

Escala breve de búsqueda de sensaciones (BSSS): estructura latente de las versiones de 8 y 4 ítems en adolescentes peruanos

Brief Sensation Seeking Scale: Latent structure of 8-item and 4-item versions in Peruvian adolescents

CÉSAR MERINO-SOTO*; EDWIN SALAS-BLAS*.

* Universidad de San Martín de Porres.

Resumen

El presente estudio tuvo el propósito de validar con adolescentes peruanos dos Escalas Breves de Búsqueda de Sensaciones: el de ocho ítems (BSSS8; Hoyle, Stephenson, Palmgreen, Lorch, y Donohew, 2002) y el de cuatro ítems (BSSS4; Stephenson, Hoyle, Slater, y Palmgreen, 2003). Los cuestionarios se aplicaron a 618 adolescentes que participaron voluntariamente, de 13.6 años de edad promedio, de diferentes niveles de estudios de la secundaria, de colegios de gestión estatal y privada, pertenecientes a un distrito del sur de Lima. Se analizó la estructura interna de ambas versiones breves a través de tres modelos: unidimensional (M1), dimensiones relacionadas u oblicuas (M2) y el modelo bifactor (M3); los resultados hallados indican que ambos instrumentos tienen una sola dimensión que representa mejor la variabilidad de los ítems, hecho que puede ser explicado tanto por la complejidad del concepto como por la pequeña cantidad de ítems que representan a cada factor; aspecto que se potencia en el BSSS4; la fiabilidad cae dentro de los niveles que los estudios anteriores hallaron (Alfa: BSSS8= .745 y BSSS4= .643) y (Coeficiente Omega: .747 del BSSS8 y .651 del BSSS4) los mismos que se consideran adecuados para el tipo de instrumentos estudiados. A partir de la correlación entre ambos instrumentos, se encontró que existen niveles satisfactorios de equivalencia entre el BSSS8 y BSSS4. Se recomienda sin embargo que el BSSS4 se utilice fundamentalmente para trabajos de investigación y con propósitos de describir poblaciones.

Palabras clave: búsqueda de sensaciones; adolescente; estructura interna; validación; fiabilidad; equivalencia.

Abstract

This research intended to validate two brief scales of sensations seeking with Peruvian adolescents: the eight item scale (BSSS8; Hoyle, Stephenson, Palmgreen, Lorch, y Donohew, 2002) and the four item scale (BSSS4; Stephenson, Hoyle, Slater, y Palmgreen, 2003). Questionnaires were administered to 618 voluntary participants, with an average age of 13.6 years, from different levels of high school, state and private school in a district in the south of Lima. It analyzed the internal structure of both short versions using three models: a) unidimensional (M1), b) oblique or related dimensions (M2), and c) the bifactor model (M3). Results show that both instruments have a single dimension which best represents the variability of the items; a fact that can be explained both by the complexity of the concept and by the small number of items representing each factor, which is more noticeable in the BSSS4. Reliability is within levels found by previous studies: alpha: .745 = BSSS8 and BSSS4 = .643; omega coefficient: .747 in BSSS8 and .651 in BSSS4. These are considered suitable for the type of instruments studied. Based on the correlation between the two instruments, it was found that there are satisfactory levels of equivalence between the BSSS8 and BSSS4. However, it is recommended that the BSSS4 is mainly used for research and for the purpose of describing populations.

Keywords: Sensation Seeking; Adolescents; Internal Structure; Validation; Reliability; Equivalence.

Recibido: Julio 2016; Aceptado: Diciembre 2016.

Enviar correspondencia a:

César Merino-Soto, Instituto de Investigación de Psicología, Universidad de San Martín de Porres, Av. Tomás Marsano 242 (5to piso), Surquillo - Lima 34, Perú; email: sikayax@yahoo.com.ar

Zuckerman (1981) empieza a investigar búsqueda de sensaciones (BS) a inicios de los 60, el primer instrumento diseñado para medirlo como constructo independiente fue el SSS (*Sensation Seeking Scale*. Zuckerman, Kolin, Price y Zoob, 1964), revisado en sucesivas versiones (II, III y IV). La última modificación, el SSS-V (*Sensation Seeking Scale*, versión V) propuesta por Zuckerman, Eysenck y Eysenck (1978), es la más usada en la investigación (Carretero-Dios y Salinas, 2008); y, ha sido adaptada en diferentes lugares y culturas como España (Pérez y Torrubia, 1986), Italia (Maná, Faraci y Como, 2013; Primi, Narducci, Benedetti, Donati y Chiesi, 2011), Canadá (Rowland y Franken, 1986) e Israel (Birenbaum y Montag, 1987), -una revisión extensa se encuentra en Aluja, García y García (2004)-. Luego, surgió otra versión (SSS-VI) publicada por el propio Zuckerman (1984) y que fue utilizada por Torki (1993) en un estudio intercultural entre poblaciones americanas y kuwaitíes.

El SSS-V ha modelado la construcción de otras medidas de BS que adoptaron parcialmente su estructura o contenidos para estudiarlo en diferentes poblaciones; uno muy conocido es el Arnett Inventory of Sensation Seeking (AISS) (Arnett 1994; Carretero-Dios y Salinas, 2008; Ferrando y Chico, 2001), que demostró medir el mismo constructo; recientemente Palacios (2015) reportó las propiedades psicométricas del Inventario de búsqueda de sensaciones para adolescentes en México (IBS-Mx), encontró ocho factores incluyendo a los cuatro reportados por Zuckerman. Otros instrumentos consideran BS como un factor de personalidad; por ejemplo el ZKPQ (*Zuckerman-Kuhlman Personality Questionnaire*, Aluja et al., 2006; Aluja, Kuhlman, y Zuckerman, 2010; Ledesma, Poó y Peltzer 2007), o, la Escala de Comportamiento Impulsivo (UPPS-P), en cuya estructura factorial se encuentra la BS (Cándido, Orduña, Perales, Verdejo-García, Billieux, 2012)

Problemas prácticos de la investigación condujeron a reducir el tamaño del instrumento sin perder confiabilidad ni validez. En este sentido, Madsen, Das, Bogen y Grossman (1987) reportaron el *Short Sensation-Seeking Scale*, derivado del SSS-IV, compuesto por 10 ítems ipsativos. Luego, Hoyle, Stephenson, Palmgreen, Lorch y Donohew (2002) reportaron el *Brief Sensation Seeking Scale* (BSSS-8), derivado del SSS-V, que demostró buenas propiedades psicométricas. Esta versión se adaptó para estudiar a trabajadores latinos en Norteamérica (Stephenson, Velez, Chalela, Ramirez y Hoyle, 2007). Posteriormente, Stephenson, Hoyle, Palmgreen y Slater (2003) crearon una más reducida de cuatro ítems (BSSS-4) para usarlo en estudios epidemiológicos o en los que BS no es el constructo principal. Sus creadores consideran que ambas versiones exhiben propiedades psicométricas estables en género y grado de instrucción, que sus puntajes se relacionan conceptualmente con las medidas más largas, y, que a pesar de su brevedad, pierden muy poco de su poder predictivo y confiabilidad.

El SSS-V, sus adaptaciones y modificaciones se concentraron en el estudio de poblaciones de adolescentes, con resultados poco satisfactorios porque se encontró que los ítems tienen bajo poder discriminativo, una estabilidad estructural poco aceptable y una débil recuperación de factores (por ejemplo, Maná et al. 2013). Hay varias razones que pueden explicar estos problemas: el uso con adolescentes de contenidos elaborados originalmente para adultos, el uso de conceptos que ya no son socialmente utilizados (Palacios, 2015) y, la obtención de medidas moderadamente válidas y las bajas confiabilidades (por ejemplo, Kafry, 1982; Pérez, Ortet, Pla y Simó, 1987; Russo, Lahey, Christ, Frick, Mc Burnnett, Loeber, Stouthamer-Loeber y Green, 1991; Russo, Stokes, Lahey, Christ., McBurnett, Loeber, Stouthamer-Loeber y Green, 1993). Evidencias que ponen límites al uso del SSS-V con poblaciones de adolescentes, pues añade varianza irrelevante al constructo, reduce la varianza común entre los ítems del mismo y no se garantiza su replicabilidad intercultural.

Para superar las críticas al SSS-V al aplicarse a adolescentes, se crearon instrumentos como el Arnett Sensation Seeking Inventory (AISS; Arnett, 1994) y otros, que consideraron las observaciones como: los contenidos de los ítems (por ejemplo, ítems que producen reacciones adversas como los referidos al uso de drogas o actividades sexuales); invalidez del constructo cuando se compara con impulsividad; la actualidad de los contenidos de varios ítems, su extensión, así como el formato de respuesta (Hoyle et al., 2002; Jensen, Weaver, Ivic, e Imboden, 2011; Palacios, 2012). Sin embargo, la estabilidad de las propiedades psicométricas del AISS con adolescentes de algunas culturas parece aún mostrar limitaciones respecto a su confiabilidad (por ejemplo, Smorti y Guarnieri, 2013), problema que se replica en estudios con poblaciones de adultos (Carretero-Dios y Salinas, 2008). Por otro lado, Stephenson, Palmgreen, Hoyle, Donohew y Colon (1999) crearon una versión de 20 ítems partiendo de dos instrumentos que fueron diseñados para adolescentes (Huba, Newcomb y Bentler, 1981; Zuckerman et al., 1978), reportando una confiabilidad por consistencia interna de 0.82 para el puntaje total; pero, el reporte sobre este instrumento mostró dos debilidades importantes: no informa la confiabilidad de las subdimensiones, ni aportó evidencias de la validez de la estructura interna del instrumento.

Otros instrumentos para niños y adolescentes surgieron con la idea de superar los problemas metodológicos del SSS-V (Michel et al., 1998; Palacios, 2015; Pérez et al., 1987; Russo et al., 1993), o como creaciones independientes pero secundarios a los objetivos principales del estudio (por ejemplo, Sargent, Tanski, Stoolmiller y Hanewinkel, 2010). Respecto a este último (Sargent et al., 2010), su brevedad y eficiencia psicométrica es comparable con otra medida breve, que es el objetivo del presente estudio (BSSS4; Stephenson et al., 2003), pero su construcción parece que

siguió un método esencialmente racional y guiado por conveniencias prácticas (brevedad) en lugar de una reestructuración completa la medición de BS y de la aplicación de análisis psicométricos multivariados.

La propuesta de medición BSSS (*Brief Sensation Seeking Scale*; Hoyle et al., 2002), trató de superar problemas relacionados con los contenidos de los ítems y con aspectos psicométricos del SSS-V (por ejemplo, Ridgeway y Russell, 1980), y parece ser el más utilizado actualmente para cuantificar la BS, pues su cobertura de edad se acomoda bien en adultos y adolescentes, sus ítems son apropiadamente relevantes a ambos grupos etarios y su contenido está relacionado con las experiencias actuales. En algunas investigaciones con adultos angloparlantes (Eachus, 2004; Litvin, 2008) y españoles (López-Bonilla y López-Bonilla, 2010) se han encontrado que posee propiedades psicométricas satisfactorias respecto a la validez interna de sus ítems, su relación con otros constructos y su consistencia interna. También, algunos resultados no publicados (e.g., Cheah, 2003) indicaron buenas propiedades psicométricas y similares al estudio original de Hoyle et al. (2002).

Hasta la fecha, los trabajos reportados con adolescentes (Banerjee, Greene y Yanovitzky, 2011; Donohew et al., 2000; Hoyle et al., 2002; Jensen, Imboden, e Ivic, 2011; Primi, Narducci, Benedetti, Donati y Chiesi, 2011; Stephenson et al., 2003) provienen principalmente de muestras anglosajonas, en estos estudios los grupos hispanos fueron identificados como minorías étnicas, cuyo estatus de origen es el de inmigrantes. Aun cuando se han construido instrumentos en español para medir BS (por ejemplo, Arnett, 1994; Palacios, 2015; Palacios, Sánchez y Andrade, 2010) que han tratado de superar los problemas del SSS-V; en Perú y en otros países sudamericanos no se han encontrado publicaciones sobre construcción, validación o análisis psicométrico de los mismos. Tampoco existen estudios publicados sobre el BSSS con adolescentes hispanos en su propia cultura.

En este estudio se presentan resultados psicométricos de la validez estructural de dos versiones breves del BSSS: de ocho (BSSS8; Hoyle et al., 2002) y de cuatro ítems (BSSS4; Stephenson et al., 2003), con muestra de adolescentes peruanos. Se justifica por su novedad, por la utilidad de las escalas breves para realizar investigaciones epidemiológicas (Stephenson et al., 2003); porque la BS está asociada con problemas sociales cada vez más amplios, como los comportamientos de riesgo, el uso problemático de alcohol y de sustancias, o el abuso de internet, de videojuegos, etc. (Cándido y Perales, 2014; Chóliz y Marco, 2011; Cortés Tomás, Giménez Costa, Motos Sellés y Cadaveira Mahía, 2014; Motos Sellés, Cortés Tomás, Giménez Costa y Cadaveira Mahía, 2015; Navas, Torres, Cándido y Perales, 2014), y también porque éstas escalas pueden usarse para predecir conductas de riesgo en diferentes actividades de la vida cotidiana, más, en contextos sociales sumamente proble-

máticos como el peruano y latinoamericano. Estas medidas breves de búsqueda de sensaciones pueden ser bien valoradas para la práctica profesional y economizar recursos en la investigación, pues los pocos estudios reportados sobre este constructo pueden estar asociados con la ausencia de instrumentos que tengan respaldo internacional, sean económicas y dimensionalmente claras. Aunque el presente estudio únicamente plantea su objetivo sobre adolescentes peruanos, también aporta con una línea base de propiedades psicométricas que son contrastables y potencialmente generalizables en otros contextos hispanos.

Método

Participantes

Fueron 618 adolescentes (mujeres: 50.6%; sin datos: 17, 2.8%) provenientes de instituciones de educación básica regular de nivel secundario, cuatro de gestión pública (495, 80.1%) y nueve de gestión privada; todos ubicados en un distrito costero al sur de Lima Metropolitana (Perú). Las instituciones educativas fueron seleccionadas por su disposición para participar en el estudio, las facilidades de acceso, y la autorización de su personal directivo con respecto al tiempo en que se desarrollaría el estudio y al cumplimiento de los aspectos éticos del mismo. Los estudiantes participantes aceptaron responder al cuestionario voluntariamente solo se incluyeron los que estuvieron presentes el día de la administración del instrumento. En las instituciones se muestrearon todos los niveles escolares (en Perú, hay cinco grados o niveles en la educación secundaria para adolescentes) para procurar la mayor potencia estadística respecto a la covarianza inter-ítem. La distribución de los estudiantes en los grados escolares fue: 1ro (238, 38.5%), 2do (107, 17.3%), 3ro (81, 13.1%), 4to (102, 16.5%) y 5to (90, 14.6%). La edad promedio fue 13.6 (DT = 1.79), y estuvo entre 10 y 21 años; esta edad extrema más alta ocurrió en uno de las instituciones públicas. No se detectaron diferencias en la distribución de la edad de acuerdo al sexo (Kolmogorov-Smirnov $Z = 0.839$, $p > 0.05$), pero sí de acuerdo a la gestión ($t[616] = 4.84$, $p = 1.58E-6$), en que los adolescentes de las instituciones privadas ($M = 14.28$, $DE = 1.52$) tuvieron moderadamente más edad ($d = 0.49$) que los de instituciones públicas ($M = 13.42$, $DE = 1.81$). Fueron excluidos los adolescentes que no aceptaron participar voluntariamente, o que no llenaron al menos el 80% de los ítems.

Instrumentos

Escala Breve de Búsqueda de Sensaciones, BSSS8 (Hoyle et al., 2002). Escala creada para adolescentes, consta de 8 ítems derivados del SSS-V que representan parsimoniosamente los cuatro factores identificados por Zuckerman para BS: búsqueda de experiencias (ítems 1 y 5), búsqueda de aventura y emociones (2 y 6), desinhibición (3 y 7), y

susceptibilidad al aburrimiento (4 y 8) (ver Apéndice). Se responde mediante un formato ordinal de respuesta de cinco opciones, desde *Muy en desacuerdo* hasta *Muy de acuerdo*. Las instrucciones de respuesta solicitan al examinado valorar sus gustos y preferencias sin referencia a un momento específico. La consistencia interna para el puntaje total en estudios previos con adolescentes es de alrededor de 0.75 (Banerjee, Greene y Yanovitzky, 2011; Donohew et al., 2000; Hoyle et al., 2002; Jensen, Imboden, e Ivic, 2011; Primi et al., 2011; Stephenson et al., 2003). La versión usada en el presente estudio fue obtenida del trabajo de Stephenson et al. (2007), en que se describen la etapa piloto de muestro de los ítems traducidos al español y las decisiones tomadas respecto a un ítem problemático para la muestra de jóvenes adultos latinos.

Escala Breve de Búsqueda de Sensaciones, BSSS4 (Stephenson et al., 2003). Es una versión super-breve, desarrollada para propósitos de ser utilizada en trabajos epidemiológicos. Fue creada mediante la selección de ítems con mayores cargas factoriales en cada dimensión en su estudio. Consta de cuatro ítems (1, 2, 7 y 8), cada uno representado a las cuatro áreas de contenido del BSSS8. En el estudio original se halló una fiabilidad de $\alpha = 0.66$, y una alta correlación ($r = .89$) con el BSSS8, del cual forman parte sus ítems. Del mismo modo, la validez de constructo con factores conductuales de riesgo y de protección se hallaron en dirección y magnitud similar al obtenido para el BSSS8. Otros estudios también han hallado similar nivel de fiabilidad ($\alpha = .65$; Vallone, Allen, Clayton, y Xiao, 2007) a los hallados por (Stephenson et al., 2003).

Procedimiento

La investigación fue aprobada por la institución de pertenencia de los investigadores, y por los directivos de las instituciones educativas de los adolescentes. Considerando que es la primera aplicación del BSSS8 en el contexto peruano, los ítems fueron examinados en un grupo pequeño de 6 adolescentes para explorar la comprensión de los ítems. En una sola entrevista semi-estructurada, los adolescentes expresaron que todos los ítems fueron percibidos como perfectamente comprensibles, respecto al contenido, extensión, opciones de respuesta e instrucciones. Posteriormente, la recolección de datos se hizo en las aulas de clase de los adolescentes, en horario regular matutino; la administración del instrumento fue supervisada por dos examinadores en cada grupo evaluado. Las instrucciones enfatizaron la respuesta honesta, anónima y concentrada en el contenido de los ítems. Todos los examinados aceptaron resolver el instrumento, luego de consentir participar.

Previo al análisis, se detectó que en cada ítem el porcentaje de valores perdidos fue menos del 1% y aparentemente aleatorio, por lo tanto, los valores ausentes se reemplazaron por el valor modal correspondiente. Respecto al análisis cuantitativo, se aplicó un análisis factorial confirma-

torio desde el modelamiento de ecuaciones estructurales (*Structural Equation Modeling, SEM*; Bentler y Dugeon, 1996; Jöreskog, 1969), mediante el cual se verificó la fuente de varianza latente de los ítems del BSSS. El método utilizado fue el de máxima verisimilitud con el escalamiento de Satorra y Bentler (1994; $SB-c^2$), pues es un procedimiento efectivo cuando ocurren distribuciones no normales de los ítems (Boomsma, 2000; Lei y Wu, 2012; Tong y Bentler, 2013), y permite aproximar mejor la prueba de bondad de ajuste a la distribución c^2 de (Bentler y Dugeon, 1996). Los análisis estructurales se basaron en la matriz de covarianzas, considerando que el número de alternativas de respuesta de los ítems (cinco), es una característica suficiente para aproximarse a variables continuas, sin que produzca sesgos sustanciales en los parámetros estimados aun usando el método maximun likelihood (Beauducel y Herzberg, 2006; Dolan, 1994; Rhemtulla, Brosseau-Liard y Savalei, 2012). Se usó la matriz de covarianzas S para el análisis (Tabla 1), estimada mediante el programa EQS 6.2 (Bentler y Wu, 2012); este programa fue utilizado para todos los análisis SEM.

En el procedimiento factorial confirmatorio, como es habitual, se impusieron especificaciones estructurales a priori (MacCallum y Austin, 2000): covariación cero entre los términos de error de los ítems y de los factores, cada ítem perteneciente a una variable latente, y el primer indicador de cada factor fue fijado con 1.0. Ya que esta especificación inicial puede requerir ser relajada durante el análisis en un marco a posteriori (Boomsma, 2000), se tomaron dos criterios para ello: uno de tipo estadístico mediante el estudio de los *índices de Lagrange* (Sörbom, 1989), llamados también *índices de modificación*; y otro de tipo racional, el mismo que tiene una base conceptual y teórica, y que se considera relativamente más importante (Boomsma, 2000; Lei y Wu, 2012) que el criterio estadístico. La cuantificación del ajuste se hizo mediante índices descriptivos, como fueron el *Comparative Fit Index* (CFI ≈ 0.95), *Standardized Root Mean Square Residual* (SRMR ≈ 0.08) y *Root Mean Square Error of Approximation* (RMSEA ≈ 0.05), con sus intervalos de confianza al 90% (McDonald, 1989). Este conjunto de índices de ajuste es recomendado para ayudarse en la toma de decisiones sobre los modelos evaluados (Jackson, Gillaspay y Purc-Stephenson, 2009). La calidad relativa del modelo también se evaluó mediante *Akaike information criterion* (AIC; Akaike, 1974).

Para verificar que las propiedades estadísticas se mantienen equivalentes entre grupos, se efectuó un análisis de la invarianza de medición (Meredith, 1993) del BSSS; el grupo clasificador fue el sexo de los adolescentes. Esto se hizo mediante el procedimiento de *análisis factorial confirmatorio de grupo múltiple* (AFCGM), en que se comparan consecutiva y acumulativamente los parámetros de los ítems, como el número de dimensiones (*invarianza configuracional*) las cargas factoriales (*invarianza débil*), interceptos (*invarianza fuerte*) y residuales (*invarianza estricta*), bajo la

hipótesis nula de igualdad de las mismas entre los grupos comparados. El AFCGM se inició con la evaluación de la invarianza configuracional o del modelo de línea base, es decir la estructura unidimensional conjuntamente verificada en varones y mujeres (invarianza configuracional). Este modelo unidimensional tuvo como referencia lo hallado previamente en la muestra total. También, se fijó con 1 la varianza de factor para permitir estimar totalmente los parámetros de los ítems. La comparación entre las diferentes formas de invarianza se hizo aplicando el criterio de Cheung y Rensvold (2002): $\Delta_{CFI} \leq 0.01$

La estimación de la fiabilidad se hizo mediante el coeficiente α (Cronbach, 1951) y sus intervalos de confianza mediante el método de Fisher (Romano, Kromrey, Owens y Scott, 2011), y el coeficiente w (McDonald, 1999). Ambos coeficientes se identifican con dos modelos diferentes de fiabilidad: tau equivalente esencial y congénico, respectivamente (Haertel, 2006), los mismos que fueron modelados por el método CFA-SEM. La precisión también se estimó usando la métrica del puntaje directo, mediante el error estándar de medición (Nunnally y Bernstein, 1995), el cual óptimamente debería ser menor que 0.5(DT) para lograr el máximo tolerable error de medición alrededor de los puntajes observados (Wyrwich, Nienaber, Tierney y Wolinsky, 1999). Asumiendo que la variabilidad del error no es necesariamente constante en los diferentes niveles de puntuación de los instrumentos de medición (Fedlt y Brennan, 1989), para el BSSS8 se examinó esta variación de error a lo largo del puntaje, mediante el error estándar de medición condicional (*Conditional Standard Error of Measurement, CSEM*; Feldt y Brennan, 1989). Éste se calculó con el método de regresión polinómica de Mollenkopf (1949),

que requiere obtener dos mitades equivalentes del instrumento. Estas mitades se formaron por el procedimiento par-impar, desde los ítems ordenados por sus respuestas promedio. El método CSEM presenta la información en la métrica del puntaje observado.

Resultados

Análisis de ítems

No hubo efectos de piso o techo en los ítems pues todas las opciones de respuesta fueron usadas por los participantes, y la dispersión de las respuestas puede considerarse similar entre ellos. Excepto el ítem 1 y 5, que muestran moderada asimetría, los ítems se acercan a la simetría distribucional, aspecto que puede verse en la magnitud del coeficiente SSI (coeficiente estandarizado de asimetría; Malgady, 2007), que varía entre 0 (simetría) a 1 (fuerte asimetría). Las respuestas a ambos ítems tienen más densidad en las opciones que indican más intensidad de BS. La curtosis mostró ser moderadamente heterogénea en los ítems.

Estructura interna y modelo de medición

Se probaron varias hipótesis de la estructura interna del BSSS: unidimensional (M_1), dimensiones relacionadas u oblicuas (M_2) y el modelo bifactor (M_3); este último modelo permite separar la varianza de los ítems en aquella relacionada con un factor común general (F_c), y otra derivada de factores específicos, F_i (Reise et al., 2010; Reise, 2012). Como se observa en la Tabla 2 respecto a BSSS8, de acuerdo a los índices de ajuste los modelos son altamente satisfactorios, aspecto que se confirma también porque la diferencia entre ellos son pequeña, y por lo tanto, estos

Tabla 1. Estadísticos descriptivos, correlaciones (Pearson) y covarianzas para los ítems.

	bsss1	bsss2	bsss3	bsss4	bsss5	bsss6	bsss7	bsss8
bsss1	1,459	0,255	0,169	0,210	0,299	0,247	0,291	0,217
bsss2	0,390	1,575	0,341	0,254	0,275	0,278	0,322	0,329
bsss3	0,272	0,559	1,726	0,264	0,265	0,221	0,363	0,267
bsss4	0,319	0,406	0,436	1,577	0,187	0,133	0,202	0,139
bsss5	0,489	0,464	0,461	0,335	1,808	0,338	0,338	0,342
bsss6	0,403	0,439	0,374	0,207	0,583	1,699	0,338	0,338
bsss7	0,469	0,521	0,628	0,331	0,592	0,497	1,736	0,439
bsss8	0,324	0,494	0,427	0,207	0,558	0,339	0,697	1,493
M	3,872	2,670	2,848	3,076	3,215	3,697	2,730	2,709
DT	1,207	1,253	1,311	1,251	1,340	1,309	1,319	1,225
As.	-0,963	0,370	0,201	-0,139	-0,126	-0,734	0,277	0,289
Cu.	0,002	-0,806	-0,989	-0,968	-1,156	-0,643	-0,998	-0,759
SSI	-0,331	0,118	0,058	-0,044	-0,035	-0,214	0,080	0,096

Nota. Todas las correlaciones son $p < 0.01$. En la diagonal y triángulo inferior, matriz de covarianzas; en el triángulo superior, correlaciones Pearson. SSI: índice estandarizado de asimetría. As.: coeficiente de asimetría. Cu.: coeficiente de curtosis.

modelos pueden considerarse adecuados. Sin embargo, comparativamente el modelo M_3 muestra el mejor ajuste, en el que los residuales entre los ítems mostraron un rango entre 0.079 y -0.019, magnitudes que pueden considerarse muy pequeñas. Todos los parámetros de los ítems en el factor general del modelo M_3 fueron estadísticamente significativos (t entre 14.08 y 7.53), mientras que los parámetros de los factores específicos no fueron estadísticamente significativos. Excepto la varianza del factor general (1.536, $t = 13.32$, $p < 0.05$), la varianza de los factores fue entre 0.00 y 0.271, y ninguna fue estadísticamente significativa ($t < 1.147$). Este modelo bi-factor claramente expone que los factores específicos (F_i) pierden poder discriminativo ante la presencia de un factor general (F_c). Efectivamente, las cargas factoriales (Tabla 3) de los factores específicos varían entre 0.000 y 0.414, mientras que en el factor general las cargas son moderadamente altas (excepto el ítem 4). Aunque el modelo M_2 también presenta un buen ajuste y es levemente bajo respecto al M_3 , debe observarse que las correlaciones interfactoriales (Tabla 3) en este modelo variaron entre 0.808 y 1.000, lo que supone una fuerte pérdida de validez discriminativa entre ellos, por lo tanto, este modelo no fue aceptado. El modelo M_1 presenta también un ajuste satisfactorio y considerando su parsimonia y la alta similitud de las cargas factoriales con M_3 , será elegido para los siguientes análisis.

En cuanto al modelamiento del BSSS4, su estructura interna exhibió un excelente ajuste. Excepto el ítem 1, el resto muestra buen nivel discriminativo (carga factorial > 0.50). Se observa también que, en general, las cargas de los ítems en el BSSS4 corresponden a los ítems que tuvieron elevadas cargas en el BSSS8, lo que sugiere que la elección

de Hoyle et al. (2002) para formar el BSSS4 (ítems con las cargas más elevadas en el BSSS8 en su estudio) se mantiene también en la presente muestra. Todos los ítems del BSSS4 fueron estadísticamente significativos ($t > 2.00$).

Una vez identificado el modelo (unidimensional), se examinó la invarianza de medición entre varones y mujeres. La invarianza configuracional arrojó que la prueba de bondad de ajuste fue estadísticamente significativa, $SB-c^2(40) = 70.78$ ($p < 0.01$). Sin embargo, los índices de ajuste pueden considerarse satisfactorios: CFI = 0.954, RMSEA = 0.050 (IC 90% = 0.030, 0.069). Habiéndose obtenido que la dimensionalidad entre los grupos es consistente, la siguiente etapa fue probar la igualdad de las cargas factoriales, en que se obtuvo $SB-c^2(48) = 81.202$ ($p < 0.01$); los indicadores de ajuste fueron CFI = 0.951, RMSEA = 0.047 (IC 90% = 0.030, 0.069), y se mantuvieron también satisfactorios. La diferencia entre ambos modelos ($D_{CFI} = 0.003$) estuvo debajo del criterio (0.01) y por lo tanto se puede mantener que ambos grupos son invariantes respecto a la métrica de sus ítems. Dado que se mantuvo este tipo de invarianza, se examinó la igualdad de los interceptos, obteniéndose $SB-c^2(56) = 94.583$ ($p < 0.01$); CFI = 0.949, RMSEA = 0.047 (IC 90% = 0.030, 0.063). La diferencia entre este modelo y el anterior fue $D_{CFI} = 0.002$, y no degradar el ajuste al introducir esta restricción adicional. Finalmente, se aplicó la restricción de igualdad de residuales entre los ítems para probar la invarianza estricta; el ajuste basado en la prueba $SB-c^2(64) = 108.625$ ($p < 0.01$), y los índices de ajuste fueron CFI = 0.944, RMSEA = 0.048 (IC 90% = 0.032, 0.062). La diferencia entre ambos modelos fue $D_{CFI} = 0.005$, por lo que se mantiene la invarianza de los residuales entre los ítems.

Tabla 2. Resultados del ajuste de la estructura interna y del modelo de medición.

Estructura interna	SB-c2 (g)	RMSEA (IC 90%)	CFI	TLI	SRMR	AIC
BSSS8						
M_1	49,170** (20)	0,049 (0,032; 0,066)	0,955	0,937	0,038	9,17
M_2	35,270** (14)	0,05 (0,030; 0,071)	0,967	0,934	0,058	7,270
M_3	26,039** (11)	0,047 (0,024; 0,071)	0,977	0,941	0,036	4,04
BSSS4						
M_1	4,564 (2)	0,046 (0,000; 0,102)	0,99	0,97	0,021	0,56
Modelo de fiabilidad (BSSS8)						
Paralelo	96,459** (33)	0,056 (0,043; 0,069)	0,915	0,910	0,068	30,45
Tau-Equivalente	77,475** (26)	0,057 (0,042; 0,071)	0,921	0,921	0,036	25,47

Nota. ** $p < 0.01$. Modelos: M_1 , unidimensional; M_2 , factores oblicuos; M_3 , bi-factor. BSSS8: Escala de 8 ítems. BSSS4: escala de 4 ítems.

Tabla 3. *Parámetros individuales del modelo bifactor y del unidimensional.*

	BSSS8						BSSS4			
	Bifactor			Unidimensional			F		h^2	
	F ₁	F ₂	F ₃	F ₄	F ₆	h^2	F	h^2	F	h^2
bsss1	0,211				0,447	0,244	0,459	0,210	0,417	0,174
bsss2		0,414			0,557	0,482	0,554	0,307	0,509	0,259
bsss3			0,295		0,510	0,347	0,520	0,271	-	-
bsss 4				0,000	0,363	0,132	0,363	0,132	-	-
bsss 5	0,250				0,555	0,371	0,562	0,316	-	-
bsss 6		0,022			0,465	0,217	0,467	0,218	-	-
bsss 7			0,128		0,637	0,422	0,644	0,414	0,679	0,462
bsss 8				0,000	0,569	0,324	0,564	0,318	0,626	0,392
Correlación interfactorial										
F1	1				-	-	-	-	-	-
F2	0,958	1			-	-	-	-	-	-
F3	0,808	0,913	1		-	-	-	-	-	-
F4	1,000	1,000	1,000	1	-	-	-	-	-	-
α	-	-	-	-	-	-	0,745	-	0,643	-
w	-	-	-	-	-	-	0,747	-	0,651	-
Estadísticos descriptivos										
M	-	-	-	-	-	-	24,77	-	11,98	-
DE	-	-	-	-	-	-	6,17	-	3,47	-
As.	-	-	-	-	-	-	-0,05	-	0,09	-
Cu.	-	-	-	-	-	-	-0,09	-	-0,27	-

Nota. h^2 : comunalidad. α : coeficiente alfa. w : coeficiente omega. BSSS8: Escala de 8 ítems. BSSS4: escala de 4 ítems.

Respecto a la evaluación del modelo de fiabilidad para el BSSS8 (Tabla 2), se encontró que el modelo tau-equivalente se ajusta levemente mejor que el modelo paralelo, pues el cambio en los índices de ajuste pueden considerarse pequeños. Pero comparado con el modelo congénico (modelo M_1), la diferencia del ajuste es notoria. Esto indica que los ítems pueden ser mejor representados por un modelo congénico, y el coeficiente de fiabilidad más exacto debería obtenerse por medio de w . Dado que los ítems del BSSS4 están contenidos en el BSSS8, entonces se asumió que los resultados de la invarianza y del modelamiento de la fiabilidad pueden ser trasladados al BSSS4.

Fiabilidad

En la Tabla 3 aparecen los coeficientes de fiabilidad α y w . Se observa que aunque el modelo congénico muestra un mejor ajuste estadístico, la significancia práctica de este ajuste no parece observarse en los coeficientes de fiabilidad calculados, por lo tanto α y w pueden considerarse muy similares. El efecto de la reducción de los ítems sobre el coeficiente α fue evaluado de dos maneras; primero, se probó la hipótesis inferencial que el coeficiente α de la for-

ma corta es igual a la de la forma larga, mediante un programa informático ad hoc (ALPHATEST; Lautenschlager y Meade, 1987), que implementa el procedimiento asintótico basado en la distribución F (Feldt, 1980; Feldt, Woodruff y Salih, 1987). Usando las correlaciones entre-formas no corregidas y corregidas por espuriedad (ver la siguiente sección), se halló una diferencia estadística entre ambos coeficientes: $\chi^2[1] = 55.79$ ($p < 0.0001$) y $\chi^2[1] = 21.97$ ($p < 0.001$), respectivamente. En segundo lugar, se comparó heurísticamente el intervalo de confianza al 95% (método Feldt et al., 1987; Romano, Kromrey, Owens y Scott, 2011) para α del BSSS8 (0.71: 0.77) y BSSS4 (0.59: 0.68). Ambos análisis muestran que la reducción de la consistencia interna en el BSSS4 es sustancial y estadísticamente significativa. La fiabilidad del BSSS8 es igual o trivialmente diferentes a los reportados en los estudios americanos de validación (Hoyle et al., 2002; Stephenson et al., 2003). Al examinar la precisión del instrumento en la métrica de los puntajes observados, se halló que el error estándar de medición en el BSSS8 (3.11) y BSSS4 (2.07) comparados con el máxima dispersión tolerable (3.08 y 1.73, respectivamente) se encuentran levemente sobre ellos.

Respecto al CSEM, la ecuación de regresión polinómica resultó en multiplicar cada puntaje por el coeficiente $b = 0.133$ ($R^2 = 0.475$, $F[2, 615] = 89.73$, $p < 0.01$), pues la constante ni el componente cuadrático y cúbico fueron estadísticamente significativos. En la Tabla 4 se muestra el CSEM y el CSEM estandarizado (CSEM-S). Se observa

Tabla 4. Error estándar de medición condicional estandarizado, para BSSS8 y BSSS4.

Puntaje	CSEM	
	BSSS8	BSSS4
4	-	1,38
5	-	1,54
6	-	1,69
7	-	1,83
8	1,06	1,95
9	1,20	2,07
10	1,33	2,18
11	1,46	2,29
12	1,60	2,39
13	1,73	2,49
14	1,86	2,58
15	2,00	2,67
16	2,13	2,76
17	2,26	2,84
18	2,39	2,64
19	2,53	2,69
20	2,66	2,74
21	2,79	-
22	2,93	-
23	3,06	-
24	1,79	-
25	1,82	-
26	1,86	-
27	1,89	-
28	1,93	-
29	1,96	-
30	2,00	-
31	2,03	-
32	2,06	-
33	2,09	-
34	2,13	-
35	2,16	-
36	2,19	-
37	2,22	-
38	2,25	-
39	2,28	-
40	2,31	-

Nota. CSEM: error estándar de medición condicional.

que los puntajes muestran menos variación por error en los niveles más bajos y cerca del promedio; y el declive se hace sustancial pues luego del puntaje 26 la fiabilidad cae debajo de 0.70.

Equivalencia BSSS8 y BSSS4

Se examinó la equivalencia y consistencia entre el BSSS8 y su versión abreviada (BSSS4). La consistencia entre las medidas, mediante su correlación entre ellas, fue corregida por espuriedad o errores correlacionados (Bashaw y Anderson 1967; Levy, 1967). Esta correlación corregida por espuriedad debería ser alta (> 0.70 ; Petrides et al., 2003; Putnam y Rothbart, 2006) para argumentar la dependencia lineal entre las formas (Smith et al., 2000). Entre las formas del BSSS, la correlación observada fue 0.89 (IC 95% = 0.87: 0.90) y la correlación corregida fue 0.68 (IC 95% = 0.63: 0.72). Con respecto al valor de referencia (0.70), la correlación corregida no fue estadísticamente significativa ($z = -0.94$, $p > 0.05$).

Para examinar la concordancia entre el BSSS8 y BSSS4, se usó el coeficiente *kappa ponderado* (K_w ; Fleiss, Levin y Paik, 2003) con pesos lineales imputados a los puntajes categorizados en deciles. Se obtuvo $K_w = 0.679$ (95% IC: 0.64, 0.70), valor que puede considerarse un buen nivel de acuerdo e igual a la correlación corregida obtenida en el párrafo anterior. Para identificar las características del acuerdo entre las versiones, éstas pueden observarse mejor en el *gráfico Bland-Altman* (1986), contrastando la diferencia BSSS8-BSSS4 contra el puntaje de la versión completa (Krouwer, 2008). La Figura 1 sugiere que, en general, las discrepancias entre BSSS4 y BSSS8 ocurren dentro de los límites aceptables (± 1.96 SD de las diferencias BSSS8-BSSS4); el aumento linealmente creciente de las diferencias se debe al aumento conjunto de los puntajes de ambas versiones y por lo tanto es una representación esperada.

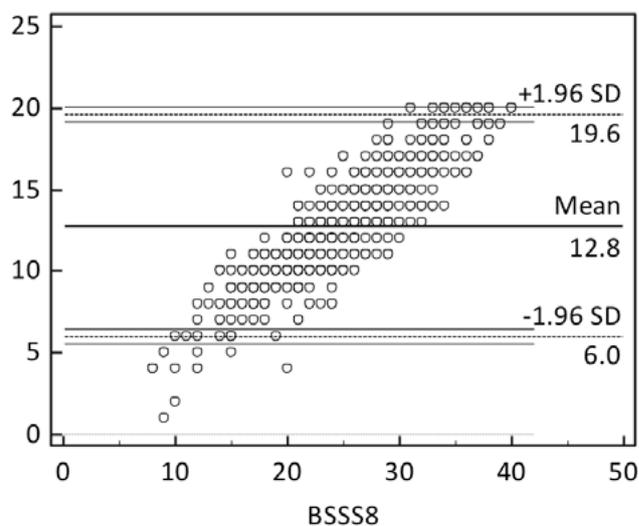


Figura 1. Gráfico Bland-Altman entre los puntajes directos de BSSS8 y BSSS4.

Sin embargo, se detecta que en los puntajes de los deciles más bajos hay más desacuerdos, aunque poco frecuentes comparado con la muestra total.

Discusión

Los resultados de la validación estructural del BSSS (8 y 4 ítems) en adolescentes peruanos arrojaron que una dimensión representa mejor la variabilidad de los ítems, este fuerte respaldo para la estructura unidimensional puede tener soporte en dos motivos: primero, las correlaciones entre los constructos originales (búsqueda de experiencias, susceptibilidad al aburrimiento, búsqueda de emociones y aventuras, y desinhibición) fueron muy elevadas, hecho que generalmente sugiere que el modelo no cumple el criterio de validez discriminativa entre las mismas (Nunnally y Bernstein, 1995). En segundo lugar, la introducción del modelo bifactor arrojó claramente que la validez de los ítems está fuertemente asociada a un factor general, donde las cargas factoriales fueron más altas que las cargas en los factores específicos. Por lo tanto, la mayor varianza común en este factor general indica inequívocamente que no existe justificación estadística para interpretar puntajes dimensionales separadamente (Reise et al., 2010; Reise, 2012). En contraste con otros estudios en los que se aplicaron métodos factoriales, este es el primero que avanza en la identificación de la estructura del BSSS mediante el modelamiento bifactor, que permitió confirmar la naturaleza unidimensional del instrumento.

Respecto a las estimaciones de consistencia interna, aquellas obtenidas en los estudios previos con el BSSS con adolescentes y adultos (citados en el presente estudio), en general no superan el 0.80, ni para los puntajes de las subescalas ni para el puntaje total. Este patrón se replica también en los resultados aquí reportados; esto puede ser consecuencia del propio origen del constructo, pues éste es considerado como un rasgo biológico y a la vez influido por la interacción con el medio social, es un constructo complejo compuesto por experiencias y sensaciones (Chico, 2000; Hoyle et al., 2002; Pérez y Torrubia, 1986; Primi et al., 2011; Stephenson et al., 2003; Zuckerman y Neeb, 1980), que agrupa y representa conductas heterogéneas con un denominador común: el riesgo; este factor puede ser relacionado conceptual y empíricamente con variables y cuestiones físicas, legales, sociales, morales, financieras, etc. (Horvath y Zuckerman, 1996; Ledesma, Poó y Peltzer, 2007). Con esta característica de base, que viene desde la definición del constructo BS y de la discusión sobre su estructura interna, es razonable que no se obtengan más altos niveles de fiabilidad. Pero, en cuanto a la comparabilidad de la consistencia interna de nuestro estudio con los estudios previos del BSSS8 y BSSS4 en adolescentes (Banerjee y Greene, 2009; Hoyle et al., 2002; Stephenson, Morgan, Lorch, Palmgreen, Donohew, y Hoyle, 2002; Stephenson

et al., 2003; Vallone et al., 2007), se obtuvieron magnitudes muy similares. Por lo tanto, existe evidencia que la variabilidad verdadera del BSSS8 y BSSS4 parece ser principalmente constante desde un ángulo intercultural. Esto tiene implicancias directas sobre la precisión de la medición de BS con estas medidas breves, que aunque no son lo suficientemente altas para uso clínico, pueden ser la mejor opción para desarrollar estudios epidemiológicos y describir grupos en un contexto de investigación científica. Por otro lado, tomando en cuenta que el error estándar de medición de los puntajes está cerca de máximo tolerable, el usuario tendrá que decidir si es adecuado aplicar un ajuste a los puntajes o estadísticos que obtenga con el instrumento, para desatenuar el de error de medición (Nunnally y Bernstein, 1995). En la interpretación de los puntajes también debe tomarse en cuenta que los puntajes debajo de la media son más fiables, y por encima hay más espacio para el error de medición. Por lo tanto, la descripción de un grupo o un sujeto en base a los puntajes del BSSS8, y más aún con el BSSS4, puede expresar menos sesgo cuando el sujeto obtiene poca intensidad de búsqueda de sensaciones en los adolescentes. En este punto no se conoce si la deseabilidad social u otro patrón irrelevante de respuesta han influenciado en este problema, pero es un asunto pendiente para resolver en una siguiente investigación.

En medidas breves como el BSSS8 y BSSS4, la medición de un constructo general puede ser preferible si lo comparamos con la medición multifactorial, debido a la cantidad de ítems en cada factor específico es reducido y repercute directamente a la magnitud de los coeficientes de fiabilidad (Heartel, 2006). Por otro lado, es conveniente precisar que en estas medidas breves de un constructo tan amplio como BS, la consistencia interna hallada puede considerarse satisfactoria para determinados contextos de uso, como los ya señalados: para investigaciones y descripciones de grupos y especialmente aquellos de tipo epidemiológico. Los niveles más exigentes de fiabilidad (por ejemplo, puntuaciones de 0.90 o superiores) de los instrumentos de medición están reservados para contextos donde se tienen que tomar decisiones sobre los sujetos individuales (por ejemplo, diagnósticos), en estas mediciones, los errores pueden ser considerados intolerables (Nunnally y Bernstein, 1995) y conducen a consecuencias que van en contra de las personas o de grupos y de instituciones (por ejemplo, un mal diagnóstico de un proceso depresivo, errores en la interpretación del clima laboral en una empresa). De acuerdo a lo argumentado, el BSSS8 y el BSSS4 pueden ser usados con seguridad en investigación o cuando se desarrollan trabajos sobre grupos de sujetos. Para evaluaciones más precisas, se requerirán una combinación de instrumentos o el uso de instrumentos con más ítems, en los cuales se pueda realizar una mejor apreciación de la consistencia interna; por ejemplo, la escala SSS-V adaptada para adolescentes (Perez, et al., 1986; Pérez, et al., 1987)

podría emplearse para ayudar al diagnóstico de este rasgo de la personalidad.

Por otro lado, otro aspecto relevante de las cargas factoriales indica que la definición del constructo en el BSSS8 y BSSS4 es similarmente ponderado por cada ítem, y que la suma simple de los ítems puede ser aceptado como un indicador del constructo. En una situación contraria, la heterogeneidad de las cargas factoriales indicaría que cada ítem influencia diferentemente en la definición del constructo, y por lo tanto una mejor representación del puntaje puede obtenerse ponderando cada ítem diferencialmente. Afortunadamente, la similaridad del poder discriminativo se cumple en cada ítem, y aunque no son muy altos (por ejemplo, 0.80 o más), contribuyen con suficiente varianza relevante al constructo para propósitos descriptivos de grupos. Esta descripción puede usar los mismos parámetros de medición hallados entre varones y mujeres, pues el instrumento ha cumplido con las formas de invarianza cada vez más exigentes. Por lo tanto, una comparación exhaustiva de las diferencias en el nivel de puntajes latentes u observados parece garantizado.

Finalmente, la equivalencia analizada entre el BSSS8 y BSSS4 ha sido satisfactoria, y puede considerarse más precisa respecto a la correlación convergente entre ambas, pues los autores (Hoyle et al., 2002; Stephenson et al., 2003; Stephenson et al., 2007) no ajustaron el coeficiente de correlación para controlar en efecto de la varianza de errores correlacionados. La escala BSSS4 puede ser utilizada para estudios epidemiológicos o en investigaciones junto a otros instrumentos, especialmente cuando la cantidad de ítems, los límites de tiempo y la disposición del evaluado interaccionan seriamente para requerir medidas abreviadas pero válidas.

El presente estudio presenta algunas limitaciones como la ausencia de evaluación de la invarianza de medición o el funcionamiento diferencial de ítems, la caracterización de los ítems mediante el modelado desde la Teoría de Respuesta al Ítem, y las relaciones externas con otros constructos convergentes y divergentes. Estos aspectos se encuentran en la agenda de posteriores investigaciones.

Reconocimientos

Los autores agradecen a la Universidad de San Martín de Porres, por las facilidades para la realización de este estudio, así como a las instituciones educativas que les dieron acceso. También agradecemos las observaciones de los revisores.

Conflictos de interés

Los autores informan que no hay conflictos de interés en, durante la concepción y desarrollo del presente manuscrito.

Referencias

- Akaike, H. (1974). A new look at statistical model identification. *IEEE Transactions on Automatic Control*, *19*, 716–723. doi:10.1007/978-1-4612-1694-0_16.
- Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.
- Aluja, A., Kuhlman, M. y Zuckerman, M. (2010). Development of the Zuckerman-Kuhlman-Aluja Personality Questionnaire (ZKA-PQ): A factor/facet version of the Zuckerman-Kuhlman Personality Questionnaire (ZKPQ). *Journal of Personality Assessment*, *92*, 416-431. doi:10.1037/t03666-000.
- Aluja, A., Rossier, J., García L. F., Angleitner, A., Kuhlman, M. y Zuckerman, M. (2006). A cross-cultural shortened form of the ZKPQ (ZKPQ-50-CC) adapted to English, French, German, and Spanish languages. *Personality and Individual Differences*, *41*, 619-628. doi:10.1016/j.paid.2006.03.001.
- Anderson, D. R., Burnham, K. P. y White, G. C. (1998). Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. *Journal of Applied Statistics*, *25*, 263-282. doi:10.1080/02664769823250.
- Arnett, J. (1994). Sensation seeking: A new conceptualization and a new scale. *Personality and Individual Differences*, *16*, 289-296. doi:10.1016/0191-8869(94)90165-1.
- Banerjee, S. C., Greene, K. y Yanovitzky, I. (2011). Sensation seeking and dosage effect: An exploration of the role of surprise in anti-cocaine messages. *Journal of Substance Use*, *16*, 1–13. doi:10.3109/14659890903490773.
- Bashaw, W. L. y Anderson, H. E. (1967). A correction for replicated error in correlation coefficients. *Psychometrika*, *32*, 435-441. doi:10.1007/bf02289657.
- Beauducel, A. y Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, *13*, 186–203. doi:10.1207/s15328007sem1302_2.
- Bentler, P. M. y Dudgeon, P. (1996). Covariance structure Analysis: Statistical practice, theory, and directions. *Annual Review of Psychology*, *47*, 563-592. doi:10.1146/annurev.psych.47.1.563.
- Bentler, P. M. y Wu, E. J. C. (2012). *EQS for Windows*. Encino CA: Multivariate Software, Inc.
- Birenbaum, M. y Montag, I. (1987). On the replicability of the factorial structure of the sensation Seeking Scale. *Personality and Individual Differences*, *8*, 403-408. doi:10.1016/0191-8869(87)90041-9.
- Bland, J. M. y Altman, D. G. (1986). Statistical method for assessing agreement between two methods of clinical measurement. *The Lancet*, *1*, 307-310. doi:10.1016/s0140-6736(86)90837-8.

- Boomsma, A. (2000). Reporting analyses of covariance structures. *Structural Equation Modeling*, 7, 461–483. doi:10.1207/s15328007sem0703_6.
- Cándido, A., Orduña, E., Perales, J.C., Verdejo-García, A. y Billieux, J. (2012). Validation of a short Spanish version of the UPPS-P impulsive behaviour scale. *Trastornos Adictivos*, 14, 73-78. doi:10.1016/s1575-0973(12)70048-x.
- Carretero-Dios, H. y Salinas, J. (2008). Using a structural equation model to assess the ambivalence between assessment instruments: The dimension of sensation seeking as measured by Zucherman's SSS-V and Arnett's AISS. *International Journal of Clinical and Health Psychology*, 8, 219-232. doi:10.1109/cisti.2014.6876961.
- Cheah, W. H. (2003). *The impact of sensation seeking and cultural orientation on the effects of fear appeal messages*. Unpublished Dissertation, University of Kentucky, Kentucky.
- Chen, X., Li, F., Nydegger, L., Gong, J., Ren, Y., Dinaj-Koci, V., ... Stanton, B. (2013). Brief Sensation Seeking Scale for Chinese - Cultural adaptation and psychometric assessment. *Personality and Individual Differences*, 54, 604–609. doi:10.1037/t26390-000.
- Cheung, G. W. y Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233-255. doi:10.1207/s15328007sem0902_5.
- Chóliz, M. y Marco, C. (2011). Patrón de uso y dependencia de videojuegos en infancia y adolescencia. *Anales de Psicología*, 27, 418-426.
- Cortés Tomás, M. T., Giménez Costa, J. A., Motos Sellés, P. y Cadaveira Mahía, F. (2014). The importance of expectations in the relationship between impulsivity and binge drinking among university students. *Adicciones*, 26, 134-145.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. doi:10.1007/bf02310555.
- Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5, and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology*, 47, 309–326. doi:10.1111/j.2044-8317.1994.tb01039.x.
- Donohew, L., Zimmerman, R., Cupp, P. S., Novak, S., Colon, S. y Abell, R. (2000). Sensation seeking, impulsive decision-making, and risky sex: Implications for risk-taking and design of interventions. *Personality and Individual Differences*, 28, 1079-1091. doi:10.1016/s0191-8869(99)00158-0.
- Eachus, P. (2004). Using the Brief Sensation Seeking Scale (BSSS) to predict holiday preferences. *Personality and Individual Differences*, 36, 141-153. doi:10.1016/s0191-8869(03)00074-6.
- Fan, H., Lin, M., Bai, C., Huang, P., Chiang, Y. y Chui, W. (2014). Validation of the Chinese-language brief sensation seeking scale: Implications for risky riding behaviors of parental motorcyclists and their child passengers. *Accident Analysis and Prevention*, 73, 333-339. doi:org/10.1016/j.aap.2014.09.015.
- Feldt, L. S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. *Psychometrika*, 45, 99-105. doi:org/10.1007/bf02293600.
- Feldt, L. S. y Brennan, R. L. (1989). Reliability. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 105-146). New York: American Council on Education/Macmillan.
- Ferrando, P. J. y Chico, E. (2001). The construct of sensation seeking as measured by Zuckerman's SSS-V and Arnett's AISS: A structural equation model. *Personality and Individual Differences*, 31, 1121-1133. doi:10.1016/s0191-8869(00)00208-7.
- Fleiss, J. L., Levin, B. y Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed). Hoboken: John Wiley y Sons.
- Haertel, E. H. (2006). Reliability. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 65–110). Westport, CT: American Council on Education/Praeger
- Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P. y Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences*, 32, 401–414. doi:10.1016/s0191-8869(01)00032-0.
- Huba, G. J., Newcomb, M. D. y Bentler, P. M. (1981). Comparison of canonical correlation and interbattery factor analysis on sensation seeking and drug use domains. *Applied Psychological Measurement*, 5, 291-306. doi:10.1177/014662168100500302.
- Jackson, D. L., Gillaspay, J. A. Jr. y Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: an overview and some recommendations. *Psychological Methods*, 14, 6–23. doi:org/10.1037/a0014694.
- Jensen, J., Imboden, K. y Ivic, R. (2011). Sensation seeking and narrative transportation: high sensation seeking children's interest in reading outside of school. *Scientific Studies of Reading*, 15, 541-558. doi:10.1080/10888438.2010.528819.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34, 183–202. https://doi.org/10.1007/bf02289343
- Kafry, D. (1982). Sensation seeking of young children. *Personality and Individual Differences*, 3, 161-166. doi:10.1016/0191-8869(82)90030-7.
- Krouwer, J. S. (2008). Why Bland-Altman plots should use X , not $(Y+X)/2$ when X is a reference method. *Statistics in Medicine*, 27, 778-780.
- Lautenschlager, G. J. y Meade, A. W. (2008). AlphaTest: a windows program for tests of hypotheses about coefficient alpha. *Applied Psychological Measurement*, 32, 502–503. doi:10.1177/0146621607312307.

- Lei, P.-W. y Wu, Q. (2012). Estimation in structural equation modeling. In: Hoyle, R. H. (Ed.), *Handbook of structural equation modeling* (pp. 164–179). New York: Guilford Press.
- Levy, P. (1967). The correction for spurious correlation in the evaluation of short-form tests. *Journal of Clinical Psychology*, 23, 84-86. doi:10.1002/1097-4679(196701)23:1<84::aid-jclp2270230123>3.0.co;2-2.
- Litvin, S. W. (2008). Sensation seeking and its measurement for tourism research. *Journal of Travel Research*, 46, 440-445. doi:10.1177/0047287507308326.
- López-Bonilla, J. y López-Bonilla, L. (2010). Sensation seeking and the use of the internet: A confirmatory analysis of the Brief Sensation Seeking Scale (BSSS) by gender. *Social Science Computer Review*, 28, 177-193. doi:10.1177/0894439309341229.
- MacCallum, R. C. y Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology* 51, 201–226. doi:10.1146/annurev.psych.51.1.201.
- Malgady, R. G. (2007). How skewed are psychological data? A standardized index of effect size. *The Journal of General Psychology*, 134, 355–359.
- Maná, G., Faraci, P. y Como, M. R. (2013). Factorial structure and psychometric properties of the Sensation Seeking Scale – Form V (SSS–V) in a sample of Italian adolescents. *European Journal of Psychology*, 9, 276–288. doi:10.5964/ejop.v9i2.500.
- McDonald, R. P. (1999). *Test theory: A unified approach*. Mahwah, NJ: Erlbaum.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58, 525-43. doi:10.1007/bf02294825.
- Michel, G., Mourén-Simeoni, M., Perez-Diaz, F., Falissard, B., Carton, S. y Jouvent, R. (1998). Construction and validation of a sensation seeking scale for adolescents. *Personality and Individual Differences*, 26, 159-174. doi:10.1016/s0191-8869(98)00059-2.
- Mollenkopf, W. G. (1949). Variation of the standard error of measurement. *Psychometrika*, 14, 189-229. doi:10.1007/bf02289153.
- Motos Sellés, P., Cortés Tomás, M. T., Giménez Costa, J. A. y Cadaveira Mahía, F. (2015). Predictores del consumo semanal de alcohol y sus consecuencias asociadas en universitarios consumidores intensivos de alcohol. *Adicciones*, 27, 119-131. doi:10.20882/adicciones.700.
- Navas, J. E., Torres, A., Cándido, A. y Perales, J. C. (2014). ¿“Nada” o “un poco”? ¿“mucho” o “demasiado”? La impulsividad como marcador de gravedad en niveles problemático y no problemático de uso de alcohol e internet. *Adicciones*, 26, 159-167. doi:10.20882/adicciones.19.
- Nunnally, J. C. y Bernstein, I. H. (1995). *Teoría psicométrica*. México, DF: McGraw-Hill.
- Feldt, L. S., Woodruff, D. J. y Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement*, 11, 93-103. doi:10.1177/014662168701100107.
- Palacios, J. R. (2015). Propiedades psicométricas del inventario de búsqueda de sensaciones para adolescentes en México (IBS-Mx). *International Journal of Psychological Research*, 8, 40-46. doi:10.21500/20112084.644.
- Palacios, J., Sánchez, B. y Andrade, P. (2010). Intento de suicidio y búsqueda de sensaciones en adolescentes. *Revista Intercontinental de Psicología y Educación*, 12, 53-75.
- Pérez, J. y Turrubia, R. (1986). Fiabilidad y Validez de la versión española de la escala de búsqueda de sensaciones (forma V). *Revista Latinoamericana de Psicología*, 18, 7-22.
- Pérez, J., Ortet, G., Plá, S. y Simó, S. (1987). Escala de Búsqueda de Sensaciones para Niños y Adolescentes (EBS-J). *Evaluación Psicológica*, 3, 283-290.
- Perez, J., Ortet, G., Pla, S. y Simo, S. (1986). A Junior Sensation Seeking Scale (J-SSS). *Personality and Individual Differences*, 7, 915-918. doi:10.1016/0191-8869(86)90095-4.
- Petrides, K. V., Jackson, C. J., Furnham, A. y Levine, S. Z. (2003). Exploring issues of personality measurement and structure through the development of a short form of the Eysenck personality profiler. *Journal of Personality Assessment*, 81, 271-280. doi:10.1207/s15327752jpa8103_10.
- Primi, C., Narducci, R., Benedetti, D., Donati, M. y Chiesi, F. (2011). Validity and reliability of the Italian version of the Brief Sensation Seeking Scale (BSSS) and its invariance across age and gender. *Testing, Psychometrics, Methodology in Applied Psychology*, 18, 231-241.
- Putnam, S. P. y Rothbart, M. K. (2006). Development of short and very short forms of the Children’s Behavior Questionnaire. *Journal of Personality Assessment*, 87, 103-133.
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47, 667-696. doi:10.1207/s15327752jpa8701_09.
- Reise, S. P., Moore, T. M. y Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment*, 92, 544-559. doi:10.1080/00223891.2010.496477.
- Rhemtulla, M., Brosseau-Liard, P. E. y Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods*, 17, 354–373.
- Ridgeway, D. y Russell, J. A. (1980). Reliability and validity of the Sensation-Seeking Scale: Psychometric problems in Form V. *Journal of Consulting and Clinical Psychology*, 48, 662-664. doi:10.1037/a0029315.
- Romano, J. L., Kromrey, J. D., Owens, C. M. y Scott, H. M. (2011). Confidence interval methods for coefficient alpha

- on the basis of discrete, ordinal response items: Which one, if any, is the best? *The Journal of Experimental Education*, 79, 382-403. doi:10.1080/00220973.2010.510859.
- Rowland G. L. y Franken R. E. (1986). The four dimensions of sensation seeking: a confirmatory factor analysis. *Personality and Individual Differences*, 7, 237-240. doi:10.1016/0191-8869(86)90062-0.
- Russo, M. F., Lahey, B. B., Christ., M. A. G., Frick, P. J., McBurnett, K., Walker, J. L., ... Green, S. M. (1991). Preliminary development of a sensation seeking scale for children. *Personality and Individual Differences*, 12, 399-405. doi:10.1016/0191-8869(91)90056-h.
- Russo, M. F., Stokes, G. S., Lahey, B. B., Christ, M. A. G., McBurnett, K., Loeber, R., ... Green, S. M. (1993). A Sensation Seeking Scale for Children: further refinement and psychometric development. *Journal of Psychopathological Behavior Assessment*, 15, 69-86. doi:10.1007/bf00960609.
- Smith, G. T., McCarthy, D. M. y Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12, 102-111. doi:10.1037//1040-3590.12.1.102.
- Smorti, M. y Guarneri, S. (2013). A study on the validity of the Arnett Inventory of Sensation Seeking (AISS) in an Italian adolescent sample. *International Journal of Advances in Psychology*, 2, 10-17.
- Sörbom, D. (1989). Model modification. *Psychometrika*, 54, 371-384. doi:10.1007/bf02294623.
- Stephenson, M. T., Hoyle, R. H., Slater, M. D. y Palmgreen, P. (2003). Brief measures of sensation seeking for screening and large-scale surveys. *Drug and Alcohol Dependence*, 72, 279-286. doi:10.1016/j.drugalcdep.2003.08.003.
- Stephenson, M. T., Morgan, S. E., Lorch, E. P., Palmgreen, P., Donohew, L. y Hoyle, R. (2002). Predictors of exposure from an antimarijuana media campaign: outcome research assessing sensation seeking targeting. *Health Communication*, 14, 23-43. doi:10.1207/s15327027hc1401_2.
- Stephenson, M. T., Palmgreen, P., Hoyle, R. H., Donohew, L., Lorch, E. P. y Colon, S. (1999). Short-term effects of an anti-marijuana media campaign targeting high sensation seeking adolescents. *Journal of Applied Communication Research*, 27, 175-195. doi:10.1080/00909889909365535.
- Stephenson, M. T., Velez, L. F., Chalela, P., Ramirez, A. y Hoyle, R. H. (2007). The reliability and validity of the Brief Sensation Seeking Scale (BSSS-8) with young adult Latino workers: implications for tobacco and alcohol disparity research. *Addiction*, 102, 79-91. doi:10.1111/j.1360-0443.2007.01958.x.
- Tong, X. y Bentler, P. M. (2013). Evaluation of a new mean scaled and moment adjusted test statistic for SEM. *Structural Equations Modeling*, 20, 148-156. doi:10.1080/10705511.2013.742403.
- Wyrwich, K. W. (2004). Minimal important difference thresholds and the standard error of measurement: is there a connection? *Journal of Biopharmaceutical Statistical*, 14, 97-110. doi:10.1081/bip-120028508.
- Zuckerman, M. y Neeb, M. (1980). Demographic influences in sensation seeking and expressions of sensations seeking in religion, smoking and driving habits. *Personality and Individual Differences*, 1, 197-206. doi:10.1016/0191-8869(80)90051-3.
- Zuckerman, M. (2007). The Sensation Seeking Scale V (SSS-V): Still reliable and valid. *Personality and Individual Differences*, 43, 1303-1305. doi:10.1016/j.paid.2007.03.021.
- Zuckerman, M., Eysenck, S. B. y Eysenck, H. J. (1978). Sensation seeking in England and American: cross-cultural age and sex comparisons. *Journal of Consulting and Clinical Psychology*, 46, 139-149. doi:10.1037/0022-006x.46.1.139.