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

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

César Merino-Soto*; Edwin Salas-Blas*.

* Universidad de San Martín de Porres.

Abstract

This research intended to validate two brief scales of sensations seeking with Peruvian adolescents: the eight item scale (BSSS8; Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002) and the four item scale (BSSS4; Stephenson, Hoyle, Slater, & 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.

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, & Donohew, 2002) y el de cuatro ítems (BSSS4; Stephenson, Hoyle, Slater, & 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.

Received: July 2016; Accepted: December 2016.

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

Zuckerman (1981) began to investigate sensation seeking (SS) in the early 1960's and the first instrument designed to measure it as an independent construct was the SSS (Sensation Seeking Scale, Zuckerman, Kolin, Price & Zoob, 1964), revised in successive versions (II, III and IV). The last modification, the SSS-V (Sensation Seeking Scale, version V) proposed by Zuckerman, Eysenck and Eysenck (1978), is the most frequently used in research (Carretero-Dios & Salinas, 2008). It has been adapted for use in different places and cultures such as Spain (Pérez & Torrubia, 1986), Italy (Maná, Faraci & Como, 2013; Primi, Narducci, Benedetti, Donati & Chiesi, 2011), Canada (Rowland & Franken, 1986) and Israel (Birenbaum & Montag, 1987). An extensive review can be found in Aluja, García and García (2004). A further version (SSS-VI) was later published by Zuckerman (1984) himself and used by Torki (1993) in an intercultural study of American and Kuwaiti populations.

The SSS-V has modeled the construction of other SS measures that partially adopted its structure or contents in order to study them in different populations. A well-known examples is the Arnett Inventory of Sensation Seeking (AISS) (Arnett 1994; Carretero-Dios & Salinas, 2008; Ferrando & Chico, 2001), which demonstrated that the same construct was being measured. Recently, Palacios (2015) reported the psychometric properties of the Sensation Search Inventory for adolescents in Mexico (IBS-Mx) and found eight factors, including the four reported by Zuckerman. Other instruments consider SS to be a personality factor, such as the ZKPQ (*Zuckerman-Kuhlman Personality Questionnaire*, Aluja et al., 2006; Aluja, Kuhlman, & Zuckerman, 2010; Ledesma, Poó & Peltzer 2007), or the Impulsive Behavior Scale (UPPS-P), the factorial structure of which covers SS (Candido, Orduña, Perales, Verdejo-García, Billieux, 2012).

Practical research problems led to the reduction in size of the instrument without loss of reliability or validity, with Madsen, Das, Bogen and Grossman (1987) developing the Short Sensation-Seeking Scale, derived from SSS-IV and composed of 10 ipsative items. Later, Hoyle, Stephenson, Palmgreen, Lorch and Donohew (2002) presented the Brief Sensation Seeking Scale (BSSS-8), derived from SSS-V and demonstrating good psychometric properties. This version was adapted to study Latino workers in North America (Stephenson, Velez, Chalela, Ramirez & Hoyle, 2007). Subsequently, Stephenson, Hoyle, Palmgreen and Slater (2003) created a smaller scale of four items (BSSS-4) for use in epidemiological studies or those in which SS is not the main construct. Its creators consider that both versions exhibit stable psychometric properties regarding gender and educational level, that their scores are conceptually related to those of more longer measures, and that despite the scale's brevity, they lose very little of their predictive power and reliability.

The SSS-V, its adaptations and modifications were centred on the study of adolescent populations, with unsatisfactory results given that the discriminatory power of the items was found to be low, structural stability unacceptable, and recovery of factors weak (e.g. Maná et al., 2013). Several reasons may underlie these problems: using content originally developed for adults with adolescents, using concepts no longer socially relevant (Palacios, 2015), and the obtainment of moderately valid measures and low reliabilities (for example, Kafry, 1982; Pérez, Ortet, Pla & Simó, 1987; Russo, Lahey, Christ, Frick, McBurnnett, Loeber, Stouthamer-Loeber & Green, 1991; Russo, Stokes, Lahey, Christ, McBurnett, Loeber, Stouthamer-Loeber & Green, 1993). These results limit the use of SSS-V with adolescent populations because it adds irrelevant variance to the construct, reduces the common variance among its items, and does not guarantee its intercultural replicability.

In order to overcome the criticism of SSS-V when applied to adolescents, instruments such as the Arnett Sensation Seeking Inventory (AISS, Arnett, 1994) and others were developed, which contemplated, for example, the contents of the items (e.g. items that produce adverse reactions, such as those related to drug use or sexual activity), the invalidity of the construct when compared to impulsivity, the up-to-dateness of various item contents , the size of the instrument as well as the response format (Hoyle, et al., 2002; Jensen, Weaver, Ivic, & Imboden, 2011; Palacios, 2012). Nevertheless, the stability of the psychometric properties of the AISS with adolescents of certain cultures still seems to be limitated in terms of reliability (e.g. Smorti & Guarnieri, 2013), a problem which is repeated in studies with adult populations (Carter-Dios & Salinas, 2008). In addition, Stephenson, Palmgreen, Hoyle, Donohew and Colon (1999) created a version with 20 items starting from two instruments that were designed for adolescents (Huba, Newcomb & Bentler, 1981; Zuckerman et al., 1978), with internal consistency reliability of 0.82 for the total score. However, the report on this instrument had two important weaknesses: it did not report the reliability of the sub-dimensions, nor did it provide evidence of the validity of the internal structure of the instrument.

Some other instruments for children and adolescents also emerged with the idea of overcoming the methodological problems of SSS-V (Michel et al., 1998; Palacios, 2015; Pérez et al., 1987; Russo et al., 1993), or as independent developments but secondary to the main objectives of the study (e.g. Sargent, Tanski, Stoolmiller & Hanewinkel, 2010). With respect to the latter (Sargent et al., 2010), its brevity and psychometric efficiency is comparable to another short measure which is the focus of the present study (BSSS4; Stephenson et al., 2003). Its construction, however, appears to have followed an essentially rational method guided by practical convenience (brevity) instead of being a complete restructuring of the

measurement of SS and the application of multivariate psychometric analyses.

The proposed BSSS (Brief Sensation Seeking Scale, Hoyle et al., 2002), attempted to overcome problems related to the contents of the items and with psychometric aspects of SSS-V (e.g. Ridgeway & Russell, 1980), and currently seems to be the most used to quantify SS, since its age coverage fits well with adults and adolescents, its items are appropriate and relevant to both age groups and its content is related to current experiences. In some studies with adult speakers of English (Eachus, 2004; Litvin, 2008) and Spanish (López-Bonilla & López-Bonilla, 2010) it was found to have satisfactory psychometric properties regarding the internal validity of its items, its relationship to other constructs and its internal consistency. Also, some unpublished results (e.g. Cheah, 2003) indicated good psychometric properties which were similar to the original study by Hoyle et al. (2002).

To date, research carried out with adolescents (Banerjee, Greene & Yanovitzky, 2011; Donohew et al., 2000; Hoyle et al., 2002; Jensen, Imboden, & Ivic, 2011; Primi, Narducci, Benedetti, Donati & Chiesi, 2011; Stephenson et al., 2003) has been mainly with Anglo-Saxon samples. In these studies Hispanic groups were identified as ethnic minorities, whose origin status is that of immigrants. Even when the instruments were created in Spanish to measure SS (e.g. Arnett, 1994; Palacios, 2015; Palacios, Sánchez & Andrade, 2010) in an attempt to overcome the shortcomings of SSS-V, no publications have been found in Peru and other South American countries describing their construction and validation or provding psychometric analysis. There are also no published studies on BSSS with Hispanic adolescents in their own culture.

In this study, we present psychometric results with structural validity of two short versions of BSSS, of eight (BSSS8, Hoyle et al., 2002) and four items (BSSS4, Stephenson et al., 2003), with a sample of Peruvian adolescents. It is justified by its novelty and by the utility of brief scales for epidemiological studies (Stephenson et al., 2003) because SS is associated with increasingly widespread social problems, such as risky behaviors, problematic use of alcohol and substances, or abuse of the internet, video games, etc. (Cándido & Perales, 2014; Choliz & Marco, 2011; Cortés Tomás, Giménez Costa, Motos Sellés & Cadaveira Mahía, 2014; Motos Sellés, Cortés Tomás, Giménez Costa & Cadaveira Mahía, 2015; Navas, Torres, Candido & Perales, 2014), and also because these scales can be used to predict risky behaviors in different activities of daily life, and moreover in highly problematic social contexts such as Peru and Latin America. These brief measures of sensation seeking can be very valuable for professional practice and to rationalize research resources, since the small number of studies on this construct can be linked to the absence of instruments that have international support, are economical and dimensionally clear. Although the present study only focuses on Peruvian adolescents, it also contributes a baseline of psychometric properties that are testable and potentially generalizable to other Hispanic contexts.

Method

Participants

The sample consisted of 618 adolescents (females: 50.6%, without data: 17, 2.8%) coming from regular basic secondary schools, four of which were state managed (495, 80.1%) and nine prívate. All were located in a coastal district south of Lima's metropolitan area in Peru. The educational institutions were selected for their willingness to participate in the study, ease of access, and the authorization of their management regarding the time when the study could be carried out and compliance with the ethical aspects of the study. Participating students agreed to respond to the questionnaire voluntarily and only those that were present on the day the instrument was administered were included. All school grades were sampled (in Peru, secondary education consists of five grades or years) in order to obtain the highest statistical power with respect to inter-item covariance. The distribution of students in the school grades was: 1st (238, 38.5%), 2nd (107, 17.3%), 3rd (81, 13.1%), 4th (102, 16.5%) and 5th (90, 14.6%) . The mean age was 13.6 (SD = 1.79), and the age range was from 10 to 21. This extreme máximum age occurred in one of the public schools. While there were no differences in age distribution according to sex (Kolmogorov-Smirnov $Z = 0.839$, $p > 0.05$), differences were found according to school management ($t[616] = 4.84$, $p = 1.58E\text{-}6$), where adolescents from private institutions (M = 14.28, SD = 1.52) were moderately older ($d = 0.49$) than those from public schools (M = 13.42, SD = 1.81). Adolescents who did not agree to participate voluntarily or who did not complete at least 80% of the items were excluded.

Instruments

Brief Sensation Seeking Scale, BSSS8 (Hoyle et al., 2002). This scale was created for adolescents and consists of 8 items derived from the SSS-V that parsimoniously represent the four factors identified by Zuckerman for SS: the experience seeking (items 1 and 5), adventure and emotion seeking (2 and 6), disinhibition (3 and 7), and susceptibility to boredom (4 and 8) (see Appendix). It has an ordinal response format of five options, from *disagree completely* to *agree completely*. The respondent is instructed to assess their likes and preferences without reference to a specific moment. The internal consistency for the total score in previous studies with adolescents was around 0.75 (Banerjee, Greene & Yanovitzky, 2011; Donohew et al., 2000; Hoyle et al., 2002; Jensen, Imboden, & Ivic, 2011; Primi et al., 2011; Stephenson et al., 2003). The version

used in the present study was obtained from the work of Stephenson et al. (2007), which describes the pilot stage of sampling items translated into Spanish and the decisions taken regarding a problematic item for samples of young Latino adults.

Brief Sensation Seeking Scale, BSSS4 (Stephenson et al., 2003). This is a super-short version, developed for purposes of use in epidemiological work. It was created by selecting items with higher factor loads in the study of each dimension. It consists of four items (1, 2, 7 and 8), each representative of the four content areas of the BSSS8. The original study yielded a reliability of $\alpha = 0.66$, and a high correlation ($r = .89$) with the BSSS8, of which the items form part. Similarly, construct validity with behavioral risk and protective factors were found which were similar in direction and magnitude to those obtained for the BSSS8. Other studies have also yielded similar levels of reliability ($\alpha = .65$; Vallone, Allen, Clayton, & Xiao, 2007) to those found by Stephenson et al., 2003.

Procedure

The research was approved by the institution to which the researchers belonged, and by the directors of the adolescents' educational establishments. Given that it was the first application of BSSS8 in a Peruvian context, the items were examined in a small group of 6 adolescents to explore how the items were understood. In a single semi-structured interview, the adolescents stated that all items were perceived as perfectly comprehensible regarding content, extent, response options and instructions. Data collection was subsequently carried out in classrooms during regular morning hours. The administration of the instrument was supervised by two researchers for each group assessed. The instructions placed emphasis on giving answers that were honest, anonymous and focused on the content of the items. All the respondents agreed to complete the quesionnaire after consenting to participate.

Prior to the analysis, it was found that for each item the percentage of missing values was below 1% and apparently random. The missing values were therefore replaced by the corresponding modal value. With regard to the quantitative analysis, a confirmatory factor analysis was applied based on structural equation modeling (SEM, Bentler & Dugeon, 1996; Jöreskog, 1969), which verified the source of latent variance of BSSS items. The method used was the one of maximum verisimilitude with Satorra and Bentler scaling (1994; *SB-c*^{2}), since this is an effective procedure with non-normal distributions of the items (Boomsma, 2000; Lei & Wu, 2012; Tong & Bentler, 2013) and allows better approximation of the goodness-of-fit test to the c^2 distribution of Bentler and Dugeon, 1996. The structural analyses were based on the covariance matrix, given that the number of response alternatives of the items (five) is a sufficient characteristic for approximating to continuous variables without producing substantial biases in the estimated parameters even when using the máximum likelihood method (Beauducel & Herzberg, 2006; Dolan, 1994; Rhemtulla, Brosseau-Liard & Savalei, 2012). The covariance matrix S was used for the analysis (Table 1), estimated using the EQS 6.2 program (Bentler & Wu, 2012), which was used for all SEM analyses.

As is the norm in confirmatory factor analysis, structural specifications were imposed a priori (MacCallum & Austin, 2000): zero covariance between the item and factor error terms, each item belonging to a latent variable, and the first indicator of each factor was set at 1.0. Since this initial specification may require some flexibility during analysis in a posteriori framework (Boomsma, 2000), two criteria were established for this: one of a statistical nature through the study of the Lagrange indices (Sörbom, 1989), also known as modification indices; and one of a rational nature, which is the same with a conceptual and theoretical basis, and which is considered relatively more important (Boomsma, 2000; Lei & Wu, 2012) than the statistical criterion. The quantification of the goodness of fit was made using descriptive indices such as the comparative fit index (CFI ^{3} 0.95), standardized root mean square residual (SRMR £ 0.08) and root mean square error of approximation (RMSEA £ 0.05), with their confidence intervals set at 90% (McDonald, 1989). This set of indices for goodness of fit is recommended in assisting decision making on the models evaluated (Jackson, Gillaspay & Purc-Stephenson, 2009). The relative quality of the model was also evaluated with the Akaike information criterion (AIC Akaike, 1974).

In order to verify that the statistical properties were maintained across groups, an analysis of the measurement invariance (Meredith, 1993) of the BSSS was carried out. The classifying group was the sex of the adolescents. This was done using multiple-group confirmatory factor analysis (MGCFA), in which the parameters of the items, such as the number of dimensions (configurational invariance), factor loadings (weak invariance), intercept (strong) invariance and residual (strict) invariance are compared consecutively and cumulatively under the null hypothesis of equality between the groups compared. MGCFA was started with the evaluation of configurational invariance or the baseline model, i.e. the unidimensional structure jointly verified in males and females (configurational invariance). The reference point for this unidimensional model was what had previously been found in the total sample. Further, the factor variance to enable the total estimation of item parameters was set at 1. The comparison between the different forms of invariance was made using Cheung and Rensvold's (2002) criterion: $\Delta_{CFI} \leq 0.01$

Reliability was estimated using the *a* coefficient (Cronbach, 1951) and its confidence intervals using the Fisher method (Romano, Kromrey, Owens & Scott, 2011), and the *w* coefficient (McDonald, 1999). Both coefficients

were identified with two different reliability models respectively, essential and congeneric tau-equivalent (Haertel, 2006), which were modeled by the CFA-SEM method. Precision was also estimated using the direct score metric with standard measurement error (Nunnally & Bernstein, 1995), which should ideally be less than 0.5 (SD) to achieve the maximum tolerable measurement error around the observed scores (Wyrwich, Nienaber, Tierney & Wolinsky, 1999). Assuming that error variability is not necessarily constant in the different scoring levels of the measurment instruments (Feldt & Brennan, 1989), this error variation was examined for the BSSS8 across all scoring using the standard error of measurement (CSEM; Feldt & Brennan, 1989). This was calculated with the Mollenkopf method of polynomial regression (1949) requiring two equivalent halves of the instrument to be obtained. These halves were formed by the odd-even procedure, with the items sorted by their average responses. The CSEM method presents the information in the observed score metric.

Results

Item analysis

There were no floor or ceiling effects on items since all response options were used by the participants, and the spread of responses among them could be considered similar. Except for items 1 and 5, which show moderate asymmetry, the items approach distributional symmetry, an aspect that can be seen in the magnitude of the SSI coefficient (standardized asymmetry coefficient, Malgady, 2007), varying from 0 (symmetry) to 1 (strong asymmetry). There is a greater density of reponses in both items in the options indicating more intense SS. The kurtosis in the items was moderately heterogeneous.

Internal structure and measurement model

Several hypotheses of the internal structure of the BSSS were tested: unidimensional (M_1), related or oblique dimensions (M_2) and the bifactor model (M_3). This last model allows the separation of item variance in one related to a general common factor (FG), and another derived from specific factors, F_i (Reise et al., 2010; Reise, 2012). In accordance with the adjustment indices, the models are highly satisfactory with regard to BSSS8, as Table 2 shows. This is also borne out by the fact that the difference between them is small, which means that these models can be considered adequate. Nevertheless, comparatively the M_3 model shows the best fit, in which the residuals between the items yielded a range from 0.079 to -0.019, magnitudes that can be considered very small. All item parameters in the general factor of the M_3 model were statistically significant (t between 14.08 and 7.53), whereas the parameters of the specific factors were not statistically significant. With the exception of the general factor variance (1.536, $t = 13.32$, $p < 0.05$), the variance of the factors was between 0.00 and 0.271, and none was statistically significant ($t < 1.147$). This bi-factor model clearly states that specific factors (F_i) lose discriminatory power in the presence of a general factor (F_G). Indeed, the factor loadings (Table 3) of the specific factors vary between 0.000 and 0.414, while in the general factor the loads are moderately high (except item 4). Although the M_2 model also shows a good fit and is slightly lower than M_3, it should be noted that the interfactor correlations (Table 3) in this model ranged from 0.808 to

Table 1. *Descriptive statistics, items correlations (Pearson) and covariance.*

	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
SD	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
Ku.	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

Note. All correlations are p < 0.01. Diagonal and lower triangle for covariance matrix; upper triangle for Pearson correlations. SSI: standardized asymmetry index. As.: asymmetry coefficient. Ku.:kurtosis coefficient.

1,000, which implies a strong loss of their discriminative validity. In consequence, this model was not accepted. The M_1 model also presents a satisfactory fit and given its parsimony and the high similarity of the factor loadings with M_3, it is selected for the following analyses.

As for BSSS4 modeling, its internal structure exhibited an excellent fit. With the exception of item 1, the rest shows good discrimination level (factor loading > 0.50). It is also observed that, in general, the BSSS4 item loadings correspond to the BSSS8 items with high loadings, which suggests that the choice of Hoyle et al. (2002) to form BSSS4 with the items with the highest loadings in BSSS8 in their study is confirmed in the present sample. All BSSS4 items were statistically significant ($t > 2.00$).

Once the model (unidimensional) was identified, the measurement invariance between males and females was examined. The configurational invariance showed that the goodness-of-fit test was statistically significant, $SB\text{-}c^2$ (40) = 70.78 ($p < 0.01$). However, the fit indices can be considered satisfactory: CFI = 0.954, RMSEA = 0.050 (90% CI = 0.030, 0.069). Having shown that the dimensionality between the groups is consistent, the next step was to test the equality of the factor loadings, in which $SB\text{-}c^2$ (48) = 81.202 ($p < 0.01$) was obtained. The goodness-of-fit indicators were CFI = 0.951, RMSEA = 0.047 (90% CI = 0.030, 0.069), and were also considered satisfactory. The difference between both models (D_{CFI} = 0.003) was below the criterion (0.01) and it can therefore be claimed that both groups are invariant with respect to their item metrics. Since this type of invariance was maintained, the equality of the intercepts was examined, obtaining $SB\text{-}c^2$ (56) = 94.583 ($p < 0.01$); CFI = 0.949, RMSEA = 0.047 (90% CI = 0.030, 0.063). The difference between this model and the previous was D_{CFI} = 0.002, and did not degrade the fit by introducing this additional constraint. Finally, the restriction of equality of residuals between the items was applied to test strict invariance. Goodness-of-fit based on the $SB\text{-}c^2$ test (64) was 108.625 ($p < 0.01$), and the fit indices were CFI = 0.944, RMSEA = 0.048 (90% CI = 0.032, 0.062). The difference between the two models was D_{CFI} = 0.005, so that residual invariance between the items is maintained.

Regarding the assessment of the reliability model for the BSSS8 (Table 2), it was found that the tau-equivalent model fits slightly better than the parallel model given that the change in goodness-of-fit indices can be considered small. But compared to the congeneric model (model M_1), the difference in fit is notable. This indicates that the items may be better represented by a congeneric model, and the more accurate reliability coefficient should be obtained by means of *w*. Since the BSSS4 items are contained in the BSSS8, it was assumed that the invariance and reliability modeling results can be transferred to the BSSS4.

Reliability

Table 3 shows the *a* and *w* reliability coefficients. Although it can be seen that the congeneric model shows a better statistical fit, the practical significance of this adjustment does not seem to be shown in the calculated reliability coefficients. Therefore *a* and *w* can be considered very similar. The effect of the reduction of the items on the *a* coefficient was evaluated in two ways. First, the inferential hypothesis that the *a* coefficient of the short form is equal to that of the long form was tested using an ad hoc computer program (ALPHATEST; Lautenschlager & Meade, 1987),

Table 2. *Results of the internal structure fit test and measurement model.*

Internal stucture	SB-c2 (gl)	RMSEA (CI 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
Reliability model (BSSS8)						
Parallel	96.459** (33)	0.056 (0.043, 0.069)	0.915	0.910	0.068	30.45
Tau-Equivalent	77.475** (26)	0.057 (0.042, 0.071)	0.921	0.921	0.036	25.47

Nota. $^{**}p < 0.01$. Models: M_1, unidimensional; M_2, oblique factors; M_3, bi-factor. BSSS8: 8 item scale. BSSS4: 4 item scale.

Table 3. *Individual parameters of the bifactor and unidimensional models.*

	BSSS8								BSSS4	
	Bifactor						Unidimensional			
	F_1	F_2	F_3	F_4	F_G	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
Interfactorial correlation										
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	-
Descriptive statistics										
M	-	-	-	-	-	-	24.77	-	11.98	-
SD	-	-	-	-	-	-	6.17	-	3.47	-
As.	-	-	-	-	-	-	-0.05	-	0.09	-
Ku.	-	-	-	-	-	-	-0.09	-	-0.27	-

Note. h^2: communality. *a*: alpha coefficient. *w*: omega coefficient. BSSS8: 8 item scale. BSSS4: 4 item scale.

which applies the asymptotic procedure based on the *F* distribution (Feldt, 1980; Feldt, Woodruff & Salih, 1987). Using the uncorrected and spurious-corrected inter-form correlations (see following section), a statistical difference between the two coefficients was found: $c^2[1] = 55.79$ ($p < 0.0001$) and $c^2[1] = 21.97$ ($p < 0001$), respectively. Second, the 95% confidence interval (Feldt et al., 1987; Romano, Kromrey, Owens & Scott, 2011) was compared heuristically for the a of BSSS8 (0.71: 0.77) and BSSS4 (0.59: 0.68). Both analyses show that the reduction of internal consistency in BSSS4 is substantial and statistically significant. The reliability of BSSS8 is the same or only very slightly different from that reported in the American validation studies (Hoyle et al., 2002; Stephenson et al., 2003). When examining instrument accuracy in the measurement of the observed scores, the standard error of measurement in BSSS8 (3.11) and BSSS4 (2.07) compared to the maximum tolerable dispersion (3.08 and 1.73, respectively) was found to be slightly higher.

For the CSEM, the polynomial regression equation was calculated by multiplying each score by the coefficient b = 0.133 ($R^2 = 0.475$, $F[2, 615] = 89.73$, $p < 0.01$), given that the neither the constant nor the quadratic and cubic component were statistically significant. Table 4 shows the CSEM and the standardized CSEM (CSEM-S). It can be seen that the scores show less error variation in the lower and near average levels, and that this decline becomes more pronounced, with reliability falling to below 0.70 with scores above 26.

Equivalence BSSS8 and BSSS4

The equivalence and consistency between BSSS8 and its abbreviated version, BSSS4, was examined. The consistency between these measures, as shown by the correlations between them, was corrected by spurious or correlated errors (Bashaw & Anderson 1967; Levy, 1967). The spurious corrected correlation should be high (≥ 0.70, Petrides et al., 2003; Putnam & Rothbart, 2006) in order to argue for the linear dependence between forms (Smith et al., 2000). The observed correlation between BSSS forms was 0.89 (95% CI = 0.87: 0.90) and the corrected correlation was 0.68 (95% CI = 0.63: 0.72). With respect to the reference value (0.70), the corrected correlation was not statistically significant ($z = -0.94$, $p > 0.05$).

To examine the concordance between BSSS8 and BSSS4, the weighted kappa coefficient (K_W; Fleiss, Levin & Paik, 2003) was used with linear weights assigned to scores categorized in deciles. We obtained K_W = 0.679 (95% CI: 0.64, 0.70), a value that can be considered a good level of agreement and equal to the corrected correlation obtained in the previous paragraph. In order to identify the characteristics of the agreement between the versions, they can be better observed in the Bland-Altman plot (1986), contrasting the difference between BSSS8 and BSSS4 against the full version score (Krouwer, 2008). Figure 1 suggests that, in general, discrepancies between BSSS4 and BSSS8 occur within acceptable limits (± 1.96 SD of BSSS8-BSSS4 differences); The linearly increasing size of the differences is due to the joint increase of the scores of both versions and is therefore expected. However, it was found that in the lower decile scores there are more discrepancies, although they are infrequent compared to the total sample.

Table 4. *Conditional standard error of measurement for BSSS8 y BSSS4.*

	CSEM	
Score	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	-

Note. CSEM: Conditional standard error of measurement.

Discussion

The results of the structural validation of the BSSS (8 and 4 items) in Peruvian adolescents showed that a single dimension better represents the variability of the items. This strong support for the unidimensional structure may be determined by two factors: First, the correlations between the constructs (experience seeking, susceptibility to boredom, emotion and adventure seeking, and disinhibition) were very high, which generally suggests that the model does not meet the criterion of discriminative validity (Nunnally & Bernstein, 1995). Second, the introduction of the bifactor model clearly showed that the validity of the items is strongly associated with a single general factor, with higher factor loadings than on the specific factors. Therefore, the larger common variance in this general factor indicates unmistakeably that there is no statistical justification for interpreting dimensional scores separately (Reise et al., 2010; Reise, 2012). In contrast to other studies in which factor methods were applied, this is the first to advance in the identification of BSSS structure using bifactor

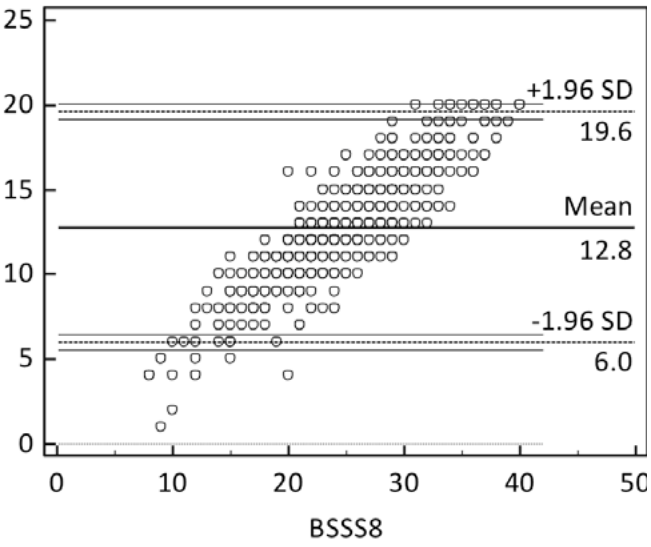

Figure 1. Bland-Altman plot of direct scores for BSSS8 and BSSS4..

modeling, which confirmed the unidimensional nature of the instrument.

Regarding the internal consistency estimates, those obtained in previous studies for the BSSS with adolescents and adults (cited in the present study), do not in general exceed 0.80, for either subscale scores or for total score. This pattern is also replicated in the results reported here, and may be a consequence of the origin of the construct since it is considered to be a biological trait influenced at the same time by interaction with the social environment; it is a complex construct composed of experiences and feelings (Chico, 2000; Hoyle et al., 2002; Pérez & Torrubia, 1986; Primi et al., 2011; Stephenson et al., 2003; Zuckerman & Neeb, 1980), which groups and represents heterogeneous behaviors with a common denominator: risk. This factor can be conceptually and empirically linked to variables and physical, legal, social, moral, financial, etc., issues (Horvath & Zuckerman, 1996; Ledesma, Poó & Peltzer, 2007). With this underlying feature, which originates in the definition of the SS construct and the discussion about its internal structure, it is reasonable that no higher levels of reliability are obtained. Yet with regard to comparing the internal consistency of our study with that of previous studies on BSSS8 and BSSS4 in adolescents (Banerjee & Greene, 2009; Hoyle et al., 2002; Stephenson, Morgan, Lorch, Palmgreen, Donohew, & Hoyle 2002; Stephenson et al., 2003; Vallone et al., 2007), very similar results were obtained. Thus there is evidence that the true variability of BSSS8 and BSSS4 seems to be mainly constant from an intercultural angle. This has direct implications for the accuracy of SS assessment using these brief measures, which, although not high enough for clinical use, may be the best option for developing epidemiological studies and describing groups in a scientific research context. Furthermore, taking into account that the standard error of measurement of the scores is close to the maximum acceptable, the user will have to decide if it is appropriate to apply an adjustment to the scores or statistics obtained with the instrument in order to mitigate the error of measurement (Nunnally & Bernstein, 1995). When interpreting the scores, it should also be taken into account that the scores below the average are more reliable, whereas higher scores are more likely to be affected by measurement error. Therefore, the description of a group or subject based on the BSSS8 scores, and even more so with the BSSS4, may be expressed with less bias when the adolescent subject's sensation seeking is of low intensity. At this point it is not known whether social desirability or other irrelevant response patterns have played a role, but it is an issue pending resolution in subsequent research.

In brief measures such as BSSS8 and BSSS4, it may be preferable to measure a general construct compared to multifactorial measurement, because the number of items in each specific factor is small and directly affects the magnitude of the reliability coefficients (Heartel, 2006). Moreover, it should be pointed out that in such brief measures of a construct as broad as SS, the internal consistency found can be considered satisfactory for certain contexts of use, such as those already mentioned: for studies and descriptions of groups, especially those of an epidemiological nature. The most demanding levels of reliability (eg scores of 0.90 or higher) for measurement instruments are reserved for contexts where decisions have to be made on the individual subject (e.g. diagnostics); measurements errors can be unacceptable in such contexts (Nunnally & Bernstein, 1995) and lead to consequences counter to the interests of individuals or groups and institutions (for example, a poor diagnosis of a depressive disorder, errors in the interpretation of a company's working environment). Given this view, BSSS8 and BSSS4 can safely be used in research or when work is done on groups of subjects. For more precise assessments, it will be necessary to use a combination of instruments or instruments with more items, which will provide a better appreciation of the internal consistency. For example, the SSS-V scale adapted for adolescents (Perez, et al., 1986, Pérez, et al., 1987) could be used as an aid in the diagnosis of this personality trait.

Furthermore, another relevant aspect of factor loadings indicates that the definition of the construct in BSSS8 and BSSS4 is similarly weighted by each item, and that the simple sum of the items can be accepted as an indicator of the construct. Conversely, the heterogeneity of the factor loadings would indicate that each item influences the definition of the construct differently, and therefore a better representation of the score may be obtained by weighting each item differentially. Fortunately, the similarity of discriminative power is satsfied in each item, and although they are not very high (e.g. 0.80 or more), they contribute with sufficient relevant variance to the construct for group descriptive purposes. This description can use the same measurement parameters found among males and females, as the instrument has fulfilled the increasingly demanding forms of invariance. Thus, an exhaustive comparison of the differences in the level of latent or observed scores appears to be assured.

Finally, the equivalence analyzed between BSSS8 and BSSS4 has been found to be satisfactory, and can be considered more precise than the convergent correlation between both, since the authors (Hoyle et al., 2002; Stephenson et al., 2003; Stephenson et al., 2007) did not adjust the correlation coefficient to effectively control the variance of correlated errors. The BSSS4 scale can be used for epidemiological studies or in research alongside other instruments, especially when the number of items, time limits and the availability of the subjects interact negatively to require abbreviated but valid measures.

The present study presents some limitations, such as the lack of evaluation of the invariance of measurement or the

differential operation of items, the characterization of the items through Item Response Theory modeling, and external links with other convergent and divergent constructs. These aspects are pending further research.

Acknowledgements

The authors thank the University of San Martín de Porres for the facilities to carry out this study, as well as the schools that gave them access. We are also grateful for the reviewers' comments.

Conflicts of Interest

The authors declare that the present manuscript, its conception and development are free from conflicts of interest.

References

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. & 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. & 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. & 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. & 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. & 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. & Dugeon, 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. & Wu, E. J. C. (2012). *EQS for Windows.* Encino CA: Multivariate Software, Inc.

Birenbaum, M. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & Paik, M. C. (2003). *Statistical methods for rates and proportions (3rd ed)*. Hoboken: John Wiley & 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. & 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. & 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. & 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. & 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. & 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. & Wu, Q. (2012). Estimation in structural equation modeling. In: Hoyle, R. H. (Ed.), *Handbook of structural equation modeling* (pp. 164–179). New York: Guildford 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. & 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. & 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. & 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. & Jouvent, R. (1998). Construction and validation of a sensation seeking scale for adolescents. *Personality and Individual Differences, 26,* 159-174. doi:org/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. & 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. & 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.

Nunally, J. C. & 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. & Andrade, P. (2010). Intento de suicidio y búsqueda de sensaciones en adolescentes. *Revista Intercontinental de Psicología y Educación, 12*, 53-75.

Pérez, J. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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. & 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 Guarnieri, 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. & 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. & 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. & 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. & 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. & 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 Statatistical, 14*, 97-110. doi:10.1081/bip-120028508.

Zuckerman, M. & 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. & 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.