

DAILY SPIRITUAL EXPERIENCES SCALE

Daily Spiritual Experiences and Well-Being among the Nonreligious, Spiritual, and Religious: A

Bifactor Analysis

Joseph H. Hammer

University of Kentucky

Ryan T. Cragun

University of Tampa

Note: This article may not exactly replicate the final version published in the journal. It is not the copy of record. Please use the DOI link on my website (<http://drjosephhammer.com>) to access the PDF through your institution, allowing full access to the published type-set article. The APA-style citation for this article can be found on the Publications page of my website.

Author Note

Joseph H. Hammer, Department of Educational, School, and Counseling Psychology, University of Kentucky; Ryan T. Cragun, Department of Sociology, University of Tampa.

Please address all correspondence to Joseph H. Hammer, Department of Educational, School, and Counseling Psychology, University of Kentucky, 243 Dickey Hall, Lexington, KY, 40506, USA. Email: jhhammer@gmail.com.

Abstract

The Daily Spiritual Experiences Scale (DSES; Underwood & Teresi, 2002) has been used in hundreds of investigations to study the role of everyday spiritual experiences in the lives of the Not Religious nor Spiritual (NRS), Spiritual but Not Religious (SNR), and Religious and Spiritual (RS). However, there is a lack of measurement equivalence/invariance (ME/I) evidence to support the use of the DSES to compare these three groups. The present study ($N = 1623$ U.S. adults) sought to (a) identify the most appropriate factor structure for the DSES within samples of NRS, SNR, and RS persons, (b) examine the ME/I of the DSES factor(s) across the three groups, and (c) investigate the relationship between the DSES and common measures of well-being (e.g., meaning in life, satisfaction with life, physical health, and mental health). Our results indicate that the DSES measured the same two independent daily spiritual experience constructs when used with SNR and RS participants, but a different array of inadequately-defined daily spiritual experience constructs among the NRS participants. This raises questions about the utility of using the DSES with NRS samples, and provides evidence that the DSES conforms to a bifactor, rather than unidimensional, structure among SNR and RS participants. Lastly, our findings paint a complex picture of the relationship between daily spiritual experiences and well-being across these three groups, such that daily spiritual experiences are not always associated with enhanced well-being. Implications for research and practice use, scoring, and interpretation of the DSES across these three groups is discussed.

Keywords: bifactor analysis; validity; daily spiritual experiences; spirituality; religion

Daily Spiritual Experiences and Well-Being among the Nonreligious, Spiritual, and Religious: A
Bifactor Analysis

The Daily Spiritual Experiences Scale (DSES; Underwood & Teresi, 2002) is a 16-item self-report instrument designed to measure everyday ordinary (cf. extraordinary or dramatic mystical experiences such as near-death experiences) spiritual experiences in daily life. It has been used in hundreds of religion and spirituality studies to measure experiences of relationship with and awareness of the divine or transcendent (Underwood, 2011). Underwood stated that the DSES is suitable for cross-cultural research with people from a variety of religious and cultural backgrounds. For example, Underwood notes that the sixteen items measure a wide variety of experiences, some focused on relationship with the divine and others focused on nontheistic experiences of feeling awe, gratitude, and connection. By allowing participants to substitute the word “God” with another term that better represents the divine to the participant, and including additional items that don’t invoke the term “God,” the developers sought to create a DSES that could validly capture the experiences of people who do not believe in a higher power.

Since its initial publication, researchers have used the DSES with individuals from across the religious/spiritual (R/S) spectrum. Sanchez, Arocena, & Ceballos (2010) used a Spanish version of the DSES to compare frequency of daily spiritual experiences (DSE) across Mexican and Basque Protestants, Catholics, Agnostics, and Atheists. Kalkstein and Tower (2009) used the DSES to examine the frequency of DSE among “religious nones” compared to other religious groups. Dong, Wu, Zhu, Jin, and Zhang (2017) used a Mandarin version of the DSES with three samples of nonreligious Chinese people. Koenig, George, and Titus (2004) used the DSES to examine the relationship between R/S and health among several groups of medical patients,

including people who self-identified as Not Religious nor Spiritual (NRS), Spiritual but Not Religious (SNR), and Religious and Spiritual (RS).

These researchers have operated on the premise that the DSES can be validly used to examine differences in DSE or compare the association between DSE and other constructs of interest across these three groups (i.e., the NRS, SNR, and RS). However, the validity of such results depends on the DSES having the same theoretical structure and meaning for each group (Miller & Sheu, 2008). This equivalence must be empirically verified by testing the DSES for measurement equivalence/invariance (ME/I) across these three groups. In the absence of evidence of ME/I, group difference and comparative association results cannot be trusted, as they may merely be the product of differences in measurement, rather than true differences in the frequency of DSE (Cheung & Lau, 2012). Given ongoing interest in using the DSES with groups across the R/S spectrum, the present study sought to provide evidence that the DSES demonstrates strong ME/I (Dimitrov, 2010) across NRS, SNR, and RS participants.

For an instrument to demonstrate ME/I, the instrument must first demonstrate the same factor structure within each group. While DSES dimensionality within NRS or SNR samples has never been directly studied, it has been examined within samples composed primarily of RS participants (for a review of DSES factor analysis studies, see Lace & Handal, 2017). Most studies have presented results that were interpreted as providing support for modeling the DSES as a unidimensional instrument (Underwood, 2011). However, at least nine studies have provided evidence of a multi-dimensional structure for the English-language DSES or its translations. Four research teams provided evidence that the DSES may be defined by two correlated factors: a theistic factor primarily consisting of items regarding connection with God and a nontheistic factor primarily consisting of items regarding spiritual connection with others

and nature (Currier, Kim, Sandy, Neimeyer, 2012; Lace, & Handal, 2017; Schuurmans-Stekhoven, 2013; Zemore & Kaskutas, 2004). Five research teams have uncovered evidence of one dominant factor upon which most items load plus a smaller factor (sometimes labeled “compassionate love”) consisting of two or three items focused on connection with others (Kalkstein & Tower, 2009; Lo, Chen, Wasser, Portenoy, and Dhingra, 2016; Ng, Fong, Tsui, Au-Yeung, & Law, 2009; Rakošec, Mikšić, Juranić, & Batinić, 2015; Underwood & Teresi, 2002). Depending on the strength of the correlation between the two factors or the degree of simple versus complex (i.e., certain items loading on more than one factor) structure, some of the nine research teams chose to retain the two-factor solution while others argued for treating the DSES as unidimensional. Finally, Kim, Martin, and Nolty (2016) found that a one-factor confirmatory factor analysis (CFA) model, which allowed the residuals of seven item pairs to correlate, provided adequate fit to the DSES data. The authors justified these correlated residuals by asserting that “each of these pairs of items functions like a testlet, a set of items that share a common passage or context, known to usher in additional sources of covariance typically not reflected in the main target construct” (p. 241). In summary, there is ongoing disagreement about the dimensionality of the DSES with samples composed primarily of RS participants, and no empirical data on the dimensionality of the DSES among the NRS and SNR. Therefore, the present study used factor analysis to determine what factor model fit best within each of the three R/S groups, and then used ME/I analysis to investigate equivalence across these groups.

Beyond issues of equivalence, Underwood (2017) suggests that published DSES research indicates that increasing a person’s DSE could prevent negative psychological effects of trauma and improve well-being. However, there are psychology of nonreligion scholars (e.g., Hammer, Cragun, & Hwang, 2013) who suggest that experiences of connection with the divine may not be

culturally relevant or palliative for NRS persons who don't believe in the existence of the divine. Thus, it is best to directly verify the relationship between DSE and well-being among the NRS. To our knowledge, while studies have examined this relationship among people who self-identify as nonreligious (e.g., Dong et al., 2017), we are unaware of any studies that have looked at this relationship among the NRS. Thus, the present study sought to offer direct empirical evidence regarding how DSE relates to well-being among the NRS, as well as among the SNR and RS.

The Present Study

The present study sought to (a) identify the most appropriate factor structure for the DSES within samples of NRS, SNR, and RS persons, (b) examine the ME/I of the DSES factor(s) across the three groups, and (c) investigate the relationship between the DSES factor(s) and common measures of well-being (e.g., meaning in life, satisfaction with life, physical health, and mental health). Extant literature guided the formation of three hypotheses. Because the preponderance of factor analysis evidence seems to support a unidimensional model for the DSES, we hypothesized (H1) that a single factor would account for the DSES item variance. Given that the DSES has been lauded for its cross-cultural replicability, we hypothesized (H2) that the DSES would demonstrate strong ME/I (Dimitrov, 2010). Strong ME/I is concluded when configural (i.e., model fits adequately in each group), metric (i.e., the 16 items load on the DSES factor(s) to a similar degree in each group), and scalar (i.e., the 16 item intercepts have similar magnitude in each group) invariance is present. Lastly, because of published evidence of positive relations between DSES scores and positive outcomes across R/S groups and cultures (e.g., Underwood, 2017), we hypothesized (H3) that the DSES factor(s) would demonstrate a positive relationship with well-being factors among all three groups.

Method

Participants and Procedure

Participants were 1623 adults. Participants were recruited via ResearchMatch, a national health volunteer registry created by several academic institutions and supported by the U.S. National Institutes of Health as part of the Clinical Translational Science Award (CTSA) program. ResearchMatch has a large population of volunteers who have consented to be contacted by researchers about health studies for which they may be eligible. Review and approval for this study and all procedures was obtained from the (*masked*) Office of Research Integrity. Participants were contacted via the registry system regarding the study, advertised as a survey about measuring spirituality across diverse groups of people. Participants were also recruited via websites (e.g., www.atheistresearch.org), listservs (e.g., Nonreligion and Secularity Research Network listserv), and social media platforms (e.g., Atheist Research Collaborative Facebook Page and Twitter Account) related to (non)religion. Interested participants were directed to an online survey that began with an informed consent page, followed by the instrument battery and demographic items, and ended with a conclusion page. Participants had the option of entering a drawing for one of several \$25 Amazon.com gift cards.

Participants (1128 women, 474 men, and 18 individuals who self-identified with a different gender identity label) ranged in age from 18 to 88 years old ($M = 46.50$, $SD = 16.77$). Approximately 85% of the sample identified as White, 4% as African American/Black, 3% as Latino/a, 3% multiracial, 2% Latino/a, 2% Asian American or Pacific Islander, and 3% as another racial/ethnic identity. Approximately 60% reported being married or in a committed relationship or civil union, 23% single, 12% separated or divorced, 3% widowed, and 1% preferred not to answer. Approximately 0.2% reported having less than a high school education,

3% earned a high school diploma or GED, 7% earned a two-year degree, 14% had some college experience, 30% earned a four-year college degree, 45% earned a graduate or professional degree, and 0.1% preferred not to answer. When asked “What label best describes how you generally identify yourself when asked what your religion or belief system is?”, participants responded as follows: 23% Atheist, 14% Agnostic, 11% Roman Catholic, 8% Christian unspecified, 8% non-denominational Christian, 4% Protestant unspecified, 2% United Methodist, 2% Jewish, 2% Southern Baptist, 2% Buddhist, 1% Episcopalian, 1% The Church of Jesus Christ of Latter-day Saints, 0.6% New Age, 0.4% Orthodox Christian, 0.4% Muslim, 0.2% Hindu, 0.1% American Indian or Native American Religion, and 19% opted to provide their own label. The five most common labels provided were Humanist ($n = 17$), Unitarian Universalist ($n = 16$), Lutheran ($n = 10$), Pagan ($n = 7$), and Spiritual ($n = 6$).

Measures

Daily Spiritual Experiences (DSE). The 16-item DSES (Underwood & Teresi, 2002) was designed to measure ordinary spiritual experiences of daily life. The first fifteen items use a 6-point Likert scale from 1 (*never*) to 6 (*many times a day*). The developers recommend opposite scaling such that 1 represents “many times a day” but we suggest that a higher score should indicate more of the construct and thus used 6 to represent “many times a day” to aid reader comprehension. The sixteenth item uses a 4-point Likert scale from 1 (*not at all*) to 4 (*as close as possible*). The developers suggest that the sixteenth item scaling can be rescaled to a 6-point Likert scale to achieve consistency with the other fifteen items; we did so. The DSES has demonstrated convergent and predictive evidence of validity and internal consistency ($\alpha > .89$; Underwood, 2011). The present study’s internal consistency estimates for the NRS, SNR, and RS were .67, .92, and .95, respectively.

Meaning in Life. The 5-item presence subscale of the Meaning in Life Questionnaire (MLQ; Steger, Frazier, Oishi, & Kaler, 2006) was used to measure felt meaning in life. Responses are rated on a seven-point Likert scale from 1 (*absolutely untrue*) to 7 (*absolutely true*), with a higher mean score indicating greater meaning. The MLQ has demonstrated convergent and discriminant evidence of validity and internal consistency ($\alpha > .82$; Steger et al., 2006). The present study's internal consistency estimates for the NRS, SNR, and RS were .89, .89, and .91, respectively.

Satisfaction with Life. The Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985) is a 5-item instrument designed to measure cognitive self-evaluation of global life satisfaction. Responses are rated on a seven-point Likert scale from 1 (*strongly disagree*) to 7 (*strongly agree*), with higher scores representing higher life satisfaction. The SWLS has demonstrated convergent and predictive evidence of validity and internal consistency ($\alpha > .79$; Pavot & Diener, 2008). The present study's internal consistency estimates for the NRS, SNR, and RS were .90, .90, and .88, respectively.

Physical and Mental Health. The 4-item global physical health subscale and the 4-item global mental health subscale derived from the 10 global health items of the National Institute of Health's Patient-Reported Outcome Measurement Information System (PROMIS; Hays, Bjorner, Revicki, Spritzer, & Cella, 2009) were used to measure physical and mental health, respectively. Responses are rated on a five-point Likert scale from 1 (e.g., *poor*) to 7 (e.g., *excellent*), with a higher mean score indicating better health. These subscales have demonstrated convergent evidence of validity and internal consistency ($\alpha > .81$; Hays et al., 2009). The present study's physical health internal consistency estimates for the NRS, SNR, and RS were .76, .73, and .80,

respectively. The mental health estimates for the NRS, SNR, and RS were .82, .83, and .81, respectively.

Religiousness and Spirituality. The two Overall Self-Ranking items from the Brief Multidimensional Measure of Religiousness/Spirituality (Abeles et al., 1999) were used to measure how religious (i.e., “To what extent to you consider yourself a religious person?”) and spiritual (“To what extent to you consider yourself a spiritual person?”) respondents considered themselves, using a 4-point Likert scale from “not at all” to “very.”

Feedback. Participants were given the opportunity to provide free-response feedback about the DSES after completing it. Participants’ anonymous feedback is provided verbatim in the Supplemental Material.

Results

Data Cleaning

The initial dataset contained 1836 individuals. Cases with 20% or more missing data ($n = 43$; Parent, 2013) or more than one incorrect response to the three attention check items ($n = 138$) were deleted. Cases that did not answer both the “To what extent do you consider yourself a religious person?” and “To what extent do you consider yourself a spiritual person?” questions ($n = 32$) were also deleted, given that they could not accurately be sorted into one of the three groups to be analyzed. In the final sample ($N = 1623$), no variables exceeded the cutoffs of 3 and 10 for high univariate skewness and kurtosis values, respectively (Weston & Gore, 2006). Missing data across all items ranged from a low of 0.1% for DSES item 1 to a high of 1.0% for DSES item 2. We used Mplus’ (version 6.11; Muthén & Muthén, 1998-2012) Full Information Maximum Likelihood (FIML) method to handle missing data. We used the maximum likelihood estimation with robust standard errors (MLR) option to estimate the scaled χ^2 test statistic, and

associated fit indices that use it, to protect against deviations from multivariate normality. We used available item analysis when using SPSS (Version 24). To sort participants into the three worldview groups of interest, the following criteria were used. First, the Not Religious nor Spiritual (NRS) group included 422 participants who indicated both “not religious at all” and “not spiritual at all” on the Religiousness and Spirituality items (see Measures section). Second, the Spiritual but Not Religious (SNR) included 467 participants who indicated “not religious at all” and at least “slightly spiritual”. Third, the Religious and Spiritual (RS) group included 734 participants who indicated at least “slightly religious” and “slightly spiritual”.

Dimensionality

The dimensionality of the DSES was tested using a series of confirmatory factor analysis (CFA) measurement models. Model fit was evaluated using the Root Mean Square Error of Approximation (RMSEA), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Standard Root Mean Square Residual (SRMR). The following fit criteria were used: $RMSEA \leq .06$, $CFI \geq .95$, $TLI \geq .95$, $SRMR < .08$ for good fit and $RMSEA \leq .10$, $CFI \geq .90$, $TLI \geq .90$, $SRMR < .10$ for adequate fit (Weston & Gore, 2006).

Underwood’s (2011) recommended unidimensional model did not demonstrate an adequate fit to the data for all three groups, nor did any of the two-factor models, including Currier et al.’s (2012) model, Lace and Handal’s (2017) model, Schuurmans-Stekhoven’s (2013) full scale model or trimmed model, Kalkstein and Tower’s (2009) model, or Lo et al.’s (2016) model. The least inadequate of these models was Schuurmans-Stekhoven’s full scale model (scaled $\chi^2 [103] = 890.89$, $p < .001$, $RMSEA = .102$ [90% CI of .096, .108], $CFI = .888$, $TLI = .870$, $SRMR = .050$) within the RS sample. Kim et al.’s (2016) unidimensional model with seven correlated residuals provided adequate fit within the RS sample according to three of the

four indices (scaled χ^2 [97] = 716.38, $p < .001$, RMSEA = .093 [90% CI of .087, .100], CFI = .912, TLI = .891, SRMR = .065). Schreiber and colleagues (2006) indicated that if the “vast majority of the indexes indicate a good fit, then there is probably a good fit” (p. 327), suggesting overall adequate fit for this correlated residual model. This model did not provide adequate fit for the NRS or SNR, however. Full model fit and item loading results for these 24 models (i.e., 8 models by 3 groups) are available from the author upon request. In summary, no previously published factor models demonstrated adequate fit among the NRS and SNR, whereas the correlated residuals model provided overall adequate fit among the RS. To maximize the chances of identifying a suitable model that could demonstrate ME/I across the three groups, we next turned to EFA with subsequent CFA to examine the factor structure of the DSES within each group separately.

Not Religious nor Spiritual Group. For the NRS sample, we first conducted an EFA with SPSS using principal axis factor extraction and promax (oblique) rotation. One thousand random Parallel Analysis data sets were also computed. Eigenvalues for the first four factors were higher in the actual data set (i.e., 3.01, 2.57, 1.79, 1.40, 1.09) than in the parallel analysis (i.e., 1.33, 1.29, 1.21, 1.17, 1.13). In conjunction with the Parallel Analysis results and inspection of the screen plot, examination of the pattern coefficients of three-factor, four-factor, and five-factor solutions suggested that the four-factor solution best approximated simple structure and had the strongest theoretical interpretability. Examination of the pattern coefficients (see Table 1) indicated that items 4 (“I find strength in my religion or spirituality.”) and 5 (“I find comfort in my religion or spirituality.”) loaded $> .90$ on Factor 1 (“Testlet”). Per Kim et al., (2016), this factor is best conceptualized as a testlet rather than a substantive factor. Factor 2 (“Nontheistic”) consisted of five nontheistic items. Factor 3 (“Theistic A”) consisted of

five theistic items and Factor 4 (“Theistic B”) consisted of three theistic items. We were unable to come up with a logical rationale for why these eight theistic items would split into two factors, as each factor contains a mix of themes and language. Importantly, these theistic items continued to load on separate factors in both the three- and five-factor solution pattern matrices, suggesting this phenomenon is not due to an idiosyncrasy of the four-factor solution. Item 11 (“I am spiritually touched by the beauty of creation”) failed to load $> .32$ on any of the four factors, loading at $.30$ on the Testlet factor and $.21$ on the Nontheistic factor. The four rotated factors accounted for 16.06%, 12.37%, 8.24%, and 5.73% of the total variance, respectively. The four factors explained a cumulative total of 42.39% of the variance, which falls below the suggested minimum of 50% offered by Peterson (2000). Examination of the descriptive statistics in Table 2 provides insight into how NRS participants’ responses may have shaped these factors. Specifically, the modal score for the eight theistic items was 1 (“never”). In other words, NRS participants overwhelmingly reported an absence of DSE related to spiritual forces (i.e., God, the divine or holy).

We used the same NRS data to specify a four correlated factors CFA model suggested by the EFA; item 11 was not used given its inability to load simply on a factor. This model did not demonstrate adequate fit to the data (scaled $\chi^2 [84] = 216.11, p < .001, RMSEA = .061$ [90% CI of $.051, .071$], CFI = $.806, TLI = .758, SRMR = .060$). Furthermore, *Mplus* indicated that standard errors of the model parameter estimates may not have been trustworthy due to a non-positive definite first-order derivative product matrix, caused by item 9 (“I feel God’s love for me, directly.”), to which 419 of the 422 NRS participants responded “never.” Dropping item 9 did not result in appreciably improved model fit (scaled $\chi^2 [71] = 205.14, p < .001, RMSEA = .067$ [90% CI of $.056, .078$], CFI = $.79, TLI = .732, SRMR = .060$). Given that the Testlet factor

and the Theistic B factor moderately correlated, we tweaked the prior model such that items 4 and 5 loaded on the Theistic B factor and their residuals were allowed to correlate, per Kim et al., (2016). This three correlated factors CFA model also failed to demonstrate adequate fit (scaled χ^2 [73] = 218.00, $p < .001$, RMSEA = .069 [90% CI of .0568, .079], CFI = .774, TLI = .719, SRMR = .066). In summary, we were unable to identify a suitable factor structure for the DSES for use with NRS respondents. It is important to note that the testlet, theistic, and nontheistic items demonstrated strong empirical independence from each other for this group, rather than demonstrating covariation that could justify the use of a single DSE factor.

Spiritual but Not Religious Group. For the SNR sample, we first conducted an EFA using the same analysis procedure described above. The first eigenvalue was 7.73, which was more than three times the size of the second eigenvalue of 2.09. Cho et al., (2015) states that this indicates the presence of a strong general factor and that traditional EFA should be abandoned in favor of bifactor EFA, which can be done using BI-GEOMIN rotation in *Mplus* (Version 7.3; Muthén & Muthén, 1998-2012). Briefly, a DSES bifactor model indicates that item variation is explained by the additive contribution of one general factor (that all 16 DSES items load onto) and one or more specific factors (that a subset of DSES items also load onto; Reise, 2012). The two factors are statistically prevented from correlating, creating independent factors that compete to account for each item's variance. Bifactor modeling allowed us to determine if the DSES is defined by both a general DSES factor and one or more specific factors (e.g., nontheistic DSES, theistic DSES), rather than one factor (i.e., unidimensional model).

The bifactor EFA eigenvalues for the first two factors were higher in the actual data set (i.e., 7.82, 2.21, 1.20) than in the parallel analysis (i.e., 1.33, 1.26, 1.21). Examination of the pattern coefficients (see Table 1) indicated that 14 of the 16 items loaded $> .32$ on a general DSE

factor in both the two- and three-factor solutions. When the testlet items were allowed to form a specific Testlet factor in the three-factor solution, the nontheistic items all achieved loadings of $> .32$ on their own specific Nontheistic factor. Thus, among the SNR respondents, it appears that the DSES is characterized by a well-defined general DSE factor characterized by both theistic and nontheistic DSE and an adequately-defined, independent specific factor characterized by nontheistic DSE. Examination of the descriptive statistics in Table 2 indicates that the SNR participants had a modal score of “1” for all eight of the theistic items, just like the NRS participants. SNR participants, in comparison to NRS participants, had higher modal scores on the three nontheistic items invoking semi-spiritual concepts such as spiritual connections and blessings.

We used the same SNR data to specify a bifactor CFA model suggested by the bifactor EFA. Given that the testlet items demonstrated strong loadings on the general factor, we specified one general factor upon which all 16 items loaded, one specific factor upon which all 6 nontheistic items loaded (including item 11), and allowed the residuals of the testlet items to correlate. Per Schreiber and colleagues' (2006) guidelines, this model appeared to demonstrate overall adequate fit (scaled $\chi^2 [97] = 418.01, p < .001, RMSEA = .084 [90\% \text{ CI of } .076, .093], CFI = .920, TLI = .902, SRMR = .064$). In summary, we were able to identify a borderline-suitable bifactor structure (i.e., one general DSE factor and one specific nontheistic factor) for the DSES for use with SNR respondents.

Religious and Spiritual Group. For the RS sample, we followed the same procedures outlined for the SNR sample. The first traditional EFA eigenvalue was 8.98, which was more than three times the size of the second eigenvalue of 1.61. The bifactor EFA eigenvalues for the first two factors were higher in the actual data set (i.e., 8.97, 1.59, 0.94) than in the parallel

analysis (i.e., 1.26, 1.21, 1.17). Examination of the pattern coefficients (see Table 1) indicated that all 16 items loaded $> .32$ on a general DSE factor in both the two- and three-factor solutions.

When the testlet items were allowed to form a specific Testlet factor in the three-factor solution, most of the nontheistic items achieved loadings of $> .32$ on their own specific Nontheistic factor. Thus, for both the SNR and RS respondents, it appears that the DSES is characterized by a well-defined general DSE factor characterized by both theistic and nontheistic experiences and an adequately-defined, independent specific factor characterized by nontheistic DSE. Examination of the descriptive statistics in Table 2 indicates that the RS participants, unlike the SNR and RS participants, endorsed both the theistic and nontheistic items to a greater degree ($M = 4.04$ across the 16 items).

We used the same RS data to specify a bifactor CFA model suggested by the bifactor EFA, which was the same as the CFA model for the NRS sample. This model (scaled $\chi^2 [97] = 463.704, p < .001, RMSEA = .072$ [90% CI of .065, .078], CFI = .948, TLI = .936, SRMR = .040) demonstrated adequate fit. In summary, we were able to identify a suitable bifactor structure (i.e., one general DSE factor and one specific nontheistic factor) for the DSES for use with RS respondents.

Measurement Invariance/Equivalence across SNR and RS Groups

Having explored plausible factor structures for the DSES within the three groups, we next used ME/I analyses to determine whether the DSES factors are measuring similar concepts across the groups. Given that the DSES did not demonstrate a bifactor structure—or any other well-defined factor structure—among the NRS, by definition the DSES cannot demonstrate configural invariance between the NRS and the other two groups. Given that the DSES did evidence a bifactor structure with one general and one specific factor explaining the same items

across the SNR and RS groups, it was possible for the DSES to attempt to demonstrate configural, metric, and scalar invariance between these two groups.

We used multiple-group CFA to compare a series of nested models, following the Sequential Constraint Imposition approach (Dimitrov, 2010). Given the limitations of using significant $\Delta\chi^2$ as an indicator of invariance (e.g., sensitivity to sample size), we used Chen's (2007) recommended cutoff for ME/I analyses using an adequate sample ($N > 300$): "For testing [metric] invariance, a change of $\geq -.01$ in CFI, supplemented by a change of $\geq .015$ in RMSEA or a change of $\geq .030$ in SRMR would indicate noninvariance... for testing [scalar] invariance, a change of $\geq -.01$ in CFI, supplemented by a change of $\geq .015$ in RMSEA or a change of $\geq .010$ in SRMR would indicate noninvariance" (p. 501). Thus, noninvariance is present when the CFI and at least one of the other two fit indices (i.e., RMSEA or SRMR) exceed the cutoffs.

First, to test for configural invariance, a baseline model was specified and tested for adequacy of data fit in each of the two groups separately (see above for those results). Second, the fit of the multiple-group model comparing the two samples was investigated. This multiple-group model showed an overall acceptable fit to the data (scaled χ^2 [194] = 881.73, $p < .001$, RMSEA = .077 [90% CI of .072, .082], CFI = .937, TLI = .923, SRMR = .051). Third, to test metric invariance, we compared a fully invariant model where each item loads on its corresponding factor to the same degree in both groups to the previous nested configural model (i.e., each item is allowed to freely load on its corresponding factor without the constraint of forced equality across groups). Results indicated that the DSES was fully metric invariant, per Chen's (2007) criteria (scaled χ^2 [214] = 1012.80, $p < .001$, RMSEA = .079 [90% CI of .074, .084], CFI = .927, TLI = .918, SRMR = .072, Δ CFI = -.010, Δ RMSEA = .002, Δ SRMR = .021). Fourth, to test scalar invariance, we compared a fully invariant model where each item intercept

is set to be equal across groups to the previous nested metric model. When scalar invariance is supported, it can be concluded that the DSES demonstrates strong ME/I (i.e., invariance of all factor loadings and intercepts) across the two groups. Results indicated that the DSES was fully scalar invariant (scaled χ^2 [228] = 1067.23, $p < .001$, RMSEA = .078 [90% CI of .074, .083], CFI = .924, TLI = .920, SRMR = .077, Δ CFI = -.003, Δ RMSEA = -.001, Δ SRMR = .005).

In summary, we can conclude that the DSES demonstrated strong ME/I across the SNR and RS groups, suggesting that the one general DSE factor and the one specific nontheistic factor are measuring the same concepts within these two groups. This made it possible to compare the general and specific latent factor means across the two groups: the RS participants reported a higher general DSE latent factor mean ($M_{diff} = 1.47$, $p < .001$), but a lower specific Nontheistic latent factor mean ($M_{diff} = -.69$, $p < .001$), than did the SNR participants.

Relationship Between DSES and Well-Being

We next examined the relationship between the DSES factors and common measures of well-being (i.e., meaning in life, satisfaction with life, physical health, and mental health). As noted above, we were unable to identify a factor structure that demonstrated adequate fit for the DSES for use with NRS respondents. However, we decided to use the NRS group's three correlated factors CFA model, which had some rough factor parallels with the SNR and RS groups' bifactor CFA model, in order to facilitate comparison of how the DSES factors relate to well-being across the three groups.

We tested one structural equation model per group. For the NRS model, the DSES items were set to load on the factors specified by the three correlated factors CFA model (see Not Religious nor Spiritual Group section above), the well-being instruments' items were set to load onto their respective well-being latent factors, and the four well-being factors were

simultaneously regressed onto the three DSES factors. For the SNR and RS groups, the DSES items were set to load on the two factors specified by the bifactor CFA model, the well-being instruments' items were set to load onto their respective well-being latent factors, and the four well-being factors were simultaneously regressed onto the two DSES factors. Cohen's (1988) guidelines were used to interpret small ($\beta = .20$), medium ($\beta = .50$), and large ($\beta = .80$) effect sizes.

Three patterns are worth noting (see Table 3). First, for all three groups, the Nontheistic factor demonstrated a small to medium positive relationship ($M_{\beta} = .43$) with well-being. Second, the general DSE factor (i.e., defined by both theistic and nontheistic DSE) demonstrated a small negative to small positive relationship ($M_{\beta} = -.03$) with well-being among the SNR and a small to medium positive relationship ($M_{\beta} = .14$) with well-being among the RS. Third, the Theist A correlated factor demonstrated a nonexistent to small negative relationship ($M_{\beta} = -.13$) with well-being among the NRS. We also conducted post-hoc multiple regression analyses using SPSS to examine how the raw mean score for the 16 DSES items relates to the raw mean scores of the four well-being measures. For all three groups, the DSES raw mean score was generally positively associated with well-being for all three groups (see Table 3).

Discussion

Researchers have used the DSES with people from across the R/S spectrum, typically operating on the premise that it measures a single DSE factor that has the same, empirically-comparable meaning for all respondents. We sought to verify this is the case. Our dimensionality results with those who are not religious nor spiritual (NRS), spiritual but not religious (SNR), and religious and spiritual (RS) indicated that the DSES did not measure a unidimensional DSE factor within any of these groups. Rather, the DSES demonstrated a

problematic correlated factors structure among the NRS, and a bifactor structure among the SNR and RS. This deviates from the conclusions of most factor analytic studies of the DSES (Underwood, 2011), but has some parallels with the multidimensional findings of four research teams (e.g., Currier et al., 2012). Our use of bifactor modeling allowed us to examine the dimensionality of the DSES with greater precision than past studies, which are characterized by contradictory findings of unidimensionality versus multidimensionality (see literature review). We will discuss the implications of the dimensionality and measurement equivalence/invariance (ME/I) results for the three groups and then discuss the implications of the DSES factors' nuanced relationships with well-being.

The DSES with the NRS

Use of EFA and CFA did not reveal a factor structure that could provide an adequate fit of the DSES data among the NRS. In other words, our data do not provide internal structure evidence of validity for the DSES within this population. The DSES appeared to tap into multiple dimensions, including a Nontheistic factor and two Theistic factors. The Nontheistic factor was completely uncorrelated with the two Theistic factors. Examination of the descriptive item-level statistic revealed the likely reason: more than 93% of NRS participants denied having DSE involving God. This extreme lack of variance occurred despite the instrument's instruction to "substitute another word that calls to mind the divine or holy for you" when the term "God" made the respondent uncomfortable. Thus, these instructions did not foster usable item scores for the theistic items among the NRS respondents. Perhaps this is not surprising, as it is common for NRS respondents to report the absence of a belief in the existence of "the divine or holy" (Kosmin & Keysar, 2009). In fact, participants' free-response feedback (see Supplemental Material) echoed this sentiment (e.g., "God, divine, holy – these do not exist in my lexicon!").

Similar feedback was given about the DSES by Chinese cancer patients, of whom 17% declined to complete the DSES citing difficulty with substituting “God” with another meaningful term or difficulty with separating the concept of spirituality from religiousness (Lo et al., 2016).

Less obvious but important is feedback from participants regarding the nontheistic items (e.g., “On the question about the beauty of ‘creation,’ if you replaced the word with ‘nature’ I would have answered ‘many times a day’.”) This feedback raises the possibility that the nontheistic items may be underestimating the true frequency of spiritual experiences that NRS participants have. This would contradict the suggestion of the scale developers that NRS people can respond to these nontheistic items “in a way that captures their spiritual experiences” (Underwood, 2006, p. 190). In fact, the capacity of culturally-irrelevant item language to lead to underestimation of spirituality constructs among the NRS has been empirically demonstrated (Hammer, Cragun, & Hwang, 2013). Thus, even if the DSES had demonstrated internal structure evidence of validity, concerns would remain about the cultural relevance of the DSES for the NRS.

Given that the DSES did not demonstrate an identifiable factor structure, the NRS was unable to demonstrate evidence of measurement equivalence/invariance with the other two groups. These results suggest that the DSES measures certain constructs when used with the NRS, but a different set of constructs when used with the SNR or RS. In summary, if these results are replicated in future research, we would conclude that the DSES, in its current form, is not suitable for use with NRS populations. The lack of ME/I further highlights the inappropriateness of comparing DSES scores between the NRS and other groups (e.g., studying whether Christians have more DSE than atheists) or comparing across groups the relations

between the DSES scores and external variables (e.g., studying whether the relationship between DSE and health is moderated by religious vs. secular identity).

The DSES with the SNR and RS

For both the SNR and RS, the DSES dimensionality was best modeled with a bifactor solution, characterized by a general DSE factor defined by a mix of theistic and nontheistic experiences as well as a narrower, independent, specific factor defined by nontheistic experiences. One could conceptualize this general factor as measuring these participants' self-reported "relationship with and awareness of the divine or transcendent" (Underwood, 2006, p. 182), a common thread that runs through the theistic and nontheistic items alike.

However, it is more difficult to conceptualize the specific factor. The stock answer is that the specific factor measures nontheistic DSE—experiences that are spiritual, but nontheistic. However, what does this really mean? Underwood (2006) states that "when the word 'spirituality' is taken out of the religious context it can become so vague as to ultimately become meaningless." (p. 184). Thus, when the general factor accounts for the theistic content, the specific factor is only able to account for non-theistic aspects of DSE. Perhaps the specific factor measures "aspects of personal life that include the transcendent, "more than" what we can see or touch or hear" (Underwood, 2006, p. 184). Understood this way, the specific factor would be measuring a spiritual aspect of DSE that is not tied to a higher power, per se.

However, other scholars may offer a different interpretation of what this specific factor measures. Schuurmans-Stekhoven (2013) argued that the nontheistic items are merely measures of civility, prosociality, appreciation, trait agreeableness, and/or virtue. Examination of the nontheistic items' language seems to allow such an interpretation. For example, "I feel a deep inner peace or harmony" seems like a candidate item for a life satisfaction or sense of meaning in

life instrument. Likewise, “I feel a selfless caring for others” and “I accept others even when they do things I think are wrong” seem like candidate items for a prosociality instrument. Thus, one might conceptualize the specific factor as an artificial amalgam of “positive psychological” (Schuurmans-Stekhoven, 2013, p. 134) experiences that form an empirical factor not because they measure a singular DSE construct, but because people who report positive psychological experiences of one kind will also tend to report other positive psychological experiences of other kinds. The empirical data cannot tell researchers how to conceptualize and name a factor; researchers must engage in this subjective exercise on the basis of theory and conceptual analysis. Thus, there appears room for debate on what this specific factor is truly measuring.

This debate is further complicated by the fact that the construct of DSE has seemingly been given permission by researchers to demonstrate significant content domain overlap with other established constructs. Each DSES item could likely be used to help measure an established psychological construct such as prosociality and spiritual well-being. However, most researchers using the DSES seem to operate on the premise that, when all these items are inserted into a single 16-item instrument, these items no longer measure those other established psychological constructs, but instead now measure this novel construct called DSE. Thus, when we attempt to conceptualize the meaning of this specific factor among the SNR and RS respondents, we struggle to convincingly answer the question of whether these nontheistic items measure a cohesive DSE-related construct or are operating as an amalgam of positive psychological experiences. Incorporation of the well-being criterion results (discussed below) cannot resolve this debate one way or the other. Whether the specific Nontheistic factor is understood as measuring (a) nontheistic DSE or (b) an amalgam of positive psychological experience, in both cases we would predict a positive association between the specific

Nontheistic factor and these well-being criteria. Thus, we are unaware of a practical way to resolve this ambiguity, as any attempt to redevelop the nontheistic items would likely be fraught with the same content domain-overlap pitfalls.

Setting aside the issues with the specific factor for now, the scalar invariance demonstrated by the DSES across SNR and RS respondents is important. The implication is that, for these two groups, the DSES had a similar theoretical structure and meaning, relations between the DSES and external variables could validly be compared, and mean differences in the DSES factors could validly be compared. Comparison of mean differences on the general DSE factor indicated that the RS participants reported more frequent experiences of relationship with and awareness of the divine and transcendent than the SNR participants. However, SNR participants reported a greater frequency of nontheistic DSE than the RS; the meaning of this depends on one's interpretation of the meaning of the specific factor (see above). In summary, when used with SNR and RS respondents, the DSES appeared to measure two independent DSE factors that have the same theoretical structure and meaning across the two groups.

However, when an instrument demonstrates a bifactor structure, scoring procedures can become more complicated. Best practices entail using structural equation modeling (SEM) to model the general DSE factor and the specific Nontheistic factor, as this allows precise measurement of these orthogonal factors. When SEM is not available, users can use an ipsatization approach to partial out the unique variance due to the specific Nontheistic factor from the variance due to the general DSE factor. Ipsatization involves subtracting each respondent's score on each of the 16 DSES items by that respondent's mean DSES total score, resulting in 16 ipsatized item scores whose values represent deviations from that respondent's mean DSES score (Greer & Dunlap, 1997). However, interpretation of ipsatized scores is more

difficult, as a positive score for an ipsatized item would, for example, indicate that respondents scored higher on that item *relative to their average score*. Normally, calculation of ancillary bifactor measures would help researchers determine if calculating a raw total score for the general DSE factor and raw subscale score for the specific Nontheistic factor is permissible (Rodriguez, Reise, & Haviland, 2016), but the lack of a specific factor representing the DSES theistic items precludes this. Thus, users who wish to follow the typical practice of calculating raw total and subscale scores must be aware of the pitfalls of doing so. Specifically, a raw total score calculated by averaging the scores of all 16 items will partially measure three things: the general DSE factor, the specific Nontheistic factor, and error variance. A raw subscale score calculated by averaging the scores of the 6 Nontheistic items will likewise partially measure those same three things, particularly since the specific Nontheistic factor is less well-defined (i.e., lower item-factor loadings) as the general DSE factor. Therefore, when a raw DSES score is used, users must be aware that this score is measuring more than the one intended factor, thereby compromising the content validity of the score to some difficult-to-quantify degree.

Relationship Between DSES and Well-Being

Participants who reported more frequent nontheistic DSE reported better well-being. This was true across all three groups, regardless of whether the nontheistic items loaded on a nontheistic correlated factor (NRS) or a nontheistic specific factor (SNR and RS). In contrast, the relationship between theistic DSE and well-being was dependent on group membership. Except for physical health, theistic DSE (tapped by the mixed nontheistic/theistic experiences general DSE factor) and well-being were positively associated among the RS. But among the NRS (via the Theistic A and B correlated factors) and SNR (via the mixed nontheistic/theistic experiences general DSE factor), experiencing more frequent theistic DSE was not associated

with enhanced well-being. This difference in predictive ability between the nontheistic and theistic items aligns with the findings of Schuurmans-Stekhoven (2013), who found that the nontheistic items were strongly associated with well-being while the theistic items were not. Thus, claims that theistic DSE are good for everyone, regardless of worldview, would be challenged by these findings. This has important practice implications. Some health professionals may assume that increasing all kinds of DSE is good for everyone. However, creating policy, programs, or interventions designed to help people increase all forms of DSE, or theistic DSE in particular, may be ineffectual for the SNR and even harmful for the NRS.

The post-hoc well-being multiple regression results highlight a potential pitfall of using a raw mean/total score for the DSES (which requires evidence of unidimensionality) when the instrument actually demonstrates multidimensionality in a sample. If we assumed the DSES was unidimensional and thus proceeded to use the DSES raw mean score (as most users do), we may have concluded that “experiences of relationship with, and awareness of, the divine or transcendent” (Underwood, 2011, p. 30) are associated with well-being for even the NRS. This would have been misleading, given the nonexistent or negative relationships between theistic experiences and well-being indicators documented among the NRS in Table 3. This highlights the issues that can arise when theistic and nontheistic experiences are lumped into a single raw mean score that is assumed to measure a unitary construct for all people.

Addressing Limitations through Future Research

The present findings should be interpreted with respect to several key limitations. First, the convenience sample means that participants in the three analyzed groups are not demographically representative of their respective national populations. For example, the overall sample was skewed female, white, and well-educated. Furthermore, self-identified atheists were

overrepresented in the NRS group. Thus, until replication of results is established, these findings should be treated as proof-of-concept evidence regarding DSES psychometrics across these R/S groups, rather than a definitive statement. Additional research focused specifically on the performance of the DSES across these groups among people of color, those with less formal education, NRS persons who don't self-identify as atheists, and those living in other countries is particularly encouraged.

Second, there are many ways to categorize people into groups on the basis of R/S identity, worldview, and behavior. Each categorization strategy has strengths and weaknesses. Because (a) we were interested in the performance of DSES across different degrees of religious and spiritual self-identity, (b) the DSES is frequently used on the assumption that it is valid for the study and comparison of most any group, (c) the Overall Self-Ranking items from the Brief Multidimensional Measure of Religiousness/Spirituality (Abeles et al., 1999) are one of the most simple and common ways for researchers to stratify people on degree of self-identified religiousness and spirituality, and (d) there is demonstrated interest among researchers and practitioners to use instruments like the DSES to understand secular individuals and the spiritual but not religious (see literature review), we chose to create three groups using participants Overall Self-Ranking responses. Limitations of this approach include: these groups do not capture multidimensional and nuanced differences in R/S identity, beliefs, and behavior that may be detectable with multidimensional R/S instruments; some scholars may be more interested in comparing psychometrics across specific affiliation groups such as Atheists, Buddhists, Catholics, and Protestant Christians; and one could argue for limiting inclusion into the SNR and RS groups to only those who indicate, at minimum, "moderately" when they rate themselves on religiousness and spirituality. Thus, interested researchers are encouraged to use different

methodologies to form groups of interest, and cross-check resultant finding with this investigation's. We should also stress that the DSES has been described as applicable across various R/S groups (Underwood, 2011). Thus, the onus is on the DSES to demonstrate its purported unidimensional structure across R/S groups, regardless of whether they were formed using one method or another. Third, guidelines for conducting bifactor modeling are rapidly evolving (Rodriguez et al., 2016), so we offer the present results with humility and the understanding that our conclusions are subject to future clarification.

Conclusion

Our results indicate that the DSES measured the same two independent DSE constructs when used with SNR and RS participants, but a different array of inadequately-defined DSE constructs among the NRS participants. This raises questions about the utility of using the DSES with NRS samples, and provides evidence that the DSES conforms to a bifactor, rather than unidimensional, structure among SNR and RS participants. This bifactor structure implies that SEM or ipsatization should be used to score the DSES, given the inaccuracy that would be introduced by using traditional raw total/subscale scoring. Lastly, our findings paint a complex picture of the relationship between DSE and well-being across these three R/S groups, such that DSE is not always associated with enhanced well-being.

References

- Abeles, R., Ellison, C., George, L., Idler, E., Krause, N., Levin, J., & Williams, D. (1999). Multidimensional measurement of religiousness/spirituality for use in health research. A *Report of the Fetzer Institute/National Institute of Aging Working Group*. Kalamazoo, MI: Fetzer Institute.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*, 464–504. doi: 10.1080/10705510701301834
- Cheung, G. W., & Lau, R. S. (2012). A direct comparison approach for testing measurement invariance. *Organizational Research Methods, 15*(2), 167-198. doi: 10.1177/1094428111421987
- Cho, S. J., Wilmer, J., Herzmann, G., McGugin, R. W., Fiset, D., Van Gulick, A. E., ... & Gauthier, I. (2015). Item response theory analyses of the Cambridge Face Memory Test (CFMT). *Psychological Assessment, 27*(2), 552-566. <http://dx.doi.org/10.1037/pas0000068>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Currier, J. M., Kim, S. H., Sandy, C., & Neimeyer, R. A. (2012). The factor structure of the Daily Spiritual Experiences Scale: Exploring the role of theistic and nontheistic approaches at the end of life. *Psychology of Religion and Spirituality, 4*(2), 108-122. doi: 10.1037/a0027710
- Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71-75. https://doi.org/10.1207/s15327752jpa4901_13

- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development, 43*(2), 121-149.
<https://doi.org/10.1177/0748175610373459>
- Dong, M., Wu, S., Zhu, Y., Jin, S., & Zhange, Y. Secular examination of spirituality-prosociality association: Survey research in nonreligious-based populations in China. *Archive for the Psychology of Religion, 39*(1), 61-81. doi: 10.1163/15736121-12341332
- Greer, T., & Dunlap, W. P. (1997). Analysis of variance with ipsative measures. *Psychological Methods, 2*(2), 200-207.
- Hammer, J. H., Cragun, R. T., & Hwang, K. (2013). Measuring spiritual fitness: Atheist military personnel, veterans, and civilians. *Military Psychology, 25*(5), 438-451. doi: 10.1037/mil0000010
- Hays, R. D., Bjorner, J. B., Revicki, D. A., Spritzer, K. L., & Cella, D. (2009). Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. *Quality of Life Research, 18*(7), 873-880. doi: 10.1007/s11136-009-9496-9
- Kalkstein, S., & Tower, R. B. (2009). The Daily Spiritual Experiences Scale and well-being: Demographic comparisons and scale validation with older Jewish adults and a diverse internet sample. *Journal of Religion and Health, 48*(4), 402-417. doi: 10.1007/S10943-008-9203-0
- Kim, S. H., Martin, B. J., & Nolty, A. T. (2016). The factor structure and measurement invariance of the daily spiritual experiences scale. *The International Journal for the Psychology of Religion, 26*(3), 240-251.

- Koenig, H. G., George, L. K., & Titus, P. (2004). Religion, spirituality, and health in medically ill hospitalized older patients. *Journal of the American Geriatrics Society*, *52*(4), 554-562. <https://doi.org/10.1080/10508619.2015.1029404>
- Kosmin, B., & Keysar, G. (2009). *American religious identification survey 2008 (ARIS 2008): Summary report*. Hartford, CT: Trinity College.
- Lace, J. W., & Handal, P. J. (2017). Psychometric properties of the daily spiritual experiences scale: Support for a two-factor solution, concurrent validity, and its relationship with clinical psychological distress in university students. *Religions*, *8*(7), 123-140. doi: 10.3390/rel8070123
- Lo, G., Chen, J., Wasser, T., Portenoy, R., & Dhingra, L. (2016). Initial validation of the Daily Spiritual Experiences Scale in Chinese immigrants with cancer pain. *Journal of Pain and Symptom Management*, *51*(2), 284-291. <https://doi.org/10.1016/j.jpainsymman.2015.10.002>
- Miller, M. J., & Sheu, H. (2008). Conceptual and measurement issues in multicultural psychology research. In Brown, S. D., & Lent, R. W. (Eds.), *Handbook of Counseling Psychology* (4th ed.) (pp. 103-120). Hoboken, NJ: John Wiley & Sons.
- Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide (7th ed.)*. Los Angeles, CA: Authors.
- Ng, S. M., Fong, T. C., Tsui, E. Y., Au-Yeung, F. S., & Law, S. K. (2009). Validation of the Chinese version of Underwood's Daily Spiritual Experience Scale—transcending cultural boundaries?. *International Journal of Behavioral Medicine*, *16*(2), 91-97. doi: 10.1007/s12529-009-9045-5

- Parent, M. C. (2013). Handling item-level missing data: Simpler is just as good. *The Counseling Psychologist, 41*(4), 568-600. doi: 10.1177/0011000012445176
- Pavot, W., & Diener, E. (2008). The satisfaction with life scale and the emerging construct of life satisfaction. *The Journal of Positive Psychology, 3*(2), 137-152. doi: 10.1080/17439760701756946
- Peterson, R. A. (2000). A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Marketing Letters, 11*(3), 261-275.
- Rakošec, Ž., Mikšić, Š., Juranić, B., & Batinić, L. (2015). Psychometric characteristics of Croatian version of the Daily Spiritual Experience Scale. *Religions, 6*(2), 712-723. doi: 10.3390/rel6020712
- Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research, 47*(5), 667-696. doi: 10.1080/00273171.2012.715555
- Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016a). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137-150.
<http://dx.doi.org/10.1037/met0000045>
- Sánchez, E. G. M., Arocena, F. A. L., & Ceballos, J. C. M. (2010). Daily spiritual experience in Basques and Mexicans: A quantitative study. *Journal of Transpersonal Research, 2*(1), 10-25.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research, 99*(6), 323-338. doi: 10.3200/JOER.99.6.323-338
- Schuermans-Stekhoven, J. B. (2013). "As a shepherd divideth his sheep from the goats": Does the Daily Spiritual Experiences Scale encapsulate separable theistic and civility

- components?. *Social Indicators Research*, 110(1), 131-146. doi: oi.org/10.1007/s11205-011-9920-8
- Steger, M. F., Frazier, P., Oishi, S., & Kaler, M. (2006). The meaning in life questionnaire: Assessing the presence of and search for meaning in life. *Journal of Counseling Psychology*, 53(1), 80-93. doi: 10.1037/0022-0167.53.1.80
- Underwood, L. G. (2017). *Summary of Daily Spiritual Experiences Scale (DSES) studies*. Retrieved from <http://www.dsesclear.org/published-dses-studies/>
- Underwood, L. G. (2011). The daily spiritual experience scale: overview and results. *Religions*, 2(1), 29-50. doi: 10.3390/rel2010029
- Underwood, L. G. (2006). Ordinary spiritual experience: Qualitative research, interpretive guidelines, and population distribution for the Daily Spiritual Experience Scale. *Archive for the Psychology of Religion*, 28(1), 181-218. doi: 10.1163/008467206777832562
- Underwood, L. G., & Teresi, J. A. (2002). The daily spiritual experience scale: Development, theoretical description, reliability, exploratory factor analysis, and preliminary construct validity using health-related data. *Annals of Behavioral Medicine*, 24(1), 22-33. https://doi.org/10.1207/S15324796ABM2401_04
- Weston, R., & Gore, P. A. (2006). A brief guide to structural equation modeling. *The Counseling Psychologist*, 34, 719-751. doi: 10.1177/0011000006286345
- Zemore, S. E., & Kaskutas, L. A. (2004). Helping, spirituality and Alcoholics Anonymous in recovery. *Journal of Studies on Alcohol*, 65(3), 383-391. <https://doi.org/10.15288/jsa.2004.65.383>

Table 1

Exploratory Factor Analysis Standardized Loadings for the DSES

#	Item	NRS EFA				SNR Bifactor EFA			RS Bifactor EFA		
		Factor 1: Testlet	Factor 2: Nontheistic	Factor 3: Theistic A	Factor 4: Theistic B	General Factor: General DSE	Specific Factor: Testlet	Specific Factor: Nontheistic	General Factor: General DSE	Specific Factor: Nontheistic	Specific Factor: Testlet
1	I feel God's presence.	.04	.02	.70	.15	.88	-.20	-.03	.86	-.05	.00
2	I experience a connection to all of life.	-.01	.62	-.09	.11	.33	.18	.54	.49	.55	.04
3	During worship, or at other times when connecting with God, I feel joy which lifts me out of my daily concerns.	.03	.02	.12	.61	.81	-.02	-.06	.80	-.01	.13
4	I find strength in my religion or spirituality.	.91	-.01	-.04	.00	.77	.56	.00	.85	.00	.49
5	I find comfort in my religion or spirituality.	.97	-.03	-.01	-.02	.79	.58	.00	.83	.00	.44
6	I feel deep inner peace or harmony.	.04	.56	-.20	.06	.33	.13	.45	.66	.36	.12
7	I ask for God's help in the midst of daily activities.	-.07	-.01	.61	-.11	.78	-.19	-.12	.77	-.25	.05
8	I feel guided by God in the midst of daily activities.	-.03	.00	-.09	.84	.85	-.20	.00	.88	-.14	-.02
9	I feel God's love for me, directly.	.03	-.06	.42	-.01	.89	-.24	.02	.91	-.13	-.05
10	I feel God's love for me, through others.	-.01	.06	.62	-.09	.84	-.21	.01	.77	-.01	-.07
11	I am spiritually touched by the beauty of creation.	.30	.21	.11	-.03	.56	.11	.39	.69	.35	-.02
12	I feel thankful for my blessings.	.14	.43	.12	-.01	.40	-.01	.54	.63	.38	-.01
13	I feel a selfless caring for others.	-.02	.68	.00	-.07	.20	-.08	.60	.47	.38	-.09
14	I accept others even when they do things I think are wrong.	-.07	.58	.13	-.08	.15	-.10	.40	.35	.32	-.12
15	I desire to be closer to God or in union with the divine.	.01	-.02	.42	.21	.74	-.02	-.12	.71	-.27	.00
16	In general, how close do you feel to God?	-.02	-.02	-.03	.49	.84	-.14	-.04	.78	-.07	.00

Note: NRS = Not Religious nor Spiritual, SNR = Spiritual but Not Religious, RS = Religious and Spiritual, EFA = Exploratory Factor Analysis, DSES = Daily Spiritual Experiences Scale, DSE = daily spiritual experiences. Loadings are standardized and based on MLR estimation. Loadings > .32 are bolded.

Table 2

Means, Standard Deviations, and Modes for DSES Items

#	Item	NRS			SNR			RS		
		<i>M</i>	<i>SD</i>	<i>Mode</i>	<i>M</i>	<i>SD</i>	<i>Mode</i>	<i>M</i>	<i>SD</i>	<i>Mode</i>
1	I feel God's presence.	1.05	0.25	1	1.99	1.39	1	3.88	1.510	5
2	I experience a connection to all of life.	3.33	1.54	2	3.82	1.40	3	4.01	1.281	4
3	During worship, or at other times when connecting with God, I feel joy which lifts me out of my daily concerns.	1.10	0.44	1	1.95	1.30	1	3.75	1.406	4
4	I find strength in my religion or spirituality.	1.35	0.96	1	2.62	1.58	1	4.31	1.331	5
5	I find comfort in my religion or spirituality.	1.36	0.94	1	2.64	1.60	1	4.34	1.353	5
6	I feel deep inner peace or harmony.	3.24	1.40	4	3.37	1.33	4	3.87	1.285	4
7	I ask for God's help in the midst of daily activities.	1.06	0.34	1	1.79	1.30	1	3.86	1.622	6
8	I feel guided by God in the midst of daily activities.	1.02	0.22	1	1.63	1.18	1	3.39	1.538	3
9	I feel God's love for me, directly.	1.01	0.08	1	1.71	1.36	1	3.72	1.672	5
10	I feel God's love for me, through others.	1.07	0.36	1	1.94	1.39	1	3.75	1.478	4
11	I am spiritually touched by the beauty of creation.	2.02	1.37	1	3.53	1.70	5	4.48	1.301	5
12	I feel thankful for my blessings.	3.23	1.70	1	4.28	1.38	4	4.96	1.022	6
13	I feel a selfless caring for others.	3.99	1.28	4	4.14	1.11	4	4.33	1.136	4
14	I accept others even when they do things I think are wrong.	3.98	1.18	4	4.03	1.06	4	4.15	1.056	4
15	I desire to be closer to God or in union with the divine.	1.12	0.50	1	2.39	1.69	1	4.36	1.530	5
16	In general, how close do you feel to God?	1.060	0.48	1	1.78	1.23	1	3.44	1.2807	4.3

Note: NRS = Not Religious nor Spiritual, SNR = Spiritual but Not Religious, RS = Religious and Spiritual, DSES = Daily Spiritual Experiences Scale.

Because item 16 was rescaled to match the Likert scaling of the other 15 items, the item responses are not limited to whole numbers (e.g., 4.3 for RS).

DAILY SPIRITUAL EXPERIENCES SCALE

Table 3

Standardized Betas Quantifying the Relationship between DSES and Well-Being

Group	Predictor	Meaning in Life	Satisfaction with Life	Mental Health	Physical Health
NRS					
	Nontheistic Correlated Factor	.56	.37	.42	.20
	Theistic A Correlated Factor	-.21	-.18	-.11	-.03
	Theistic B Correlated Factor	.08	-.06	-.04	-.04
	DSES Raw Mean Score	.30	.21	.21	.11
SNR					
	Nontheistic Specific Factor	.60	.52	.56	.28
	DSE General factor	.12	-.03	-.06	-.13
	DSES Raw Mean Score	.24	.11	.11	-.04
RS					
	Nontheistic Specific Factor	.45	.42	.49	.24
	DSE General factor	.39	.15	.13	-.12
	DSES Raw Mean Score	.41	.20	.20	-.07

Note: DSES = Daily Spiritual Experiences Scale, NRS = Not Religion nor Spiritual, SNR = Spiritual but Not Religious, RS = Religious and Spiritual. Standardized betas are based on MLR estimation, with the exception of those for the DSES raw mean scores, which were derived from SPSS multiple regression analyses. All bolded standardized betas were significant at $p < .05$.