

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Educational Psychology Papers and
Publications

Educational Psychology, Department of

2020

Does the Spiritual Values/Religion Subscale of the Self-Description Questionnaire III Function Differentially Across Heterosexual and Non-Heterosexual Young Adults? A Measurement Invariance Study

Thai Q. Ong

Deborah L. Bandalos

Susan M. Swearer

Follow this and additional works at: <https://digitalcommons.unl.edu/edpsychpapers>



Part of the [Child Psychology Commons](#), [Cognitive Psychology Commons](#), [Developmental Psychology Commons](#), and the [School Psychology Commons](#)

This Article is brought to you for free and open access by the Educational Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Educational Psychology Papers and Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Does the Spiritual Values/Religion Subscale of the Self-Description Questionnaire III Function Differentially Across Heterosexual and Non-Heterosexual Young Adults? A Measurement Invariance Study

Thai Q. Ong, MA,¹ Deborah L. Bandalos, PhD,¹
& Susan M. Swearer, PhD²

¹ Department of Graduate Psychology, James Madison University,
Harrisonburg, Virginia, USA

² Department of Educational Psychology, University of Nebraska-
Lincoln, Lincoln, Nebraska, USA

Corresponding author: Thai Q. Ong, Department of Graduate Psychology,
James Madison University, Harrisonburg, VA 22807, USA; ongtq@dukes.jmu.edu

Abstract

We evaluated the dimensionality and measurement invariance of the Spiritual Values/Religion (SVR) subscale from the Self-Description Questionnaire III across heterosexual and non-heterosexual young adults. We found a one-factor model provided adequate fit to the data for each group, with the SVR items exhibiting configural, metric, and scalar invariance across the two groups. Given that we established measurement invariance, we examined the latent mean difference on the construct and found the heterosexual group reported significantly higher levels of spiritual value/religion than the non-heterosexual group. Our results provided empirical support for the theorized factor structure of the SVR items and the use of the SVR subscale across heterosexual and non-heterosexual young adults, making the scale a viable option for researchers studying religiosity in these specific subpopulations.

Keywords: Religiosity, measurement invariance, non-heterosexual, heterosexual, spiritual, SDQ-III, young adults

Published in *Journal of Homosexuality* 67:10 (2020), pp 1367–1385.

DOI: 10.1080/00918369.2019.1591785

Copyright © 2019 Taylor & Francis Group, LLC. Used by Permission.

The role of religiosity on well-being has been a subject of research for many decades. To that end, researchers have found a strong association between adolescent religiosity and positive health outcomes (e.g., Wong, Rew, & Slaikou, 2006). Young adults who reported higher levels of religiosity tended to have overall better health outcomes, such as fewer mental problems and lower risks of cigarette smoking, heavy drinking, and marijuana use, than those with lower levels (Ellison, 1995; Plante & Sherman, 2001; Rew & Wong, 2006; Wills, Yaeger, & Sandy, 2003). Researchers, however, have not found the positive benefits of religiosity to be consistent across all young adults, specifically for those self-identifying as lesbian, gay, bisexual, or queer (LGBQ). For example, Rostosky, Danner, and Riggle (2007) found religion to act as a buffer against substance use for heterosexual young adults but not for non-heterosexual young adults in a longitudinal study, and Longo, Walls, and Wisneski (2013) found non-heterosexual young adults who identified as Christians had greater psychosocial risks (e.g., self-harming) compared to heterosexual young adults who identified as Christians. These findings suggest that the positive benefits from religiosity may be moderated by one's sexual orientation.

Potential differences in conceptualizations of religiosity

Because religiosity has not been clearly conceptualized in the literature (Zinnbauer, Pargament, & Scott, 1999), one explanation for the differential role of religiosity may be due to differences in how heterosexual and non-heterosexual young adults conceptualize religiosity. For example, because non-affirming religious denominations tend to view same-sex behaviors as sinful (Clark, Brown, & Hochstein, 1990; Morrow, 2003; Sherkat, 2002), non-heterosexual young adults may view religion as a stressor, particularly during the period of confirming a new sexual self-identify. This stress, associated with conflict regarding religion, may serve to offset the positive health outcomes associated with religion. In fact, non-heterosexual young adults were less likely to report a religious affiliation compared to their heterosexual peers and had greater declines in religiosity from early adolescence to young adulthood (Rostosky, Danner, & Riggle, 2008). These findings suggest that, at least for some non-heterosexual young adults, the stress associated with reconciling religious beliefs with emerging

sexual identity may lead to a move away from religious beliefs, which could ultimately lead non-heterosexual young adults to conceptualize religiosity less positively than their heterosexual counterparts.

The stress from being affiliated with a non-affirming religious denomination could serve as a foundation for internalized homophobia and shame. For example, Ream and Savin-Williams (2005) found that those whose religion made it difficult to accept their sexual identity reported higher internalized homophobia. Eliason, Burke, van Olphen, and Howell (2011) found that lesbian, gay, and bisexual (LGB) young adults who were high in religiosity and who were not comfortable labeling their sexual identity were at the greatest risk for negative drinking outcomes. It may be that these non-heterosexual young adults had extreme difficulty reconciling their sexual orientation with their religious beliefs, such that religion became a source of stress rather than comfort. Thus, in contrast to heterosexual young adults, non-heterosexual young adults may view religiosity as a negative, rather than a positive, component of their lives.

Given that heterosexual and non-heterosexual young adults may differ in their conceptualizations of religiosity and that religiosity is related to positive health outcomes, researchers studying religiosity in heterosexual and non-heterosexual young adults should ensure measurement scales used to measure religiosity function equivalently across the two populations. This is important for two reasons. From a substantive standpoint, if heterosexual and non-heterosexual young adults conceptualize religiosity differently, the use of the same religiosity scale to conduct group comparisons may not be appropriate because the items on the scale may have different relative meaning and importance for heterosexual and non-heterosexual young adults. From a statistical standpoint, if heterosexual and non-heterosexual young adults conceptualize religiosity differently and the same religiosity scale is used to measure the construct, the scores obtained from the scale may represent different conceptualizations of religiosity. Any mean differences based on these scores would be biased, to some extent, by differences in the meaning of scores across groups. Similarly, any comparisons of correlations with external variables across heterosexual and non-heterosexual young adults based on the scores may also be invalid.

Because of the two reasons above, researchers should determine whether measurement scales are invariant across subpopulations of

interest (e.g., heterosexual and non-heterosexual young adults) when studying certain constructs that may differ in meaning across these subpopulations (e.g., religiosity). If measurement invariance is established, researchers can be more confident that any group differences in the scores are due to true differences on the construct and not due to differences in the way the two groups interpret or respond to the items. We therefore evaluated the comparability of scores on a religiosity scale across groups of heterosexual and non-heterosexual young adults in the current study.

Spiritual values/religion subscale

One religiosity scale that has not been heavily studied is the Spiritual Values/ Religion (SVR) subscale from the Self-Description Questionnaire III (SDQ-III; Marsh, 1992). The SDQ-III is one of three measures developed to assess self-concept for the young adult population, which conceptualize self-concept as a multidimensional construct consisting of 13 underlying dimensions, as measured by 13 subscales. The SVR subscale consists of 12 items and is used to measure a unidimensional construct of religiosity (see Figure 1 for wording of all 12 items), which is thought to be one of the dimensions of a person's self-concept. Participants are assumed to fall along a single continuum of religiosity ranging from high to low religiosity, and a single total score is calculated for each participant. Calculation of a single score is based on the assumption that the items from the SVR subscale are unidimensional and reflect a single construct of religiosity. Researchers interested in comparing scores across different subpopulations must also assume that the SVR items function equivalently across different subpopulations. Although these assumptions may be valid, they have not been tested empirically in the literature.

Measurement invariance of spiritual values/religion subscale

Based on previous research on the role of religiosity for heterosexual and non-heterosexual young adults, there are several reasons why one would hypothesize that heterosexual and non-heterosexual young adults would conceive of religiosity, as measured by the SVR subscale, in different ways.

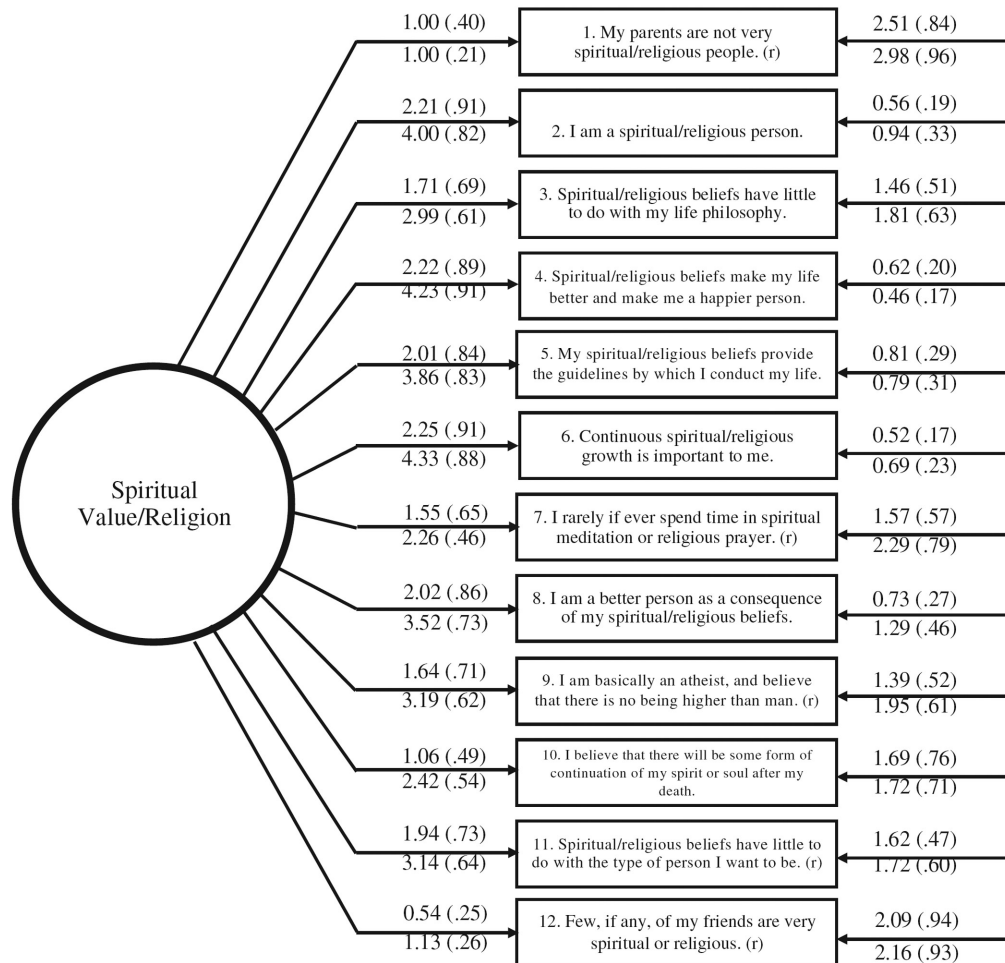


Figure 1. Unstandardized factor pattern coefficients, standardized factor pattern coefficients, and error variances for the one-factor model were estimated separately for the two groups. Values for the heterosexual group are above the arrows and values for the non-heterosexual group are below the arrows. Standardized parameter estimates are in parentheses. The estimated factor variances for the two groups were 0.481 and 0.120 for the heterosexual group and nonheterosexual group, respectively. 24 parameters were estimated in the one-factor model for each group: 11 factor pattern coefficients, 12 error variances, and 1 factor variance. The total number of observations was 78. Therefore, the degrees of freedom were 54.

These differences in perceptions could be manifested in three different ways when investigating how the scale functions across the two groups (measurement invariance), termed *configural*, *metric*, and *scalar invariance* in the psychometric literature (Cheung & Rensvold, 2002).

Configural invariance

First, non-heterosexual young adults may feel more stigmatized and oppressed by mainstream religious denominations and, as a result, may have a different general conceptualization of religiosity than heterosexual young adults. For example, heterosexual young adults may conceptualize religiosity as a unidimensional construct, whereas non-heterosexual young adults may conceptualize religiosity as a multidimensional construct. If so, this would have large implications for how the SVR scale is used, administered, and scored. Such differences in overall conceptualizations of religiosity would be manifested as differences in factor structures when testing for configural invariance between the two groups. If configural invariance is established, this would indicate that the two groups are conceptualizing religiosity in the same way in the most general sense—the items are unidimensional (as theorized) in structure for both groups (Cheung & Rensvold, 2002).

Metric invariance

Second, given that religiosity is associated with positive health outcomes for heterosexual young adults, heterosexual young adults may find some SVR items to be more salient than non-heterosexual young adults, particularly those regarding the importance and benefits of religiosity. For example, heterosexual young adults may find the item “Continuous spiritual/religious growth is important to me” more salient to them than non-heterosexual young adults because of the potential positive benefits the former group obtains from being involved in religious affiliations and activities. In contrast, non-heterosexual young adults may find the item “Spiritual/religious beliefs make my life better and make me a happier person” less salient because of the potential negative viewpoints about their sexual orientation from different religious affiliations. Such differences in saliency would be manifested as differences in items’ factor loadings when testing for metric invariance between the two groups. If metric invariance is established, this would indicate that the strength, or saliency, of each item’s relation to the factor is similar across the groups. In other words, the 12 items of the SVR subscale relate to the construct in the same manner and with the same strength across groups (Cheung & Rensvold, 2002).

Scalar invariance

Third, even if the two groups have the same underlying level of religiosity, heterosexual and non-heterosexual young adults may endorse certain items more or less strongly. For example, the item “Few, if any, of my friends are very spiritual or religious” could evoke weaker responses from non-heterosexual young adults than heterosexual young adults because even though a non-heterosexual young adult responder may perceive himself or herself as religious, many of his or her friends (other non-heterosexual young adults) may be turned off by religion. The item “Spiritual/religious beliefs have little to do with the type of person I want to be” could also evoke more positive responses from non-heterosexual young adults than heterosexual young adults, even if the two groups have the same level of religiosity, because non-heterosexual young adult responders may associate the type of person they are more with their sexual identity than their religious beliefs. Such differences in the scenarios above would be manifested as differences in items’ intercepts when testing for scalar invariance between the two groups. If scalar invariance is established, this would indicate that the observed mean difference of each item between the groups is reflective of true mean differences at the latent level (Cheung & Rensvold, 2002; Steenkamp & Baumgartner, 1998).

Purpose of study

As mentioned, there has been no effort put forth to confirm the theorized factor structure of the SVR subscale and to test whether the items are invariant across different subpopulations, which limits the use and interpretations of the scores obtained from the SVR subscale. The lack of psychometric work on the SVR subscale is even more problematic in instances when certain subpopulations of interest may differ in conceptualization of the construct, such with heterosexual and non-heterosexual young adults as discussed above. Thus, additional psychometric research on the SVR subscale is warranted. The purpose of our study was to evaluate the psychometric properties of the SVR subscale. Specifically, we had two primary research questions.

Research question 1: How well does a one-factor model fit the SVR data for two independent samples of heterosexual and non-heterosexual young adults?

Given that the SVR items were written to reflect a unidimensional construct of religiosity, the items are expected to fit a one-factor model. If a one-factor model fails to provide a good fit to the data, this would indicate that (1) the SVR items are multidimensional, and (2) a single total score should not be calculated. Even if a one-factor model provides adequate fit to the data, which would support the theorized dimensionality of the SVR subscale, researchers should aim to evaluate the invariance of the SVR subscale across different subpopulations of interest.

Research question 2: What level of invariance (configural, metric, or scalar) do Spiritual Values/Religiosity items exhibit across two independent samples of heterosexual and non-heterosexual young adults?

Although empirical research has suggested that, on average, heterosexual young adults have higher religiosity scores than non-heterosexual young adults (Rostosky et al., 2008), this difference could be an artifact of differences in factor structure or item-level interpretations across groups. Thus, unless measurement invariance can be established on scales used to assess religiosity, such as the SVR subscale, the interpretation of results comparing mean levels of religiosity across groups is ambiguous. If strong measurement invariance (i.e., scalar invariance) is established, however, latent mean differences on the construct between heterosexual and non-heterosexual young adults could be examined.

Method

Participants and procedure

Participants were recruited in 2014 during national and international concert tours headlined by top recording artist Lady Gaga,

through her social media outlets, and via the Born This Way Foundation (BTWF). Participants were handed postcards with the online link to the Born Brave Experiences Survey,¹ which included the SVR items, at the BTWF booth outside of the concert venues. Lady Gaga further recruited participants by publishing a message with the online link to the Born Brave Experiences Survey on her Twitter account encouraging her followers to complete the online survey. The participants logged onto the foundation's Web site, gave consent to participate in the study, and completed the Born Brave Experiences Survey. Data were cleaned for redundant IP addresses, duplicated participant names, and duplicated e-mail addresses. There was a total of 3,433 young adults (11th grade or younger) who completed the SVR subscale. For our study, we analyzed data only from young adults from the United States and excluded those who had incomplete data on the SVR subscale ($n = 1,623$), those who did not report their sexual orientation ($n = 103$), and those who reported living outside of the United States ($n = 1,752$).

Our final sample consisted of 740 heterosexual and non-heterosexual young adults, with 329 self-identifying as heterosexual ($M_{\text{age}} = 20.64$ years, $SD_{\text{age}} = 3.29$, age range: 13–25 years), which included 100% straight young adults, and 411 self-identifying as non-heterosexual ($M_{\text{age}} = 20.16$ years, $SD_{\text{age}} = 3.07$, age range: 13–26 years), which included 11% lesbian, 49% gay, 36% bisexual, and 4% queer young adults. Of the 329 heterosexual young adults, 10% self-identified as male, 89% self-identified as female, and 1% self-identified as genderqueer. More so, 77% self-identified as White, 16% self-identified as mixed (two or more races), 3% self-identified as African American, 2% self-identified as American Indian or Alaska Native, 2% self-identified as Asian, and 4% self-identified as Native Hawaiian or Pacific Islander. Of the 411 non-heterosexual young adults,² 56% self-identified as male, 39% self-identified as female, 1% self-identified as transgender, and 3% self-identified as queer. More so, 72% self-identified as White, 20% self-identified as mixed (two or more races), 4% self-identified as African American, 1% self-identified as American Indian or Alaska Native, and 2% self-identified as Asian.

Measure

The SVR subscale was developed as part of the SDQ-III (Marsh, 1992). The subscale consists of 12 items developed to evaluate respondents' feelings about the importance of religiosity (see Figure 1 for wording of all 12 items). Participants indicated the extent to which they agreed with each statement on a scale of 1 (*strongly disagree*) to 6 (*strongly agree*). Six out of the 12 items on the SVR subscale were reverse-scored such that high scores on the scale imply strong spiritual and religion beliefs. Previous studies administering the SVR subscale have reported adequate reliability values ($\alpha > .80$; Marsh & O'Neill, 1984; Faria, 1996). The reliability values of the SVR items were .91, .92, and .89 for the total group, heterosexual group, and non-heterosexual group, respectively.

Data analysis

We conducted a series of confirmatory factor analyses to assess the invariance of the SVR subscale across heterosexual and non-heterosexual young adults. Prior to testing measurement invariance, we fit a one-factor model to both groups independently and evaluated model fit. We then conducted a series of nested multiple-group CFA models to test the measurement invariance of the SVR subscale across heterosexual and non-heterosexual young adults. Given that the data are approximately normal and that there were five response options, we used maximum likelihood (ML) estimation, which has been found to function best with continuous and normally distributed data (Finney & DiStefano, 2013). ML estimation is more sensitive to model misspecification than other normal theory estimators (Olsson, Foss, Troye, & Howell, 2000; Olsson, Troye, & Howell, 1999). We used LISREL 9.2 (Jöreskog & Sörbom, 2015) for all analyses.

Single-group CFA

Given the lack of psychometric work on the SVR subscale, it was important to evaluate how well the SVR subscale functions for both groups separately prior to testing measurement invariance. Thus, we fit a one-factor CFA model (the theorized factor structure underlying

the SVR subscale) to data from each group independently and evaluated the fit of the model.

Multiple-group CFA

We evaluated measurement invariance of the SVR subscale by conducting a series of nested multiple-group CFA models (i.e., multiple-group invariance models). We followed the invariance testing method recommended by Cheung and Rensvold (2002), where in each step we added equality constraints for different sets of parameters. The model fitting steps are summarized below.

First, we tested for configural invariance, which assesses the underlying factor structure of the SVR across the two groups. In the multiple-group configural invariance model, we allowed all parameter estimates (e.g., factor pattern coefficients, error variances, and factor variances) to be freely estimated across both groups while fitting the same factor structure to both sets of data. The configural model serves as a baseline model for the additional invariance constraints outlined below (Cheung & Rensvold, 2002).

Second, we tested for metric invariance, which assesses the degree of saliency of the SVR items to the religiosity construct for the two groups. In the multiple-group metric invariance model, we constrained the unstandardized factor pattern coefficients (i.e., loadings) for the items to be equal across groups. The metric model is considered nested within the configural model because the metric model can be obtained from the configural model by setting the unstandardized factor pattern coefficients (loadings) to be equal across the two groups.

Third, we tested for scalar invariance, which assesses the observed item means (or intercepts) across the groups. In the multiple-group scalar invariance model, we constrained the item intercepts to be equal across groups in addition to constraining the factor pattern coefficients. The scalar model is considered nested within the metric model because the scalar model can be obtained from the metric model by setting the item intercepts to be equal across groups.

Referent item. An important decision when conducting both single and multiple-group invariance models is model identification. One important aspect of model identification is setting the scale of the latent

factor(s). Because these factors have no inherent metric, it is necessary to provide one by either setting the factor variance to one or setting one of the factor pattern coefficients to one. In single-group CFA models, either approach can be taken. However, in multiple-group CFA models, this decision is often more complicated (Bontempo & Hofer, 2007). For example, setting the factor variance to one for both groups would imply that the factor variances are invariant across groups, which is unlikely. The alternative is to set the loading of the same indicator (i.e., a referent indicator) of the factor to one across both groups. The chosen referent item, however, must be tested for invariance across both groups. In this study, we found the loading of item 1 to be invariant across the two groups.³ Thus, we scaled the factor by fixing item 1's factor pattern coefficient to one for both groups and allowed the factor variances to be freely estimated.

Model fit

There are a variety of fit indices that can be used to evaluate model fit. Researchers should appropriately choose which fit indices to use and provide a justification for each (Hoyle & Panter, 1995). In this study, we evaluated model fit using the model chi-square (χ^2), standardized root mean square residuals (SRMR), comparative fit index (CFI), and root mean square error of approximation (RMSEA). The model χ^2 is a measure of exact model fit to the data and was assessed using a significance test. The model χ^2 is extremely stringent and can be overly sensitive to sample size (Kline, 2011). Thus, it is common practice to examine approximate model-fit statistics as well such as the SRMR, CFI, and RMSEA. The SRMR is a measure of the standardized squared residuals between the predicted correlations and observed correlations; the RMSEA is a measure of the amount of misfit per degree of freedom; and the CFI is a measure of the proportional improvement of fit when compared to a more restricted, baseline model, where all variables are uncorrelated (Kline, 2011).

Hu and Bentler (1998) recommended the use of SRMR paired with another index such as the CFI or RMSEA because the fit indices are sensitive to different model misspecifications. The SRMR is sensitive to misspecified factor correlations, whereas the CFI and RMSEA are sensitive to misspecified factor loadings (Hu & Bentler, 1998). Based on Hu and Bentler's (1998) recommendations, we examined the SRMR,

RMSEA, and CFI with the following guidelines: SRMR \leq .08, RMSEA \leq .06, CFI \geq .95 (Hu & Bentler, 1999). These guidelines, however, were created under very specific conditions that may not be generalizable (Marsh, Hau, & Wen, 2004). Therefore, we also evaluated model fit by examining correlation residuals (i.e., the discrepancy between the model-implied correlations and observed correlations). Ideally, if the CFA model is correctly specified, the correlation residual for each pair of variables should be small. Correlation residuals greater than $|.15|$ were taken to be indicative of local misfit.

Model comparison

Given the more constrained invariance models are nested within the less constrained invariance models, we tested the measurement invariance of the SVR items by comparing the invariance models using the chi-square difference test ($\Delta\chi^2$). Similar to the model χ^2 , however, the $\Delta\chi^2$ may also be too stringent (Kline, 2011; Quintana & Maxwell, 1999). Thus, we adopted guidelines proposed by Cheung and Rensvold (2002) stating that parameter estimate invariance is established if the difference in CFI values (Δ CFI) is less than or equal to 0.1.

We tested for configural invariance by examining the overall model-data fit of the configural model. If the configural model fits the data well, configural invariance is established. We tested for metric invariance by comparing the fit of the configural model to the fit of the metric model based on the Δ CFI criterion. If the metric model fits significantly better than the configural model based on the Δ CFI criterion, metric invariance is established. We tested for scalar invariance by comparing the fit of the metric model to the fit of the scalar model. If the scalar model fits significantly better than the metric model based on the Δ CFI criterion, scalar invariance is established.

Results

Data screening and descriptive statistics

Prior to conducting the CFA analyses, we screened the data for outliers and normality. We found no outliers, and univariate skewness and kurtosis values were less than the recommended cutoff value of

|3| for skewness and |10| for kurtosis in both groups (Kline, 2011). Given the univariate skewness and kurtosis values, we treated our data as normally distributed. Descriptive statistics and correlation matrices for both groups are presented in Table 1.

Single-group CFA

A one-factor model was tested and evaluated for both groups separately prior to testing for measurement invariance. Overall, the one-factor model provided adequate fit for both groups (see Table 2). Although the model's χ^2 values were significant, the fit indices were acceptable in both groups (CFI close to .95, RMSEA close to .06, and SRMR below .06). For the heterosexual group, we found only one large correlation residual between items 9 and 10 (.22). For the non-heterosexual group, we found two large correlation residuals between items 3 and 7 and items 9 and 10 (.16 and .20, respectively). Because the large correlation residual between items 9 and 10 was present in both groups, we examined the two items to determine whether the item wording might be similar. We were not able to identify any theoretical reasons for why the two items shared more variance than other pairs of items, and therefore did not respecify the model. The large correlation residual may be due to an item-order effect. This could be examined in the future by randomizing item order and assessing the stability of the residual. All unstandardized factor pattern coefficients were positive (after reverse coding negatively keyed items) and statistically significant different from 0 ($ps < .05$; see Figure 1), and a small to moderate amount of the items' variances was explained by the factor in both groups (see Table 3). For the heterosexual group, the factor accounted for, on average, 52% of variance in the items, and McDonald's (1999) omega reliability coefficient was .71. For the non-heterosexual group, the factor accounted for, on average, 43% of variance in the items, and McDonald's omega reliability coefficient was .86.

Multiple-group CFAs

Configural invariance model

Although configural invariance was already established because a one-factor model provided good fit to the data for both groups,

Table 1. Descriptive statistics and correlation matrices for heterosexual and non-heterosexual data.

Item	Heterosexual (n = 326)															
	1	2	3	4	5	6	7	8	9	10	11	12	M	SD	Skew	Kurt
1	-	0.399	0.326	0.324	0.317	0.356	0.252	0.332	0.300	0.157	0.335	0.116	3.687	1.726	-0.173	-1.308
2	0.175	-	0.615	0.814	0.724	0.830	0.602	0.749	0.647	0.481	0.630	0.208	3.706	1.706	-0.301	-1.173
3	0.181	0.483	-	0.612	0.604	0.604	0.537	0.585	0.458	0.297	0.614	0.258	3.500	1.689	0.037	-1.236
4	0.168	0.767	0.531	-	0.740	0.812	0.589	0.767	0.640	0.440	0.628	0.183	3.574	1.722	-0.147	-1.243
5	0.181	0.629	0.514	0.759	-	0.786	0.500	0.779	0.547	0.363	0.608	0.194	3.258	1.661	0.096	-1.157
6	0.147	0.718	0.494	0.802	0.757	-	0.593	0.778	0.618	0.427	0.631	0.197	3.469	1.718	-0.067	-1.244
7	0.200	0.427	0.442	0.362	0.372	0.374	-	0.523	0.428	0.268	0.526	0.251	2.936	1.651	0.335	-1.232
8	0.137	0.586	0.434	0.691	0.644	0.628	0.276	-	0.538	0.408	0.638	0.243	3.390	1.643	-0.038	-1.217
9	0.104	0.549	0.422	0.538	0.474	0.540	0.323	0.384	-	0.567	0.548	0.150	4.816	1.637	-1.217	0.192
10	0.081	0.506	0.348	0.469	0.408	0.468	0.285	0.369	0.542	-	0.328	0.103	4.650	1.493	-0.967	-0.081
11	0.100	0.477	0.503	0.560	0.567	0.545	0.410	0.463	0.442	0.256	-	0.312	3.457	1.854	-0.002	-1.453
12	0.151	0.159	0.280	0.220	0.181	0.202	0.244	0.185	0.246	0.062	0.292	-	3.276	1.496	0.214	-0.935
Non-Heterosexual (n = 410)																
M	3.68	3.078	2.971	2.927	2.693	3.027	2.837	2.937	4.22	4.427	2.849	3.127				
SD	1.762	1.695	1.698	1.616	1.609	1.715	1.703	1.667	1.783	1.558	1.703	1.522				
Skew	-0.121	0.213	0.347	0.350	0.512	0.277	-0.520	0.279	-0.545	-0.881	0.458	0.320				
Kurt	-1.348	-1.275	-1.166	-1.070	-1.002	-1.239	-1.044	-1.249	-1.109	-0.209	-1.124	-0.966				

M = mean; SD = standard deviation; Skew = skewness; Kurt = kurtosis. The values above the diagonal represent the correlation matrix for the heterosexual group. The values below the diagonal represent the correlation matrix for the non-heterosexual group.

Table 2. Single-group and multiple-group CFA results.

Model	ML χ^2	df	$\Delta\chi^2$	CFI	Δ CFI	SRMR	RMSEA	RMSEA 90% CI
Single-Group CFA								
Heterosexual	181.119*	54	–	0.951	–	0.044	0.085	(0.072–0.099)
Non-Heterosexual	242.520*	54	–	0.924	–	0.056	0.092	(0.080–0.104)
Multiple-Group CFA								
Configural Invariance ^a	423.068*	108		0.971		0.055	0.089	(0.080–0.098)
Metric Invariance ^b	443.239*	119	20.171*	0.970	–0.001	0.063	0.086	(0.077–0.094)
Scalar Invariance ^c	468.288*	130	25.049*	0.969	–0.001	0.063	0.084	(0.076–0.092)

ML χ^2 = maximum likelihood chi-square; CFI = comparative fit index; SRMR = standardized root mean square residual
 RMSEA = root mean square error of approximation; CI = confidence interval; $\Delta\chi^2$ and Δ CFI tests were conducted between the model and the previous model.

- 48 parameters were estimated in the multiple-group configural invariance model: 22 factor pattern coefficients, 24 error variances, 2 factor variances, and 0 intercepts. The total number of non-redundant covariances was 156. Therefore, the degrees of freedom were 108.
- 37 parameters were estimated in the multiple-group metric invariance model: 11 factor pattern coefficients, 24 error variances, two factor variances, and zero intercepts. The total number of non-redundant covariances was 156. Therefore, the degrees of freedom were 119.
- 50 parameters were estimated in the multiple-group scalar invariance model: 11 factor pattern coefficients, 24 error variances, two factor variances, 12 intercepts, and one latent mean difference.

The total number of non-redundant was 180. Therefore, the degrees of freedom were 130.

*p < .05

Table 3. Variance accounted (R^2) for in the items for the single-group one-factor CFA model.

	Heterosexual	Non-Heterosexual
Item 1	0.16	0.04
Item 2	0.81	0.67
Item 3	0.49	0.37
Item 4	0.80	0.83
Item 5	0.71	0.70
Item 6	0.83	0.77
Item 7	0.43	0.21
Item 8	0.73	0.54
Item 9	0.48	0.39
Item 10	0.24	0.30
Item 11	0.52	0.41
Item 12	0.06	0.06

we estimated the configural invariance model to obtain the pooled fit indices. As expected, the multiple-group configural invariance model provided adequate fit to the data (see [Table 2](#)). The CFI was above the cut-off value of .95, and the RMSEA was near the cut-off value of .06 (Hu & Bentler, 1999), indicating that configural invariance was established. This implies that both groups conceptualized religiosity, as measured by the SVR, as unidimensional in nature. As expected, the same large correlation residuals found in the two separate one-factor models were also present in the configural model.

Metric invariance model

Given that configural invariance was established, we constrained the factor pattern coefficients to be equal across both groups to test metric invariance. The $\Delta\chi^2$ between the configural and metric models was significant, suggesting the metric model fit significantly worse than the configural model. Recall, however, that the $\Delta\chi^2$ is a stringent test of model comparison and should be evaluated in conjunction with other model comparison indices such as the Δ CFI and change in correlation residuals. The Δ CFI between the configural and metric models was minimal ($< .01$; see [Table 2](#); Cheung & Rensvold, 2002). Additionally, we found only one large correlation residual in the non-heterosexual group between items 9 and 10 (.23). Note that this was the same large correlation residual found in the configural model. Given that the CFI difference was less than .01, we determined that metric invariance was established, indicating that the 12 items have equal saliency to the factor across groups.

Scalar invariance model

After metric invariance was established, we added the constraint that item intercepts must be equal across groups to test scalar invariance. Although the $\Delta\chi^2$ between the metric and scalar models was significant, the Δ CFI was very minimal ($< .01$; see [Table 2](#)). This suggested the scalar model did not fit the data practically worse than the metric model. As in the metric invariance testing results, we found only one large correlation residual for the nonheterosexual group between items 9 and 10 (.22). We therefore determined that scalar invariance was established in addition to configural and metric invariance. This indicates that observed mean differences between the two

groups are reflective of true differences on the latent construct of religiosity, as measured by the SVR, between the two groups.

Latent mean difference

Given that configural, metric, and scalar invariance were established across the two groups, we estimated the latent mean difference⁵ between the two groups on the construct. We found the non-heterosexual group had significantly lower levels of spiritual value/religion than the heterosexual group. The unstandardized latent mean difference ($-.195$) is difficult to interpret because it is on the scale of the referent indicator (item 1). To better interpret the latent mean difference, we computed a latent mean effect size. The latent effect size has the same interpretation as Cohen's d (difference in means in standard deviation units); however, the standard deviation difference is on the continuum of the latent construct. The heterosexual group was .37 standard deviation units higher than the non-heterosexual group on the latent continuum of spiritual value/religiosity.

Discussion

There were two purposes to this study. The first purpose was to evaluate how well the SVR subscale functioned for heterosexual and non-heterosexual young adults independently. The second purpose was to gather additional validity evidence for the SVR subscale by testing measurement invariance across heterosexual and non-heterosexual young adults. Our findings have several substantive and measurement implications. In the following sections, we first discuss the implications of our study from a measurement perspective. That is, how do our findings contribute to the current research on the SVR subscale? Then we discuss the implications of our study from a researcher perspective. That is, how do our findings inform researchers currently conducting religiosity research on heterosexual and non-heterosexual young adults?

Implications from the measurement perspective

We found a one-factor model provided adequate fit to the data for both groups, which supported the theorized factor structure of

the SVR subscale. Further, we found the SVR items to exhibit configural, metric, and scalar invariance across both groups. This suggested that (1) heterosexual and non-heterosexual young adults in this sample conceptualized the construct of spiritual value/religion in a similar manner (configural invariance), (2) each item had equal saliency to the construct (metric invariance), and (3) observed mean differences between the two groups were representative of true mean differences at the latent level (scalar invariance; as shown in Table 1). In other words, the scores obtained from the SVR in the two samples represented the same construct of interest, allowing for valid group comparisons based on the scores. Our findings provided empirical evidence for the use of the SVR subscale for heterosexual and non-heterosexual young adults, which has great implications for researchers studying religiosity in these specific subpopulations. For example, researchers interested in comparing SVR scores or exploring how SVR scores correlate with other variables across the two groups should feel more comfortable doing so given the results of our study. Note that if we had instead found the SVR items to lack measurement invariance, it would not be appropriate for researchers to compare scores across heterosexual and non-heterosexual young adults because the scores could represent different constructs.

Implications from the researcher perspective

With configural, metric, and scalar invariance established, we examined the latent mean difference between the two groups on the construct. We found nonheterosexual young adults reported significantly lower levels of spiritual value/ religion than heterosexual young adults. This finding is not surprising given the literature on the role of religiosity for non-heterosexual young adults, thus providing possible validity evidence for the SVR subscale. It is important to consider and explore why this difference in religiosity occurs. One possibility is that religion could provide stress rather than comfort for non-heterosexual groups, especially for those struggling to accept their new sexual identity, which may be at odds with the teachings of their religion (e.g., Newman & Muzzonigro, 1993). This stress could deter non-heterosexual young adults from being involved in religious practices and activities, which would explain why non-heterosexual young adults reported lower levels of spiritual value/religion than heterosexual

young adults. Another possibility is that the stress associated with being affiliated with a non-affirming religious denomination could result in greater internalized homophobia (e.g., Barnes & Meyer, 2012), leading to decreases in spiritual value/ religion. The latter may be one of the reasons that religiosity was found to be related to negative health outcomes rather than positive health outcomes for nonheterosexual young adults. Researchers should further evaluate the underlying mechanisms of why heterosexual young adults have higher levels of religiosity than non-heterosexual young adults, as evident by the SVR subscale.

The SVR subscale conceptualizes religiosity and spirituality as a single construct that underlies a person's self-concept on the SDQ-III. This conceptualization may be outdated given current research on the distinction between religiosity and spirituality (Zinnbauer et al., 1999). We could not, however, test a two-factor CFA model in our study because of how the original items on the scale were written (e.g., "Spiritual/religious beliefs make my life better and make me a happier person"). Thus, it may be plausible that a person's self-concept consists of both spirituality and religion and that these two constructs are distinct aspects of self-concept. Researchers should evaluate the appropriateness of measuring religiosity as a unidimensional construct, as conceptualized by the SVR subscale, versus measuring religiosity as a multidimensional construct, as potentially conceptualized by other religiosity measures.

Limitations

There were several limitations to our study. First, our measurement invariance results should be taken with some reservations because some possible revisions to the SVR subscale may be needed. For example, we found items 1 and 12 functioned poorly in both groups with standardized factor pattern coefficients less than .30 (i.e., items 1 and 12 were weakly related to the construct). It is possible that removing these items from the analyses may change the fit of the models used in our study. Second, due to the relatively small number of LGBQ young adults in the study, we collapsed these four groups into a single non-heterosexual group and compared them to a single heterosexual group. Because the majority of the non-heterosexual group consisted

of self-identified gay and bisexual young adults, the correlation matrix obtained from the non-heterosexual group was likely more representative of gay and bisexual young adults' responses than those of queer and lesbian young adults, which may have influenced our measurement invariance results. Finally, our sample of young adults was recruited at concert venues and through social media outlets sponsored by the BTWF and may not be representative of the broader young adult population.

Conclusion and future research

In conclusion, we conducted a psychometric evaluation of the SVR subscale by first investigating the dimensionality of scale and then testing the measurement invariance of the scale across heterosexual (straight) and non-heterosexual (lesbian, gay, bisexual, and queer) young adults. Our results provided support for the theorized factor structure of SVR subscale (unidimensional) and use of the SVR subscale across heterosexual and non-heterosexual young adults (configural, metric, and scalar invariance). Future researchers studying the SVR subscale should aim to test measurement invariance between lesbian, gay, bisexual, and queer young adults individually compared to heterosexual young adults because it is possible for the SVR subscale to be invariant for only a subset of non-heterosexual young adults, but not for all.

Notes

1. The purpose of the Born Brave Experiences research study was to develop, assess, and evaluate a kindness and bravery scale for lesbian, gay, bisexual, transgender, and queer (LGBTQ) young adults. Data for the broader study were collected in two phases. The first phase of data collection was in 2013, and the second phase of data collection was in 2014.
2. We compared the demographics of our non-heterosexual young adults to the demographics of non-heterosexual young adults who participated in GLSEN's 2017 National School Climate Survey (N = 20,236; GLSEN, 2017). With respect to sexual orientation and race, the breakdowns of our non-heterosexual sample were similar to those reported by GLSEN, with most self-identifying as either gay or lesbian and White.

3. We tested the invariance of item 1 by first fitting a configural model with item 1 as the reference item. Then we estimated 11 additional models where we constrained item 1's factor pattern coefficients to be equal across both groups with each of the other 11 items as the reference item. We then compared the fit of the latter 11 estimated models to the fit of the configural model. We found the 11 estimated models did not fit significantly worse than the configural model, which provided evidence for the invariance of item 1 (Rensvold & Cheung, 2001).
4. In a follow-up analysis, we allowed the error variances of items 9 and 10 to correlate and reran the invariance analyses. We found no substantial differences in the results. Thus, for the sake of parsimony, we reported the results from the models without the correlated error parameter. These results are available from the first author by request.
5. To estimate the latent mean difference, we fixed the heterosexual group's factor mean to 0 and allowed the non-heterosexual group's factor mean to be freely estimated. The factor mean estimated in the non-heterosexual group represents the latent mean difference between the two groups (Thompson & Green, 2006).

Disclosure No potential conflict of interest was reported by the authors.

References

- Barnes, D. M., & Meyer, I. H. (2012). Religious affiliation, internalized homophobia, and mental health in lesbians, gay men, and bisexuals. *American Journal of Orthopsychiatry*, *82*, 505–515. doi:10.1111/j.1939-0025.2012.01185.x
- Bontempo, D. E., & Hofer, S. M. (2007). Assessing factorial invariance in cross-sectional and longitudinal studies. In A. D. Ong & M. H. M. van Dulmen (Eds.), *Oxford handbook of methods in positive psychology* (pp. 153–175). New York, NY: Oxford University Press.
- 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
- Clark, J. M., Brown, J. C., & Hochstein, L. M. (1990). Intuitional religion and gay/lesbian oppression. *Homosexuality and Family Relations*, *14*, 265–284.
- Eliason, M. J., Burke, A., van Olphen, J., & Howell, R. (2011). Interactions of sexual identity, sex/ gender, and religious/spiritual beliefs on college student substance use. *Sexuality Research and Social Policy*, *8*, 117–135. doi:10.1007/s13178-011-0046-1
- Ellison, C. (1995). Race, religious involvement and depressive symptomatology in a southeastern U.S. community. *Social Science and Medicine*, *40*, 1561–1572.
- Faria, L. (1996). Marsh's self description questionnaire III (SDQ III): Adaptation study with Portuguese college students. *Social Behavior & Personality: An International Journal*, *24*, 343. doi:10.2224/sbp.1996.24.4.343

- Finney, S. J., & DiStefano, C. (2013). Nonnormal and categorical data in structural equation modeling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: A second course* (2nd ed., pp. 439–492). Charlotte, NC: Information Age. Hill, P. C.
- GLSEN. (2017). *The 2017 national school climate survey: The experiences of lesbian, gay, bisexual, transgender, and queer youth in our nation's school*. Retrieved from <https://www.glsen.org/sites/default/files/GLSEN%202017%20National%20School%20Climate%20Survey%20%28NSCS%29%20-%20Full%20Report.pdf>
- Hoyle, R. H., & Panter, A. T. (1995). Writing about structural equation models. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 158–176). Thousand Oaks, CA: Sage.
- Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424–453. doi:10.1037/1082-989X.3.4.424
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55. doi:10.1080/10705519909540118
- Jöreskog, K. G., & Sörbom, D. (2015). *LISREL 9.20 for Windows* [Computer software]. Skokie, IL: Scientific Software International, Inc.
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York, NY: Guilford.
- Longo, J., Walls, N. E., & Wisneski, H. (2013). Religion and religiosity: Protective or harmful factors for sexual minority youth? *Mental Health, Religion & Culture*, 16, 273–290. doi:10.1080/13674676.2012.659240
- Marsh, H. W. (1992). *The Self Description Questionnaire (SDQ) III: A theoretical and empirical basis for the measurement of multiple domains of late adolescent self-concept: An interim test manual and a research monograph*. New South Wales, Australia: University of Western Sydney, Faculty of Education.
- Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers to overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling*, 11, 320–341. doi:10.1207/s15328007sem1103_2
- Marsh, H. W., & O'Neill, R. (1984). Self description questionnaire III (SDQ-III): The construct validity of multidimensional self-concept ratings by late-adolescents. *Journal of Educational Measurement*, 21, 153–174.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- Morrow, D. F. (2003). Cast into the wilderness: The impact of institutionalized religion on lesbians. *Journal of Lesbian Studies*, 7, 109–123. doi:10.1300/J155v07n04_07
- Newman, B. S., & Muzzonigro, P. G. (1993). The effects of traditional family values on the coming out process of gay male adolescents. *Adolescence*, 28, 213–226.

- Olsson, U. H., Foss, T., Troye, S. V., & Howell, R. D. (2000). The performance of ML, GLS, and WLS estimation in structural equation modeling under conditions of misspecification and nonnormality. *Structural Equation Modeling*, 7, 557-595. doi:10.1207/S15328007SEM0704_3
- Olsson, U. H., Troye, S. V., & Howell, R. D. (1999). Theoretic fit and empirical fit: The performance of maximum likelihood versus generalized least square estimation in structural equation modeling. *Multivariate Behavioral Research*, 34, 31-58. doi:10.1207/s15327906mbr3401_2
- Plante, T. G., & Sherman, A. C. (Eds.). (2001). *Faith and health: Psychological perspectives*. New York, NY: Guilford Press.
- Quintana, S. M., & Maxwell, S. E. (1999). Implications of recent developments in structural equation modeling for counseling psychology. *The Counseling Psychologist*, 27, 485-527. doi:10.1177/0011000099274002
- Ream, G. L., & Savin-Williams, R. C. (2005). Reconciling Christianity and positive nonheterosexual identity in adolescence, with implications for psychological well-being. *Journal of Gay and Lesbian Issues in Education*, 2, 136-159. doi:10.1300/J367v02n03_03
- Rensvold, R. B., & Cheung, G. W. (2001). Testing for metric in variance using structural equation models: Solving the standardization problem. In C. A. Schriesheim & L. L. Neider (Eds.), *Research in management: Equivalence in measurement* (pp. 25-50). Greenwich, CT: Information Age.
- Rew, L., & Wong, Y. J. (2006). A systematic review of associations among religiosity/ spirituality and adolescent health attitudes and behaviors. *Journal of Adolescent Health*, 38, 433-442. doi:10.1016/j.jadohealth.2005.02.004
- Rostosky, S. S., Danner, F., & Riggle, E. D. B. (2007). Is religiosity a protective factor against substance use in young adulthood? Only if you're straight! *The Journal of Adolescent Health*, 40, 440-447. doi:10.1016/j.jadohealth.2006.11.144
- Rostosky, S. S., Danner, F., & Riggle, E. D. (2008). Religiosity and alcohol use in sexual minority and heterosexual youth and young adults. *Journal of Youth and Adolescence*, 37, 552-563.
- Sherkat, D. E. (2002). Sexuality and religious commitment in the United States: An empirical examination. *Journal for the Scientific Study of Religion*, 41, 313-323.
- Steenkamp, J. E., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-90.
- Thompson, M. S., & Green, S. B. (2006). Evaluating between-group differences in latent variable means. In G. R. Hancock & R. O. Mueller (Eds.), *A second course in structural equation modeling* (pp. 119-169). Greenwich, CT: Information Age.
- Wills, T. A., Yaeger, A. M., & Sandy, J. M. (2003). Buffering effect of religiosity for adolescent substance use. *Psychology of Addictive Behaviors*, 17, 24-31.
- Wong, Y. J., Rew, L., & Slaikou, K. D. (2006). A systematic review of recent research on adolescent religiosity/spirituality and mental health. *Issues in Mental Health Nursing*, 27, 161-183.

Zinnbauer, B., Pargament, K. I., & Scott, A. B. (1999). The emerging meanings of religiousness and spirituality: Problems and prospects. *Journal of Personality*, 67, 889-919.