SSOAR

Open Access Repository
www.ssoar.info

Supernatural Belief Scale (SBS) - German Version
Bluemke, Matthias; Merkle, Belinda; Jong, Jonathan; Halberstadt, Jamin

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der

Leibniz-Gemeinschaft

Supernatural Belief Scale (SBS)
– German Version

Matthias Bluemke, Belinda Merkle,
Jonathan Jong, Jamin Halberstadt

1 Overview

Abstract

The 10-item scale measures the tendency to believe in supernatural entities such as immaterial agents, afterlife/otherworld places, and miraculous earthly events. The scale is intended to be a valid indicator of the cognitive component of religiosity, capturing ten cross-cultural beliefs that appear as religious themes in various religious contexts in the world. The scale was first tested in English by Jong, Bluemke, & Halberstadt (2013). The SBS aims to be applicable to people of different religious standing (atheists and believers alike), across various countries, with a multitude of religious backgrounds, if necessary by adapting the labels pertaining to the supernatural entities, while not altering the essence of the item content or response format. After its seminal application in New Zealand, and later in Croatia, among student samples with mostly a Christian cultural background, it was applied in Germany, yielding good reliability, a replicable factor structure, and cross-cultural measurement invariance (strict invariance across Germans and New Zealanders). The scale showed signs of construct validity and criterion validity in all samples (NZ, CRO, GER).

2 Instrument

Instruction

Please indicate your agreement with the following statements, using the scale below:

Items

Table 1. SBS items

No.	Item
1	There exists an all-powerful, all-knowing, loving God.
2	There exists an evil personal spiritual being, whom we might call the Devil.
3	There exist good personal spiritual beings, whom we might call angels.
4	There exist evil, personal spiritual beings, whom we might call demons.
5	Human beings have immaterial, immortal souls.
6	There is a spiritual realm besides the physical one.
7	Some people will go to Heaven when they die.
8	Some people will go to Hell when they die.
9	Miracles—divinely-caused events that have no natural explanation—can and do happen.
10	There are individuals who are messengers of God and/or can foresee the future.

Note. Alternative wordings for afterlife places (#7; #8) are preferred nowadays, in line with later research on cross-cultural applicability (see Scale Development).

Response specifications

Answers are provided on 9-point-rating-scales, ranging from −4 to +4 (−4 = strongly agree; 0 = neither agree nor disagree; +4 = strongly agree).

Scoring

The SBS score is computed as the mean of the raw item scores (ratings). Positive scores represent supernatural belief, negative scores represent disbelief, and middle scores represent agnosticism.

Application field

The SBS measures exclusively the cognitive aspect of religiosity – an individual's tendency to believe in supernatural entities, places, and events. The scale has been sucessfully applied as a paper-pencil questionnaire, computer-based test form, and web-based questionnaire. Metric invariance across cultures (with Christian heritage across New Zealand, Croatia, and Germany), and partial scalar invariance, allows researchers to compare across contexts SBS's correlations, regression coefficients, factor means, and—within limits and only with careful consideration—even effect sizes on the basis of observed means (see Quality Criteria). So far the SBS has been predominantly tested with students, mostly from a Christian cultural background, though it has been developed to be applicable across

various countries and religions. It is unique, firstly, by its choice of religious themes surfacing in major world religions around the globe, and secondly, by allowing for cultural sensitivity and careful adaptation that allows replacing the volatile item label in the item wordings (e.g., substituting "Allah" for "God" in predominantly Muslim cultural contexts) while keeping the central tenet of an item constant.

3 Theory

Religiosity, or religiousness, can be conceptualized in many ways, with cognitive (belief), affective (belonging, identity), and behavioral (individual activity, social groups) aspects being the prime candidates (Hill & Pargament, 2003). Arguably, religiosity is a peculiar psychological and sociological phenomenon with interindividual variability. Far from considering religiosity as a simple construct, this scale focuses on the cognitive domain of religiosity with its basically untestable, yet often costly beliefs about the world (Norenzayan et al., 2016). Different from what the secularization hypothesis states, religious belief shows no signs of decrease at a global level; in fact people with traditional religious views are a growing percentage of the world's population (Norris & Inglehart, 2011). Nonetheless, religious belief is often changing at a country level. For instance, in industrial countries religious belief is continuously waning (Smith, 2012). Not only is religious belief a globally important topic for many people, but the ongoing changes among societies and religions make the scientific study of the etiology, the correlates, and the consequences of religious belief a field of universal importance. Hence the need for a reliable scale that assesses religiosity uncontaminated by highly-specific religious content, free from religious activities and belongingness to social groups, as well as accompanying affective qualities.

The SBS assesses one's tendency to believe in supernatural entities such as immaterial agents, otherworld places (or afterlife conceptions), and miraculous earthly events. By focusing on supernatural belief, the SBS captures specifically the cognitive aspect of religiousness (Jong, Bluemke, & Halberstadt, 2013; Bluemke, Jong, Grevenstein, Mikloušić, & Halberstadt, 2016). Intended to be a unidimensional measure of supernatural belief, the SBS is supposed to correlate with other aspects of religiousness such as religious identity and behavior as well as a positive orientation towards spiritual values; yet it does not blend all facets into one (Gorsuch, 1984). Furthermore, the SBS was designed to be applicable among people from all cultures, independent of specific religious orientations, say, monotheistic religious worldviews (e.g., Judaism, Christianity, Islam), polytheistic religious backgrounds (e.g., Hinduism), agnostic worldviews (e.g., Buddhism) or atheistic philosophies (e.g., Atheism, Humanism).

To identify cross-culturally recurring religious themes, recent literature on psychological, anthropological, and religious studies—in particular monographs on Cognitive Science of Religion (e.g,. Whitehouse & Laidlaw, 2007)—were consulted and discussed with a New Zealand-based religious studies scholar. The following five themes were identified that are recurring in many religions around the world: high-order agents, low-order agents, afterlife entities/spiritual dimension, afterlife places, and supernatural events (on earth).

4 Scale development

Item generation and selection

The SBS items were initially developed in English (Jong; see item wordings in Table 1). The German translation of the questionnaire is available as a download, appended to this documentation. Items to represent the five themes were drafted, consisting of two components: (1) an affirmation of the belief in the existence (or occurrence) of the supernatural entity (or event), which can be either of positive, neutral, or negative valence; and (2) an additional label to help respondents from specific dominant religious contexts to construe the content (Jong et al., 2013). The items were grouped in pairs to represent specific aspects of supernatural belief high-order agents (God vs. Devil), low-order agents (angels vs. demons), afterlife entities/spiritual dimension (souls and realm), afterlife places (heaven vs. hell), and supernatural events on earth (miracles and prophecy). The ten SBS items are preferably presented in a fixed, logical order that maintains the facetted structure of five intended religious belief facets.

Cross-cultural validations (e.g., Bluemke et al., 2016, and this documention) have shown that the scale can be translated and incorporated into other cultures without loss of generalizability. For example, in predominantly Muslim contexts the label "God" might be exchanged by "Allah" without loss of information. While the word "soul" was used in the original English version, for use among urban New Zealanders the word "atta" or "atman" might be used instead in contexts where Pali- or Sanskrit-based religious traditions (e.g., Hinduism, Buddhism) are more culturally influential. Similarly, while "prophet" was originally used, "shaman" or "medium" might be preferable in other cultural contexts. Any changes to the item wordings are to be documented and made explicit, and measurement invariance has to be inspected first before inferring validity at the construct level after changes have been introduced.

In the more recent history of the SBS, the wordings for afterlife places (items #7 and #8) had to be changed to be more in line with research on cross-cultural supernatural beliefs and the applicability of the SBS. Instead of focusing on afterlife *places,* and to obtain better translatability and higher cultural meaningfulness, the items should rather convey general positive or negative afterlife *conceptions*. Thus, the SBS-item #7 might be replaced by "Some people will be rewarded in an afterlife when they die", and item #8 might be replaced by "Some people will be punished in an afterlife when they die" in cross-cultural comparisons. These variants are also more comparable to the item wordings introduced with the recent SBS-6 short scale (Bluemke et al., in preparation).

Samples

The scale was initially administered in multiple studies at the University of Otago in New Zealand, resulting in a total sample size of $N = 477$ New Zealanders for whom SBS scores were available (Jong et al, 2012; Jong et al., 2013). The second sample for the cross-cultural validation consists of $N = 642$ Croatian students, five of whom did not answer any of the SBS items and were excluded from the analyses below (Bluemke et al., 2016). The third sample encompasses three German subsamples from social psychological online experiments on death anxiety and supernatural belief, resulting in a

total German sample of N = 403 for whom SBS baseline measurements were available (Bluemke, Gonzalez, Laukenmann, Jong, & Halberstadt, 2015). In all the samples, the original item wordings (Table 1) were used, not the updated items #7 and #8. They were checked by backtranslation carried out by two speakers who were competent in their native language and in English (but see Harkness's, 2003, TRAPD approach that recommends translation, review, adjudication, pretest, documentation).

Table 2. Sample Demographics (NZ, CRO, GER)

	New Zealand	Croatia	Germany
N_{Total}	477 (100%)	637 (100%)	403 (100%)
Sex			
Men	176 (26.9%)	188 (29.5%)	136 (33.7%)
Women	300 (62.9%)	440 (69.1%)	265 (65.8%)
Religious Identification			
Religious	207 (43.4%)	428 (67.2%)	217 (53.8%)
Non-religious/Atheist/Agnostic	270 (56.6%)	187 (29.4%)	186 (46.2%)

Note. Differences to 100% are due to "other" codings, or due to missing data.

Item analyses

Item parameters of the manifest items were obtained from the three samples, presented in chronological order (Table 3). The correct interpretation of these item parameters needs to take into account that the items are not univariate normally distributed; rather bimodal distributions are typical in the field or religiosity (due to a mixture of "believers" and "non-believers" and few people covering the middle ground of agnosticism).

Table 3. Descriptives of manifest items (NZ, CRO, GER)

		New Zealand			Croatia			Germany		
		M	SD	$r_{\text{item-total}}$	M	SD	$r_{\text{item-total}}$	M	SD	$r_{\text{item-total}}$
1	God	−0.05	2.96	.85	0.50	2.97	.84	−0.59	2.96	.80
2	Devil	−1.04	2.89	.84	−0.51	2.89	.83	−2.24	2.47	.76
3	Angels	−0.07	2.75	.89	0.22	2.73	.85	−0.97	2.75	.84
4	Demons	−0.95	2.74	.82	−0.77	2.72	.74	−2.09	2.41	.80
5	Souls	0.28	2.57	.71	1.31	2.61	.72	−0.23	2.73	.74
6	Spiritual Realm	1.21	2.46	.70	1.75	2.45	.69	0.60	2.65	.63
7	Heaven	0.08	2.89	.85	−0.01	2.92	.86	−1.29	2.84	.80
8	Hell	−0.77	2.77	.82	−0.39	2.85	.81	−2.27	2.35	.74
9	Miracles	1.03	2.61	.71	0.74	2.83	.82	−0.44	2.91	.82
10	Prophecy	−0.99	2.60	.72	−1.29	2.54	.60	−1.79	2.58	.77

Note. Depending on missing values, *Ns* vary between 474-477, 635-637, and 403 for NZ, CRO, and GER, respectively. Item-total correlations are part-whole corrected.

Factorial validity

Dimensionality was investigated in the original New Zealand sample with exploratory factor analysis (EFA) of the ten SBS items in a pretest sample of New Zealand students. EFA indicated that none of the items should be eliminated (Jong et al., 2013). Given the generally high inter-itemcorrelations, initially a one-factorial solution with high factor loadings (.72–.91) and high communalities (.51–.84) was supported (Jong et al., 2013). Confirmatory factor analysis (CFA) on additional NZ samples provided support for the factorial validity of the scale with a dominant first factor (Jong et al., 2012; Jong et al., 2013). The SBS captures a general factor underlying all ten items. Non-negligible item variance can be attributed to an orthogonal method factor for negative entitites and another five orthogonal content facets (Figure 1).

The items rather follow bimodal distributions (believers and disbelievers), which necessitates robust maximum likelihood (ML) estimation (MLR), which was used for CFA and later invariance testing with multiple-groups (MGCFA). The SBS is intended to be a unidimensional instrument, yet different from what the seminal EFA suggested, CFAs converged on the fact that more than one dimension is necessary to replicate the data structure sufficiently well (Jong et al., 2013; Bluemke et al., 2016). The strongest factor is reflected in the latent variable for the general supernatural belief tendency. Five secondary factors explain item uniqueness due to the influence of the five themes or facets that underlie the item generation: high-order agends (HOA), low-order agents (LOA), afterlife entities (ALE), afterlife places (ALP), and events (E). An additional method factor reflected variability specific for negative supernatural item content (see Figure 1).
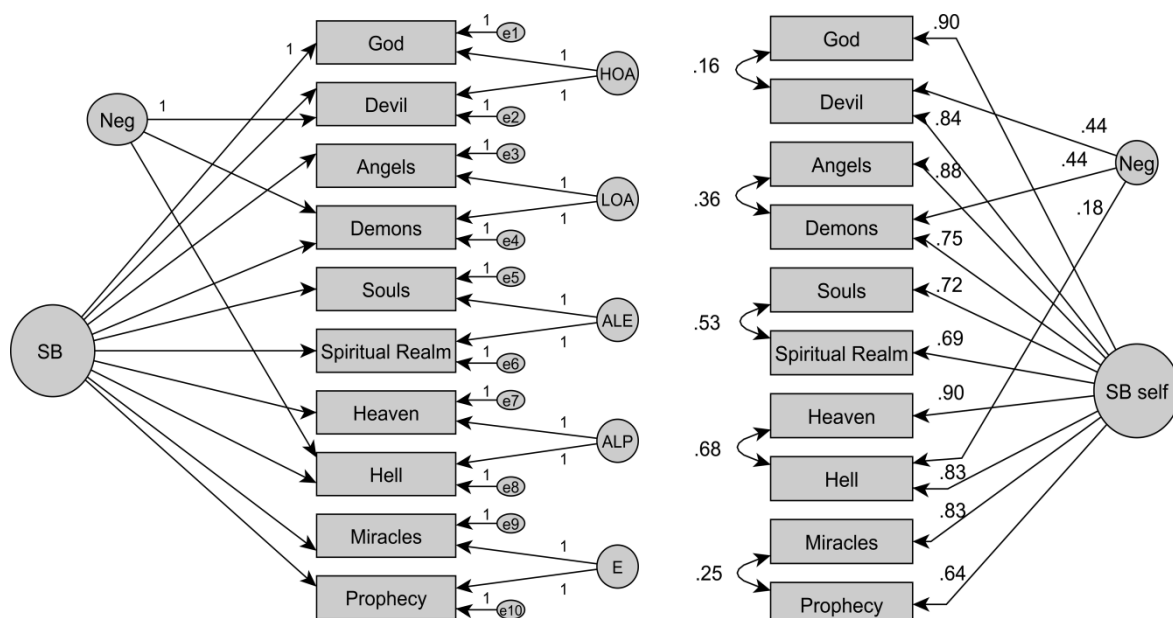

Figure 1. SBS factor model with essentially unidimensional structure. HOA = high-order agends, LOA = low-order agents, ALE = afterlife entities, ALP = afterlife places, and E = events. The left panel shows the intended structure (with five facets). The right panel shows the estimated standardized

paths on the basis of Croatian Supernatural Belief self- and peer-reports (five facets represented as correlated residuals).

Figure 1 shows the model specifications of the accepted measurement model. The estimated parameters are taken from the full measurement invariance model across Croatian self- and peer-reports, which had very good model fit, $\chi^2(180) = 535.57^{***}$, CFI = .96, RMSEA = .055. Left and right model present two specifications of the measurement model; they result in the same degrees of freedom and the same model fit. The difference is that facets are represented either as hypothesized latent variables or modelled as (a priori specified) correlated item residuals. The measurement model replicated well across Croatian self- and peer-reports, and across Croatian and New Zealand samples (Bluemke et al., 2016), confirming the factorial validity of the SBS *essentially unidimensional* structure. This model fitted the NZ data best, $\chi^2(27) = 57.13^{***}$, CFI = .97, RMSEA = .072, SRMR = .033; and it replicated the German data: $\chi^2(27) = 104.81^{***}$, CFI = .97, RMSEA = .085 [.06–.10], SRMR = .035.

5 Quality criteria

Objectivity
The SBS can by applied, evaluated, and interpreted objectively.

Reliability
About 80% item variance can—on average—be explained by the accepted SBS measurement model according to the Average Variance Extracted (AVE; Fornell & Larcker, 1981), and as shown in Table 4 more than 60% of item variance can—on average—be attributed to the dominant supernatural belief factor, indicating an essentially unidimensional scale. Questionnaire reliability is often evaluated by the internal consistency of the used manifest variables, Cronbach's alpha (Cronbach, 1951). Given that the SBS is only essentially unidimensional, the scale reliability according to Raykov's (1997) Rho, or McDonald's Omega (McDonald, 1999), should be computed, as the assumptions for Cronbach's Alpha are violated, which can lead to reliability misestimation (Raykov, 1998). As secondary factors influence the covariance structure too, Cronbach's Alpha is only a poor estimate of scale reliability. Below, we present the conventionally used reliability index Alpha, next to Rho and Omega-hierarchical (Brunner, Nagy, & Wilhelm, 2012). Rho is a more appropriate reliability indicator, reflecting the reliable amount of variance accounted for both by general construct and specific facets underlying the scale score. Given the facetted structure of the SBS, Omega-hierarchical is the most appropriate estimator of reliability of the target construct, supernatural belief. All estimates indicate excellent reliability across the three cultures.

Table 4. Reliability estimates according to different underlying model assumptions

	Cronbach's α	Rho ρ	Omega-h Ω_h	AVE (1st factor)
New Zealand	0.95	0.97	0.91	0.64
Croatia	0.95	0.97	0.92	0.64
Germany	0.94	0.97	0.91	0.62

Note. AVE = Average variance explained: percentage of item variability that is explained by the underlying factor(s). Reliability estimates are based on unstandardized factor loadings from Mplus.

Validity

Three types of validity were inspected: known-groups validity, construct validity (convergent and discriminant validity), and criterion validity. Regarding known-groups validity, the SBS behaved as expected, with SBS scores differing between one and two pooled standard deviations between self-declared believers and non-believers (d = 1.38–2.27, see Table 5).

Table 5. Differences between religious and non-religious people on SBS mean scores

	$M_{Religious}$	$M_{Non-religious}$	$t(df)$	p	Cohen's d
New Zealand	1.48	−1.37	−17.40 (475)	< .001	1.60
Croatia	1.13	−2.15	−23.63 (432.54)	< .001	2.27
Germany	0.00	−2.44	−13.73 (398.32)	< .001	1.38

Note. Corrected *t*-tests (adjusted degrees of freedom), whenever the Levene test indicated significant heterogeneity of variances (p < .01).

In terms of convergent validity, the SBS has been validated mostly with several one-item (9-point Likert-type) measures (see Table 6). Across all countries the SBS correlated substantially, albeit differently, with the self-ascribed importance of religion in one's life, that is, one's religious identity. In terms of discriminant validity, we report the correlation of Germans' SBS scores with intrinsic and extrinsic religiosity. The latter aspect resulted in somewhat weaker correlations—a mild indication for discriminant validity: Stronger belief sensu SBS is more like being inherently religious, and thus different from religiosity that individuals might maintain for mere social reasons. Likewise, the SBS's agreement with being "religious" was somewhat higher than that with being "spiritual". To establish further aspects of construct validity, the Croatian peer-ratings of participants' supernatural belief levels were correlated with participants' self-reports on the same measure, yielding a satisfactory degree of association, r(628) = .75 (Kendall's τ = .56).

Table 6. Convergent validity: Pearson correlations (Kendall's Tau)

	Religious Identity (Importance)	Intrinsic Religiousness	Extrinsic Religiousness	Self-Description: Religious	Self-Description: Spiritual
New Zealand	.41 (.32)	-	-	-	-
Croatia	.81 (.65)	-	-	-	-
Germany	.51 (.37)	.72 (.53)	.40 (.29)	.72 (.56)	.58 (.46)

*Note. N*s vary depending on availability per country and variable: N_{NZ} = 330; N_{CRO} = 610; N_{GER} = 339, 59/61, and 227/227, for Identity, Religiousness, and Religiosity/Spirituality variables; all *p* < .01.

The SBS's criterion validity was satisfactory, given that SBS correlated with the frequency of self-reported religious behaviors reported on 5-point rating scales with verbal frequency anchors. The criteria were visiting religious services, participating in holy communion, or praying (see Table 7), though the latter two criteria were only examined in Croatia.

Table 7. Criterion validity regarding religious behavior: Pearson correlations (Kendall's Tau)

	Service (Mass)	Communion	Prays
New Zealand	.63 (.46)	-	-
Croatia	.70 (.57)	.68 (.55)	.77 (.60)
Germany	.62 (.43)	-	-

*Note. N*s = 276, 630-637, and 339 for NZ, CRO, and GER, respectively; all *p* < .01.

Descriptive statistics (scaling)

The distributions of aggregated scale means differed significantly from the normal distribution: QQ-plots and histogramms showed deviations specifically towards the extreme ends. The degree of skewness was minor and stayed close to zero, whereas the (typically bimodal) platykurtic distributions were reflected in overall negative kurtosis values. Descriptes are given in Table 8.

Table 8. Means, standard deviations, and distributional parameters for SBS scale means

	M	*SD*	Skewness (*SE*)	Kurtosis (*SE*)
New Zealand	−0.13	2.26	0.09 (0.11)	−0.89 (0.22)
Croatia	0.15	2.27	−0.21 (0.10)	−0.99 (0.19)
Germany	−1.13	2.18	0.50 (0.12)	−0.61 (0.24)

*Note. N*s = 477, 637, and 403, for NZ, CRO, and GER, respectively.

Further quality criteria

Economy. The 10-item SBS can be used economically (roughly 1–2 minutes of survey time).

Measurement Invariance. The comparability of SBS scores across different groups was inspected by a series of multigroup confirmatory factor analyses (MGCFA). Analyses were run for the comparison of

cultures and genders (Chen, 2008). First of all, the SBS is meant to be a relatively culture-fair test of religiosity. Second, to be able to inspect whether the supernatural belief tendency truly differs across the genders, the measure has to function comparably across gender groups.

Cross-cultural comparability. The NZ–CRO comparison has already been reported in Bluemke et al. (2016). Here we briefly report the main findings, then analyze in more detail the comparability of the German to the seminal NZ sample. Invariance held for English and Croatian participants not only at the level of the item-factor relationships (configural invariance); also the factor loadings appeared to be equivalent (metric invariance). Furthermore, partial scalar invariance could be demonstrated (with 8 out 10 intercepts being equivalent, so that SEM latent means can be compared in an appropriately specified partial-invariance SEM). Partial scalar invariance means that the intercepts of the items "Souls" and "Realm" differed somewhat between New Zealanders and Croatians, yet they did so in a compensatory way without biasing latent mean estimates; for details, see Bluemke et al., 2016). In combination with observed residual invariance, manifest SBS scale means can be used as a proxy for religiosity and compared across cultures (other cultures require additional invariance testing).

Table 9 displays the new invariance tests between the German sample and the seminal NZ sample. Arguably, accepting metric invariance exclusively for the dominant SBS factor is a boundary decision based on Chen's (2007) criteria, though constraining *all* loadings to equality results in an acceptable "full metric" invariance model. Except for the assumption of equal factor means, additional constraints did not produce severe model misfit. The best tradeoff between model accuracy and parsimony sensu BIC was obtained when factor variances were assumed to be equal while factor means were allowed to differ (Model 6b) rather than constrained (Model 6a). The SBS is applicable across the cultures.

Table 9. Measurement invariance between Germans and New Zealanders

Invariance Model		df	χ^2	Δdf	$\Delta\chi^2$	CFI	RMSEA	SRMR	BIC	MI
1	Configural	54	183.98***	-	-	.973	.074	.030	34639	YES
2a	Metric	63	239.08***	9	65.15***	.963	.080	.054	34641	YES
2b	Full metric	65	237.83***	11	57.59***	.964	.078	.054	34628	YES
3	Scalar	68	248.74***	3	10.87**	.962	.078	.056	34618	YES
4	Residuals	78	276.05***	10	29.71***	.958	.076	.059	34608	YES
5	Means	85	366.88***	7	106.76***	.941	.087	.109	34680	NO
6a	Means & Variances	92	369.51***	7	8.41	.942	.083	.108	34649	(NO)
6b	Variances	85	280.08***	7	7.80	.959	.072	.059	34574	YES

Note. N = 874 (missing data handled by Mplus 7.31 FIML); Metric = 1st-factor loadings equal, Full metric 1st-7th-factor loadings equal; * $p < .05$, ** $p < .01$, *** $p < .001$ (with difference-tests based on Satorra-Bentler corrected, scaled χ^2).

Gender comparability. When testing so far unreported gender differences, both the Croatian and German dataset resulted in full measurement invariance as the best trade-off between model

accuracy and parsimony according to BIC (Table 10a and 10b). Constraining parameters mostly resulted in better RMSEA values, and CFI hardly deteriorated. In other words, the SBS captures the same construct, regardless whether women or men are tested. Yet despite excellent overall model fit in Croatia (and good fit in Germany), statistical differences were noticeable when testing for equal factor means. This is a plausible outcome, given the gender differences typically found in research on religion. In all countries women reported higher religiosity (SBS mean scores) than men (Table 11). Due to cross-gender measurement equivalence, it is legitimate to conclude that women's supernatural belief tendencies are somewhat stronger than men's, both on the latent and the manifest level (d = 0.20–0.42).

Table 10a. Measurement invariance between Croatian women and men

Invariance Model		df	χ^2	Δdf	$\Delta\chi^2$	CFI	RMSEA	SRMR	BIC	MI
1	Configural	54	136.67***	-	-	.977	.070	.024	25044	YES
2a	Metric	63	146.13***	9	4.91	.976	.065	.027	24991	YES
2b	Full metric	65	153.28***	11	15.28	.975	.066	.027	24990	YES
3	Scalar	68	158.86***	3	5.26	.974	.065	.028	24976	YES
4	Residuals	78	168.92***	10	15.57	.974	.061	.031	24947	YES
5	Means	85	185.50***	7	16.91*	.974	.061	.042	24920	YES
6	Variances	92	190.39***	7	5.74	.972	.058	.049	24884	YES

Note. N_{CRO} = 628 (missing data handled by Mplus 7.31 FIML); Metric = 1st-factor loadings equal, Full metric 1st-7th-factor loadings equal; * $p < .05$, ** $p < .01$, *** $p < .001$ (Satorra-Bentler corrected).

Table 10b. Measurement invariance between German women and men

Invariance Model		df	χ^2	Δdf	$\Delta\chi^2$	CFI	RMSEA	SRMR	BIC	MI
1	Configural	54	145.55***	-	-	.959	.092	.038	16036	YES
2a	Metric	63	156.45***	9	9.06	.959	.086	.046	15992	YES
2b	Full metric	65	155.87***	11	8.90	.960	.084	.046	15980	YES
3	Scalar	68	164.01***	3	8.50*	.957	.084	.047	15970	YES
4	Residuals	78	202.48***	10	33.72***	.945	.089	.052	15979	(YES)
5	Means	85	228.44***	7	27.40***	.936	.092	.069	15967	NO
6a	Means & Variances	92	228.49***	7	6.24	.939	.086	.070	15938	(YES)
6b	Variances	85	202.37***	7	6.11	.948	.083	.056	15950	YES

Note. N_{GER} = 401 (missing data handled by Mplus 7.31 FIML); Metric = 1st-factor loadings equal, Full metric 1st-7th-factor loadings equal; * $p < .05$, ** $p < .01$, *** $p < .001$ (Satorra-Bentler corrected).

Table 11. Gender differences on SBS mean scores

	M_{Women}	M_{Men}	$t(df)$	p	Cohen's d
New Zealand	0.22	-0.74	4.55 (474)	< .001	0.42
Croatia	0.26	-0.10	1.79 (432.54)	.075	0.20
Germany	−0.98	-1.44	1.99 (399)	.047	0.20

Note. Corrected *t*-tests (adjusted degrees of freedom), whenever the Levene test indicated significant heterogeneity of variances ($p < .01$).

From a purely statistical point of view, the assumption of equal residuals between German men and women was untenable. Different amount of errors entered into their ratings, affecting scale reliability differently. Therefore, when interested in a direct comparison of religiosity among the genders in Germany, using observed SBS means is not recommended (rather use latent SEM variables).

6 Literature and data sources

Data sources

Croatian data are available through the ZPID/psychdata repository under the URL: http://psychdata.de. The DOI is: 10.5160/psychdata.bems99me29.

Contact details

Dr. Matthias Bluemke, GESIS – Leibniz Institute for the Social Sciences, 68159 Mannheim, Germany. Tel.: +49 621 1246-253; E-Mail: matthias.bluemke@gesis.org

References

Bluemke, M., Gonzalez, D., Laukenmann, R., Jong, J., & Halberstadt, J. (2015). *Assessing supernatural belief and existential death anxiety in Germany.* Unpublished raw data.

Bluemke, M., Jong, J., Grevenstein, D., Mikloušić, I., Halberstadt, J.(2016). Measuring cross-cultural supernatural beliefs with self- and peer-reports. *PLOS ONE, 11*(10): e0164291. doi:10.1371/journal.pone.0164291

Bluemke, M., Lechner, C., Danner, D., Groskurth, K., Jong, J., & Halberstadt, J. (in prep.). *German Version of the Short Supernatural Belief Scale (SBS-6).* In: Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS). Mannheim: GESIS.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*, 796–846.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14,* 464–504.

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*, 1005–1018.

Gorsuch, R. L. (1984). Measurement: The boon and bane of investigating religion. *American Psychologist, 39*(3), 228–236. doi:0.1037/0003-066X.39.3.228

Harkness, J. (2003). Questionnaire translation. In J. Harkness, F. van de Vijver, & P. Mohler (Eds), *CrossCultural Survey Methods* (pp. 35–56). Hoboken, NJ: Wiley.

Hill, P.C., & Pargament, K.I. (2003). Advances in the conceptualization and measurement of religion and spirituality: Implications for physical and mental health research. *American Psychologist, 58*(1), 64–74. doi:10.1037/0003-066X.58.1.64

Jong, J., Halberstadt, J., & Bluemke, M. (2012). Foxhole atheism, revisited: The effects of mortality salience on explicit and implicit religious belief. *Journal of Experimental Social Psychology, 48,* 983–989. doi:10.1016/j.jesp.2012.03.005

Jong, J., Bluemke, M., & Halberstadt, J. (2013). Fear of death and supernatural beliefs: Developing a new Supernatural Belief Scale to test the relationship. *European Journal of Personality, 27,* 495–506. doi:10.1002/per.1898

McDonald, R. P. (1999). *Test theory: A unified treatment.* Hillsdale, NJ: Erlbaum.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences, 39*, e1. doi:10.1017/s0140525x14001356

Norris, P., & Inglehart, R. (2011). *Sacred and secular: Religion and politics worldwide*. Cambridge University Press.

Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement, 21*(2), 173–184.

Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement, 22*(4), 375–385.

Smith, T. W. (2012). *Beliefs about God across time and countries.* NORC, University of Chicago.

Whitehouse, H., & Laidlaw, J. (Eds.) (2007). *Religion, anthropology, and cognitive science.* (Ritual studies monograph series). Durham, N.C.: Carolina Academic Press.