

In Press at the *Psychology of Religion and Spirituality*

©American Psychological Association, [2021]. This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. Please do not copy or cite without author's permission. The final article is available, upon publication, at: [DOI forthcoming]

God is Up?

Replication and Extension Attempts of Meier et al. (2007)

Brian P. Meier¹, Adam K. Fetterman², David J. Hauser³, & Michael D. Robinson⁴

¹Gettysburg College

²University of Houston

³Queen's University

⁴North Dakota State University

BP Meier and AK Fetterman acknowledge support for this project from the John Templeton Foundation (grant number 61592). Data collection plans for the studies were uploaded to the Open Science Framework website before data collection commenced: <https://osf.io/tgb9j/>. The cleaned data files in SPSS format can also be found at this website. Address correspondence to Brian P. Meier, Gettysburg College, bmeier@gettysburg.edu.

Abstract

Meier et al. (2007) examined relations between the concept of God and the vertical dimension of space. They found that God-related concepts were recognized faster if they were associated with words related to a high location (Study 1) or actually presented in a high location (Study 2). Additionally, participants recalled God-like images as appearing higher than they actually did (Study 3) and judged strangers as having a stronger belief in God if the strangers' images appeared toward the top versus bottom of a computer screen (Studies 4 and 5). Meier et al. (2007) suggested that people use the vertical dimension of space to represent the concept of God. The present investigation attempted to replicate effects from four of these studies using adequately powered designs while also examining potential individual difference moderators. We successfully replicated the effects of Studies 1 and 3, but not Studies 4 or 5. Individual differences related to God beliefs and religiosity did not moderate the findings. The results and a prior successful replication of Meier et al.'s (2007) Study 2 suggest that the concept of God is represented with the vertical dimension of space, but this representation does not appear to impact perceptions of other's traits.

KEYWORDS: Metaphor, God, Vertical Space, Embodiment, Person Perception, Replication

God is Up?

Replication and Extension Attempts of Meier et al. (2007)

God is often described in metaphoric terms, such as when Christians refer to God as “the man upstairs” or a “father” (DesCamp & Sweetser, 2005). Such metaphor-related ways of speaking about God may help speakers convey some of God’s central attributes, such as wisdom, power, or goodness.

In addition, metaphors may provide key insights into how people think about God, potentially in ways that get to the core of our representational systems. Lakoff and Johnson (1980, 1999) developed the conceptual metaphor theory (CMT) which proposed that metaphors map abstract concepts such as “love” to more concrete domains like “journeys”. These metaphoric mappings are cognitive associations, often learned through experience, that allow people to think about and represent abstract concepts (e.g., love) in ways that are more concrete and readily perceptible in a direct sense (e.g., journeys). Metaphoric mappings commonly involve associations between abstract concepts and body-based experiences, likely because we can easily understand concepts we can actually experience with our senses. Ideas along these lines are also considered in terms of embodiment or embodied cognition (Fincher-Kiefer, 2019; Meier et al., 2012).

A number of authors have suggested that CMT and embodiment views are well-suited for examining religion because religious beliefs tend to be abstract in nature and should therefore benefit from conceptions related to metaphor or bodily experience (Barsalou et al., 2005; Soliman et al., 2015; Van Capellan et al., 2021). Some supportive research on this connection has examined religion, CMT, and embodiment in tasks that involve judgment or behavior. For example, in a two-week experience sampling study, Fetterman (2016) found that people who had

a stronger belief in God reported feeling physically cleaner on days on which they reported higher self-control, consistent with metaphors linking cleanliness to Godliness. Concerning explicit thoughts and judgments, Ransom and Alicke (2013) found that participants who adopted a religious posture (e.g., kneeling) reported a higher frequency of religion-related cognitions. Adopting religious postures, that is, seemed to affect one's thinking in the religious domain.

Inspired by CMT and embodiment theory, Meier et al. (2007) examined potential associations linking representations of God to higher locations in vertical space (which presumably give rise to expressions or thoughts consistent with God watching down on us). In specific terms, Meier et al. (2007) hypothesized that people somewhat automatically link God to verticality, which should result in patterns of encoding, retrieval, and judgment that systematically (and unwittingly) favor God's higher spatial position. In these studies, participants categorized God-related stimuli more quickly when they were linked to words describing higher locations in space (Study 1) or when they were actually shown in higher spatial positions (Study 2). Furthermore, participants recalled that God-like images had been presented vertically higher than was actually the case (Study 3) and manipulations of vertical position influenced inferences concerning pictured strangers' belief in God (Study 4 and Study 5).

Meier et al. (2007) revealed that people seem to think about God in a way that reflects associations found in common metaphors. Studies have shown similar findings in other metaphoric domains (Landau, 2017; Landau et al., 2010). Although the Meier et al. (2007) paper has been cited 360 times (according to Google Scholar as of March, 2021), the sample sizes in the experiments were small ($Ns = 41, 47, 33, 27, 66, \& 55$) and there have been few conceptual replications reported in this area (although see Chasteen et al., 2010 & Guan et al., in press, for two examples). Furthermore, valid concerns have been raised concerning the replicability of

CMT and embodied cognition research (Skulmowski & Rey, 2018). For example, there have been reported failures to replicate metaphor and embodied cognition effects on social judgments and in studies involving judgments of other's traits (e.g., Ashton et al., 2014; Chabris et al., 2019; Siev et al., 2019; but see Hauser & Schwarz, in press, for a discussion of relevant boundary conditions; also see Giessner & Schubert, 2019, for a successful replication). For instance, Siev et al. (2019) reported a failure to replicate the "Macbeth Effect" which occurs when unethical thoughts trigger desires or behaviors related to cleansing oneself. Similarly, Ashton et al. (2014) reported a failure to replicate findings linking the personality trait of agreeableness to a preference for sweet foods. These findings relate to replication issues in psychological research more generally, given that (only) 50% of research results reported in the field might be replicable (Klein et al., 2018).

Nonetheless, successful replications have occurred using implicit cognitive tasks in which metaphor-related cues facilitate metaphor-consistent response latencies (e.g., Meier et al., 2015). However, it is difficult to know whether such apparent discrepancies between social judgment and response latency studies are real because there have been no attempts to replicate any particular relationship across both social judgment and cognitive tasks within the same metaphoric domain. Efforts of this type were central to the present research.

We conducted a series of studies that attempted to replicate findings reported by Meier et al. (2007) with the exception of their Study 2, which was replicated by Meier and Fetterman (in press) in a different type of project that examined multiple metaphors used in describing God (i.e., spatial location, brightness, & huamanness). In the 2007 paper, Studies 1 and 2 involved categorization latencies and Study 3 used a recall task – i.e., dependent measures that are cognitive in nature. Studies 4 and 5, on the other hand, involved a person perception task and

found that pictured strangers were inferred to believe in God to a greater extent. The latter is a “social priming” task because it involves judgments or inferences about other people in the context of a metaphor-related manipulation. Accordingly, the range of studies reported by Meier et al. (2007) seemed ideal in making conclusions about the scope and power of conceptual metaphors in representations of God and religiosity.

We also attempted to extend Meier et al.’s (2007) findings by examining potential moderators related to belief in God, religiosity, and endorsement of the God is up metaphor. Past work has shown that individual differences in metaphor use (Fetterman et al., 2016) or concerning the metaphoric relation of interest (Fetterman et al., 2019) moderate effects of the type discussed here. In the present case, for example, more religious individuals were presumed to possess stronger associations involving religious concepts, which should facilitate metaphor-linked processes related to those concepts (Lee, 2016). We should state, however, that our central focus was on replication rather than any potential moderation across individuals.

Disclosure Statement

We report all studies, data exclusions, manipulations, and measures, and how we determined sample sizes. Data-collection plans (uploaded before data collection commenced) and cleaned data files in SPSS format can be found on the Open Science Framework website:

<https://osf.io/tgb9j/>.

Study 1: Does Verticality Affect Perceptions of a Stranger’s Belief in God (on a six-point rating scale)?

Our original data-collection plan was to attempt to replicate Study 4 from Meier et al. (2007) and to determine if differences existed among three data-collection contexts that are often used in Psychology: an online context using Amazon Mechanical Turk (Mturk), traditional

samples collected in the context of Psychology research labs, and a sample collected in the context of a Psychology class project.

In Meier et al.'s (2007) Study 4, 27 participants rated strangers' belief in God as the strangers' images appeared near the top or bottom of a computer screen. Meier et al. (2007) hypothesized that strangers shown in higher versus lower vertical positions would be rated as having a stronger belief in God, which was confirmed in the original study ($d = .58$). However, the sample size was unacceptably small by modern standards. We used G*Power to determine the sample size needed to detect a small effect ($d = .20$) with 80% power in a paired-samples t -test. Although we test differences between contexts and examined individual differences, we were most interested in the replication of the original finding, which used a paired-samples t test. The sample size estimate was 199 and we attempted to collect data from approximately 200 participants in each of the contexts listed above.

Method

Participants

The initial sample consisted of 621 participants. Fifty-four participants were removed prior to analyses for the following reasons: the computer program created duplicate participant numbers for 20 participants, which precluded linking questionnaire data to rating data, 12 participants did not complete the experimental task, and 22 participants gave the same ratings for all trials. The final sample consisted of 567 participants:

- 197 participants with an average age of 36.49 ($SD = 10.68$) years from Mturk who were paid \$1.75 (92 females, 102 males, & 3 missing; 80.71% White, 6.09% Black, 5.58% Asian/Pacific Islander, 5.58% Hispanic, 1.02% Bi-Racial, .51% American Indian/Alaskan Native, & .51% missing)

- 181 participants with an average age of 19.33 ($SD = 1.41$) years who were recruited as volunteers by 21 students as part of a research project in a Statistics and Research Methods class at Gettysburg College (104 females, 74 males, 2 missing, & 1 other; 80.11% White, 6.63% Hispanic, 6.08% Asian/Pacific Islander, 3.31% Black, 1.66% Bi-Racial, 1.10% American Indian/Alaskan Native, & 1.10% missing)
- 189 participants with an average age of 19.36 ($SD = 3.07$) years who completed the study in the research labs of the second, third, and fourth authors in exchange for course credit (114 females, 71 males, 2 other, & 2 missing; 54.50% White, 29.63% Hispanic, 8.99% Asian/Pacific Islander, 3.17% Black, 1.59% missing, 1.06% Bi-Racial, & 1.06% American Indian/Alaskan Native)

Materials and Procedure

We used Inquisit software (www.millisecond.com) to present the study over the internet, following the procedures of Meier et al.'s (2007) Study 4. Stimuli consisted of one hundred (50 male & 50 female) grayscale images from the AR face database (Martinez & Benavente, 1998). These images, photographed from the shoulder up, depicted individuals who appear to range in age from 20 to 50 years. As in Meier et al. (2007), instructions indicated that the researchers were interested in how people make inferences about the extent to which strangers believe in God. On each trial, participants saw a new, unfamiliar target and their task was to make a rating of that person's belief in God using a rating scale that was shown at the vertical center of the screen. The 100 images randomly appeared near the top or bottom of the computer screen, subject to the constraint that there needed to be 50 at each position. The rating scale had a horizontal arrangement with 1 (no belief in God) toward the left and 6 (strong belief in God)

toward the right. After each rating, the image disappeared, and there was a 1,500-ms delay before the next image appeared.

Participants responded to demographic questions (age, sex, ethnicity, & religious affiliation) after the task as well as the following single-item self-reported measures:

- Belief in God ($M = 5.07$; $SD = 1.89$; Where would you place yourself on the following scale?; 1 = Strong Atheist - I am 100% certain there is no God to 7 = Strong Theist - I am 100% certain that there is a God)
- Religiosity ($M = 3.48$; $SD = 1.97$; To what extent are you a religious person?; 1 = not at all religious to 7 = very religious)
- God is Up agreement ($M = 3.98$; $SD = 2.23$; To what extent do you agree with this statement: “God is up high”?; 1 = not at all to 7 = strongly agree)

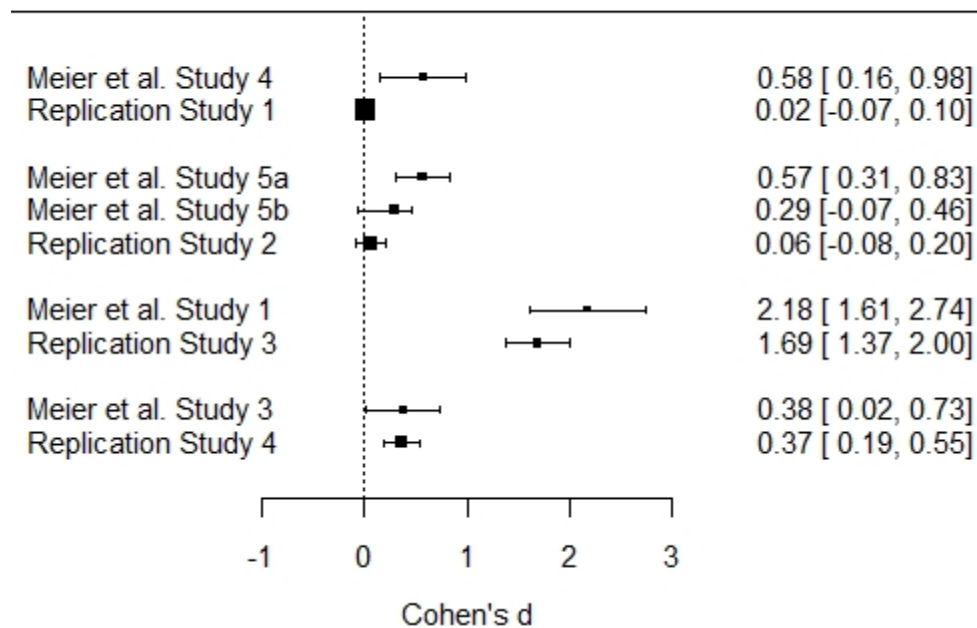
Results

We first examined the effect of position on belief in God ratings using a paired-samples t test. Contrary to Meier et al. (2007), the effect of image position was not significant, $t(566) = .36$, $p = .72$, $d = .02$ (95% CI: $-.07, .10$). The present means, as well as the means from Meier et al.’s (2007) Study 4, are shown in Table 1. A Forest Plot with the effects and confidence intervals of the original studies and the comparable replication studies from the current project are shown in Figure 1.

Table 1. *Original (Meier et al., 2007, Study 4) and Replication (Study 1) Means and Standard Deviations for the Belief in God Ratings by Image Position*

Image Position	Belief in God ratings	
	Meier et al., 2007, Study 4	Replication, Study 1
Top	3.61 (.34)	3.53 (.51)
Bottom	3.48 (.27)	3.54 (.52)

Figure 1. *Forest Plot of the Effect Sizes and Confidence Intervals of the Original Studies and Associated Replication Studies from the Current Project*



We next used a mixed-model ANOVA with image position (top, bottom) as a within-participant variable and study context (online, lab, class project) as a between-participants variable in order to examine the effect of context. The main effect of image position was not significant, $F(1, 564) = .10, p = .76, \eta_p^2 < .01$ (95% CI: .00, .00), which accords with the results of the t -test discussed above. There was a main effect for study context, $F(2, 564) = 23.71, p <$

.001, $\eta_p^2 = .08$ (95% CI: .04, .12), such that higher belief ratings occurred online ($M = 3.70$, $SD = .48$) relative to in the lab ($M = 3.54$; $SD = .47$) or with respect to a class project ($M = 3.36$, $SD = .47$). There was no image position by study context interaction, $F(2, 564) = 1.67$, $p = .19$, $\eta_p^2 = .01$ (95% CI: .00, .02).

Finally, we examined potential moderation effects by the individual difference variables. In order to simplify these analyses and their interpretation, we created a difference score for the image rating task by subtracting the mean rating of images presented at bottom from the mean rating of images presented at top ($M = -.01$; $SD = .34$). Positive (negative) scores mean that participants gave higher belief in God ratings for higher (lower) targets. We computed correlations between this difference score with the three potential moderators (note that not every participant completed all three measures). The belief in God, $r(561) = .01$ (95% CI: $-.07$, $.09$), $p = .86$, and religiosity, $r(561) = .05$ (95% CI: $-.03$, $.13$), $p = .26$, measures were not correlated with verticality effects. The correlation with the item assessing agreement that God is up was significant, $r(560) = .09$ (95% CI: $.01$, $.17$), $p = .04$. This correlation reveals that participants who generally believed God is up were more likely to ascribe stronger God beliefs to targets presented upwards (versus downwards). This correlation was small, however.

Study 2: Does Verticality Affect Perceptions of a Stranger's Belief in God (on a continuous rating scale)?

In the remaining studies, we abandoned the different data-collection contexts and focused on replication attempts of the remaining studies of Meier et al. (2007). We first attempted to replicate their Study 5, which involved a similar social perception task, albeit with some differences. One, the original paradigm randomized individual trials to one of two ratings - inferred beliefs in God and likability (their Study 5a) or power (their Study 5b), which Meier et

al. (2007) added to address alternative mechanisms. Given the results of Study 1 reported above, we only attempted to replicate the effects involving God-belief ratings (i.e., we ran one study with only God-belief ratings as the dependent measure using the paradigm described below). Two, belief-ratings were made for fewer trials. And three, ratings were made using a continuous rating bar, which might afford greater sensitivity.

We created a second data-collection plan and uploaded it to the Open Science Framework website before data collection commenced for this study and the remaining studies (<https://osf.io/tgb9j/>). In the current study, this plan revealed that we needed 156 participants to find a small effect with 80% statistical power using a paired-samples *t*-test. However, we made a mistake using G*Power and the correct figure was 199.

Method

Participants

The sample was collected at Queen's University and consisted of 189 participants. Due to a programming error, potential moderators for this study were not collected and demographics that were collected consisted of age and gender, but not race.¹ Gender and age was missing from 15 participants and, of the remaining 174 participants, 137 were female and 37 were male, with a mean age of 19.47 (*SD* = 2.66) years.

Materials and Procedure

The procedures were identical to Studies 5a and 5b from Meier et al. (2007) and involved 48 images (24 of each sex) from the AR face database (Martinez & Benavente, 1998). Half of the targets of each sex appeared near the bottom of the screen and half of the targets of each sex appeared near the top. Two counterbalanced programs ensured that all targets were presented in high vertical positions for some participants and low vertical positions for others. Target images

were presented on a white background with a rating bar arranged horizontally in the vertical center of the screen. This rating bar had “No Belief in God” toward the left and “Strong Belief in God” toward the right. Participants used the mouse and mouse cursor to select a point along the rating bar that reflected their impressions of God belief for each pictured stranger, resulting in a -1 to +1 scale with .001 increments.

Results

We used a paired-samples *t*-test to determine whether vertical placement of the images affected God-belief ratings. The effect of position was not significant, $t(188) = .88, p = .38, d = .06$ (95% CI: $-.08, .20$). The means from this study and the corresponding means from Meier et al. (2007) are shown in Table 2.

Table 2. *Original (Meier et al., 2007, Study 5a & Study 5b) and Replication (Study 2) Means and Standard Deviations of Belief in God Ratings by Image Position*

Image Position	Belief in God ratings		
	Meier et al., 2007, Study 5a	Meier et al., 2007, Study 5b	Replication, Study 2
Top	439.96 (71.39)	438.55 (68.09)	.04 (.12)
Bottom	412.47 (71.95)	425.73 (73.32)	.03 (.13)

Note: The replication Study 2 used a -1 to +1 rating scale with .001 increments for mouse-click locations while Studies 5a and 5b from Meier et al. (2007) used a 750-point rating scale with an increment of 1 for mouse-click locations.

Study 3: Are God and Vertical Space Implicitly Related?

We failed to replicate the studies from Meier et al. (2007) that focused on the perception of others' belief in God. Nonetheless, we suspected that the studies using more straightforward

cognitive procedures would replicate, so we sought to make this comparison. The current Study 3 is an attempt to replicate Meier et al. (2007) Study 1. That study used an implicit association test (IAT: Greenwald et al., 1998) to investigate whether participants associate God versus Devil concepts with verbal descriptors indicating a higher versus lower position in space. Meier et al. (2007) found a strong, hypothesis-consistent effect ($\eta_p^2 = .85$), but we sought to be conservative in our sample size estimate. A G*Power analysis indicated that 34 participants would give us 80% power to detect a medium-sized ($\eta_p^2 = .06$) IAT effect. Regardless, we sought to exceed this number as much as possible given the data-collection session timeframe and resources available. We managed to recruit 94 participants.

Method

Participants

The sample was collected at the University of Texas-El Paso and consisted of 94 participants. We removed one participant who had IAT accuracy rates less than 50%, which left a sample size of 93 participants (48 females, 44 males, 1 missing) with an average age of 21.24 ($SD = 3.84$) years. Self-reported race was as follows: 46.24% Hispanic, 22.58% Bi-Racial, 21.50% White, 7.53% Black, & 1.08% Asian/Pacific Islander.

Materials and Procedure

Meier et al.'s (2007) Study 1 procedures were used, which required us to use the original methodological procedures for the IAT (Greenwald et al., 1998). Stimuli consisted of four words from each category pertinent to the present associations: God (Almighty, Creator, Deity, & Lord), Devil (Antichrist, Demon, Lucifer, & Satan), up (High, Top, Above, & Ascend), and down (Low, Bottom, Below, Descend). Participants were told to categorize the words as quickly and as accurately as possible. Category labels appeared toward the upper left and right of the

computer screen. The choice to be made was always between “God” and “Devil” or between “Up” and “Down.” Participants were told to press the “Q” key if the word belonged to the relevant category on the left and the “P” key if it belonged to the relevant category on the right. Incorrect categorizations were followed by the word “INCORRECT” in red font for 1.5 seconds. Correct categorizations were followed by a 150-ms blank screen. Each stimulus remained on the screen until a response was made.

The IAT had seven blocks, but only two blocks comprised the critical comparisons: block four (“God or Up” versus “Devil or Down”: 48 trials) and block seven (“Devil or Up” versus “God or Down”: 48 trials). We randomly assigned participants to one of two versions of the IAT, which counterbalanced which of the two critical blocks occurred first.

Demographics (age, sex, ethnicity, & religious affiliation) were collected after the task as were other individual difference variables. These items varied slightly from Study 1. Also as in Study 2, a programming error resulted in the possibility of participants skipping these post-experiment questions, which a number of participants did. The pertinent moderator questions consisted of the following:

- Belief in God ($N = 43$; $M = 5.07$; $SD = 2.06$; I believe God Exists; 1 = strongly disagree to 9 = strongly agree)
- Religious Beliefs ($N = 62$; $M = 4.63$; $SD = 2.28$; I am a religious person; 1 = strongly disagree to 9 = strongly agree)
- God is Up belief ($N = 89$; $M = 5.64$; $SD = 1.80$; To what extent do you view or think about God as being Up or High; 1 = not at all to 7 = to a large extent)

Results

We followed procedures outlined by Greenwald et al. (1998), which involved deleting the first two trials of each block, replacement of trials with response times that were below 300 ms and above 3,000 ms with these values, deletion of inaccurate trials (3.2%), and a log-transformation of the raw latencies. For each participant, we then computed two raw means and two log-transformed means, one set pertaining to the compatible block (God + Up versus Devil + Down) and the other set pertaining to the incompatible block (God + Down versus Devil + Up). Of note, analyses were performed on transformed latency means, but patterns are reported in terms of millisecond units.

Data were analyzed with a mixed-model ANOVA, with block order and block type as independent variables. Confirming predictions, the main effect of IAT block type was significant, $F(1, 91) = 326.98, p < .001, \eta_p^2 = .78$ (95% CI: .70, .83). As shown in Table 3, this effect replicates Meier et al. (2007) and indicates that participants implicitly associate God with higher vertical positions, relative to the Devil.

Table 3. *Original (Meier et al., 2007, Study 1) and Replication (Study 3) Means and Standard Deviations for the Reactions Times (ms) for each IAT Block Type*

IAT Block Type	Response latency (in ms)	
	Meier et al. (2007), Study 1	Replication, Study 3
God-Up and Devil-Down	708.71 (83.66)	803.55 (141.93)
God-Down and Devil-Up	976.35 (184.61)	1,071.65 (227.82)

The remaining effects within the initial analysis were not of major interest, but are reported. The interaction between block order and block type was significant, $F(1, 91) = 4.94, p$

= .03, $\eta_p^2 = .05$ (95% CI: .00, .16). This interaction revealed that the IAT effect was stronger when the compatible block occurred first, $t(47) = 12.43, p < .001, d = 1.79$ (95% CI: 1.33, 2.25), relative to second, $t(45) = 10.91, p < .001, d = 1.61$ (95% CI: 1.17, 2.04). The main effect of block order was not significant, $F(1, 91) = 3.18, p = .08$.

To examine effects involving the individual difference variables, we created an IAT score by subtracting the compatible log-transformed mean from the incompatible log-transformed mean, with higher numbers indicating associations consistent with vertical metaphors (i.e., God is up). None of the potential moderators significantly correlated with the magnitude of the IAT effect: belief in God, $r(41) = -.03$ (95% CI: -.33, .27), $p = .87$, religiosity, $r(60) = -.10$ (95% CI: -.34, .15), $p = .46$, or agreement with the God is up item, $r(87) = .00$ (95% CI: -.21, .21), $p = .99$.

Study 4: Are God-Like Images Falsely Recalled as Appearing Higher in Vertical Space than Devil-Like Images?

In our final study, we attempted to replicate Meier et al.'s (2007) Study 3 in which 33 participants were presented with God-like, Devil-like, and neutral images one at a time, with the images placed in different vertical locations. Meier et al. (2007) found that God-like images were falsely recalled as appearing higher than Devil-like images.

The Meier et al. (2007) effect size was small to medium ($d = .38$), but we sought to be conservative with respect to a small ($d = .20$) effect size in the present context. The data collection plan specified that we needed 156 participants to have 80% power to detect effects in this range. However, as before, we made a mistake with G*Power and the correct number was 199. Given the data-collection session timeframe and available resources, we were only able to collect data from 125 participants. Even so, sensitivity analyses indicated that 125 participants gave us 80% power to detect an effect size of $d = .25$.

Method

Participants

The sample was collected at North Dakota State University and consisted of 125 participants (70 females, 53 males, 2 missing) with an average age of 19.25 ($SD = 1.90$) years. Self-reported race was as follows: 87.20% White, 5.60% Asian/Pacific Islander, 3.20% Black, 1.60% Hispanic, 1.60% Bi-Racial, and .80% missing.

Materials and Procedure

We used procedures identical to Meier et al.'s (2007) Study 3. Instructions indicated that the researchers were interested in memory and that the participants would see images to be encoded for a later memory test. They also indicated that the images would appear at different locations on the computer screen in order to increase participants' attention to the task. Stimuli to be encoded consisted of 5 God-like images, 5 Devil-like images, and 20 neutral images that were identical to those presented in Meier et al. (2007). Each image appeared for 1 second, which was followed by a 2-second blank screen. The images appeared in one of five vertical locations from the top to the bottom of the screen and an equal number of images from each category were presented in each location.

After all images were shown, participants were told that they would now see each image again, one at a time, in the center of the screen. On the right side of the screen was a white bar that spanned from top to bottom. Participants were instructed to use the mouse cursor to select a point on the vertical bar that reflected the vertical position at which the center of the image had been presented earlier in the task, which resulted in a 1 (top) to 768 (bottom) vertical scale. Participants were asked to recall the vertical location of all 30 images, which received a new

random ordering during the test phase. After each location choice was made, there was a 100 ms blank-screen delay until the next image was shown.

Participants responded to demographic (age, sex, ethnicity, & religious affiliation) questions after the task as well as the single-item self-reported measures used in Study 3. However, as in Study 3, a programming error resulted in a number of participants skipping these post-experiment questions related to Belief in God ($N = 60$; $M = 4.67$; $SD = 1.93$), Religious Beliefs ($N = 94$; $M = 4.62$; $SD = 2.15$), and God is Up belief ($N = 111$; $M = 5.19$; $SD = 1.81$).

Results

As in Meier et al. (2007), we subtracted the remembered location of the religious images from the remembered location of the neutral images presented at the same location, with positive (negative) numbers reflecting a higher (lower) vertical placement relative to neutral controls. As shown in Table 4, and replicating the effect of Meier et al. (2007), the God-like images were recalled as appearing higher than the Devil-like images, $t(124) = 4.15$, $p < .001$, $d = .37$ (95% CI: .19, .55).

Table 4. *Original (Meier et al., 2007, Study 3) and Replication (Study 4) Means and Standard Deviations of the Original Vertical Placement of God-Like and Devil-Like Images*

Image Type	Vertical Placement Recall	
	Meier et al., 2007, Study 3	Replication, Study 4
God-Like	36.04 (121.14)	8.87 (128.93)
Devil-Like	-21.02 (113.70)	-52.11 (138.19)

Next, we examined potential moderators. As in the prior studies, we created a difference score by subtracting the average location chosen for Devil-like images from the average location

chosen for God-like images, with positive scores indicative of metaphor-consistent biases. We then correlated this difference score with the potential moderators. Participants who believed in God to a greater extent exhibited metaphor-consistent memory biases to a greater extent, though this relationship did not reach conventional level of significance, $r(58) = .25$ (95% CI: -.004, .47), $p = .05$. Correlations with the religiosity, $r(92) = .09$ (95% CI: -.11, .29), $p = .38$, and God is up, $r(109) = .19$ (95% CI: .004, .36), $p = .21$, items were not significant.

General Discussion

We conducted a series of replication and extension attempts based upon the studies from Meier et al. (2007). Our results revealed successful replications of effects that involved cognitive tasks, such as reaction-time based categorizations and memories for God-like images that varied in vertical location. However, effects that involved the perception of others' belief in God failed to replicate. Also, individual differences in belief in God, religiosity, or belief that God is up did not consistently moderate the associations that were found. We discuss the implications of this collection of findings below.

Replication Implications

The results bear upon the replicability of metaphoric and embodied effects, including the realms of religious cognition and symbolism. There have been fewer attempted replications of embodied or conceptual metaphor effects that involve cognitive tasks (e.g., those using reaction time metrics) relative to perception tasks involving the judgment of others, the latter of which often invoke the label of "social priming". The current replications span both domains in the context of a single phenomenon - God's purported verticality. That the cognitive tasks displayed consistent evidence for metaphoric associations and the social judgments tasks did not accord with the idea that cognitive priming seems to be the more robust phenomenon (O'Donnell et al.,

2018). Thus, in assessing the replicability of metaphor-related (and embodied) phenomena, it may be important to consider methodological details as some procedures and foci are likely to produce more robust effects than others.

Our failure to replicate social judgment effects could be due multiple reasons: (1) vertical location might not affect perceptions of another's belief in God, (2) the designs used might be unreliable, or (3) there are unknown moderator variables. All of these factors and others are possible (e.g., Hauser & Schwarz, in press). Yet, other studies have found support for metaphor-linked social perception effects that are probably reliable. For example, Persich et al. (in press) examined metaphoric associations between brightness and God-related concepts. In one study, participants rated individuals wearing white shirts or dresses as more religious than those wearing black shirts or dresses. Such results are consistent with metaphors linking religiosity to lightness (rather than darkness) and we suspect that these effects would replicate. This is because people dress themselves and the colors of clothing that they wear can therefore be ascribed to personal preferences to some legitimate extent. Under such circumstances, clothing color can quite reasonably be used as a cue to the beliefs or personality of the person wearing the clothing (Gosling, 2008). In the present case, by contrast, targets did not choose vertical locations to have their pictures presented in and vertical locations on a computer screen probably do not reveal anything about a person. In other words, picture verticality was simply not “applicable” to the targets (Higgins, 1996). Future work can examine whether such attribution-related explanations can distinguish social priming effects that are likely to replicate from those that have a more fragile basis to them.

Individual Differences

Individual differences in belief in God, religiosity, and endorsement of the idea that God is up did not consistently moderate the present effects, regardless of which type of task was involved. Other work, by contrast, has shown that individual differences sometimes do moderate metaphoric and embodied effects (Fetterman et al., 2016, 2018; Persich et al., in press). For example, people who preferred white to black and lighter to darker rooms reported that they believed in God to a greater extent in several of Persich et al.'s (in press) studies. As another example, Fetterman et al. (2018) showed that higher feelings of physical warmth in a daily report study predicted higher levels of agreeable behavior (e.g., being nice, friendly, & caring on a given day), which implicates personality processes of a within-person type.

Thus, it is clear that individual differences can moderate metaphor effects, even in religious domains, though they did not generally do so in the current studies. One reasonable interpretation of these findings is that implicit cognitive associations linking God concepts to higher levels of vertical space are culturally shared, if not somewhat universal (Guan et al., in press; Kövecses, 2005), and thus even those with lesser beliefs in God would still exhibit the relevant patterns or associations. For example, Guan et al. (in press) found that participants from Tibet linked Buddhist-related words with higher locations in space. An additional possibility, though, is that there are individual difference moderators that were not assessed. In this connection, Persich et al. (in press) and Fetterman et al. (2018) created individual difference measures that were closely associated with the symbolic phenomenon of interest (e.g., brightness or warmth), but this individual difference strategy was not fully employed in the present studies. For example, one might tap individual differences in preferences for low versus high vertical spaces to determine if they moderate effects like the ones found here. Accordingly, we conclude

that there are opportunities for advancement in thinking about, and documenting, the manner in which individual differences are key to embodied phenomena (Fetterman et al., 2021).

Limitations

The current research does have limitations. First, we made errors in our sample size estimates as well as in some of the programming, which resulted in lost data. Yet, in terms of statistical power, we had adequate power, in all studies, to find small effects. The *Ns* for the analyses with the individual differences varied across studies and items, though, and this could have affected results involving these variables.

A second limitation is that the replication attempts carried out in the present context involved three authors who helped carry out the original research (though a fourth author was not involved in the original research). This allowed us to approach the studies with care and deliberation, though not as fully independent scientists. To deal with such issues, we tried to be as transparent as possible in the conduct of the research. This extended to posting data collection plans as well as analysis plans at Open Science Framework website. It also extended to the transparent reporting of all data exclusions, manipulations, and measures. Importantly, we conclude that verticality manipulations do not influence social perceptions, though we had originally believed findings of this type to be robust.

Conclusion

In four studies, we successfully replicated effects from Studies 1 and 3, but not Studies 4 or 5, of Meier et al. (2007). These results reveal that metaphors for God that involve vertical space reflect the ways in which people think about God in implicit cognitive terms. Yet, it does not appear that people use associations of this type when explicitly judging the belief in God of unknown others. Attempts to investigate and replicate metaphor and embodied effects might

consider systematically examining the cognitive versus social judgement distinction highlighted in the present work.

References

- Ashton, M. C., Pilkington, A. C., & Lee, K. (2014). Do prosocial people prefer sweet-tasting foods? An attempted replication of Meier, Moeller, Riemer-Peltz, and Robinson (2012). *Journal of Research in Personality, 52*, 42-46.
- Barsalou, L. W., A. K. Barbey, W. K. Simmons, & Santos, A. (2005). Embodiment in religious knowledge. *Journal of Cognition and Culture, 5*, 14-57.
- Chasteen, A. L. Burdzy, D. C. Pratt, J. (2010). Thinking of God moves attention. *Neuropsychologia, 48*, 627-630.
- Chabris, C. F., Heck, P. R., Mandart, J., Benjamin, D. J., & Simons, D. J. (2019). No evidence that experiencing physical warmth promotes interpersonal warmth: Two failures to replicate. *Social Psychology, 50*, 127-132.
- DesCamp, M. T., & Sweetser, E. E. (2005). Metaphors for God: Why and how our choices matter for humans? The application of contemporary cognitive linguistics research to the debate on God and metaphor. *Pastoral Psychology, 53*, 207-238.
- Fetterman, A. K. (2016). On God-belief and feeling clean: Daily experiences are related to feeling clean, particularly for those high in God-belief. *Social Psychological and Personality Science, 7*, 552-559.
- Fetterman, A. K., Bair, J. L., Werth, M., Landkammer, F., & Robinson, M. D. (2016). The scope and consequences of metaphoric thinking: Using individual differences in metaphor usage to understand how metaphor functions. *Journal of Personality and Social Psychology, 110*, 458-476.

- Fetterman, A. K., Meier, B. P., & Robinson, M. D. (2019). Dispositional properties of metaphor: The predictive power of sweet taste metaphor for trait and daily prosociality. *Journal of Individual Differences, 38*, 175-188.
- Fetterman, A. K., Robinson, M. D., & Meier, B. P. (2021). Leveraging individual differences to understand grounded procedures. *Behavioral and Brain Sciences, 44*, E6. doi: 10.1017/S0140525X20000382.
- Fetterman, A. K., Wilkowski, B. M., & Robinson, M. D. (2018). On feeling warm and being warm: Daily perceptions of physical warmth fluctuate with interpersonal warmth. *Social Psychological and Personality Science, 9*, 560-567.
- Fincher-Kiefer, R. (2019). *How the body shapes knowledge: Empirical support for embodied cognition*. Washington, DC: American Psychological Association.
- Gosling, S. (2008). *Snoop: What your stuff says about you*. New York, NY: Basic Books.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology, 74*, 1464-1480.
- Guan, F., Xu, Z., Zhang, B., Yan, W., & Peng, K. (in press). Bidirectional associations between Tibetan Buddhism and vertical space. *Psychology of Religion and Spirituality*.
- Hauser, D. J. & Schwarz, N. (in press). Situated embodiment: When physical weight does and does not inform judgments of importance. *Social Psychological and Personality Science*.
- Higgins, E. T. (1996). *Knowledge activation: Accessibility, applicability, and salience*. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (p. 133-168). New York, NY: The Guilford Press.

- Klein, R. A., Vianello, M., Hasselman, F., et al. (2018). Many labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, 4, 443-490.
- Kövecses, Z. (2005). *Metaphor in culture: Universality and variation*. New York, NY: Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago, IL: University of Chicago Press.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenges to western thought*. New York, NY: Basic Books.
- Landau, M. J. (2017). *Conceptual metaphor in social psychology: The poetics of everyday life*. New York, NY: Routledge.
- Landau, M. J., Meier, B. P., & Keefer, L. A. (2010). A metaphor-enriched social cognition. *Psychological Bulletin*, 136, 1045-1067.
- Lee, S. W. S. (2016). Multimodal priming of abstract constructs. *Current Opinion in Psychology*, 12, 37-44.
- Martinez, A. M., & Benavente, R. (1998). *The AR face database* (CVC Tech. Rep. No. 24): <http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html>.
- Meier, B. P., & Fetterman, A. K. (in press). Metaphors for God: God is high, light, and human in implicit tasks. *Psychology of Religion and Spirituality*.
- Meier, B. P., Fetterman, A. K., & Robinson, M. D. (2015). Black and white as valence cues: A large-scale replication effort of Meier, Robinson, and Clore (2004). *Social Psychology*, 46, 174-178.

- Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What's "up" with God? Vertical Space as a representation of the divine. *Journal of Personality and Social Psychology, 93*, 699-710.
- Meier, B. P., Schnall, S., Schwarz, N., & Bargh, J. A. (2012). Embodiment in social psychology. *Topics in Cognitive Science, 4*, 705-716.
- Meyer, D. E., & Schvaneveldt, R. W. (1976). Meaning, memory structure, and mental processes. *Science, 192*, 27-33.
- O'Donnell, M., Nelson, L. D., Ackermann, E., Aczel, B., Akhtar, A., Aldrovandi, S., ... Zrubka, M. (2018). Registered replication report: Dijksterhuis and van Knippenberg (1998). *Perspectives on Psychological Science, 13*, 268–294.
- Persich, M. R., Steinemann, S., Fetterman, A. K., & Robinson, M. D. (in press). Drawn to the light: Predicting religiosity using "God is light" metaphor. *Psychology of Religion and Spirituality*.
- Ransom, M. R. & Alicke, M. D. (2013). On bended knee: Embodiment and religious judgments. *Current Research in Social Psychology, 21*, <https://crisp.org.uiowa.edu/sites/crisp.org.uiowa.edu/files/2020-04/21.9.pdf>.
- Giessner, S. R., & Schubert, T. W. (2019). Power and vertical positions in an organization chart: A pre-registered replication report of Study 3a and a modification of Study 1a, Giessner & Schubert (2007). *Collabra: Psychology, 5*, 25. <https://doi.org/10.1525/collabra.216>
- Siev, J., Zuckerman, S. E., & Siev, J. J. (2018). The relationship between immorality and cleansing: A meta-analysis of the Macbeth Effect. *Social Psychology, 49*, 303-309.

- Skulmowski, A., & Rey, G. D. (2018). Adjusting sample sizes for different categories of embodied cognition research. *Frontiers in Psychology, 9*, doi: 10.3389/fpsyg.2018.02384.
- Soliman, T. M., Johnson, K. A., & Song, H. (2015). It's not "all in your head": Understanding religion from an embodied cognition perspective. *Perspectives on Psychological Science, 10*, 852-864.
- Van Cappellen, P., & Edwards, M. E. (2021). The embodiment of worship: Relations among postural, psychological, and physiological aspects of religious practice. *Journal for the Cognitive Science of Religion, 6*, 56–79.

Footnotes

¹Two data files for Study 2 were uploaded to OSF (<https://osf.io/tgb9j/>). One file presents information from the study proper and the second presents age and gender information only. Due to a programming error, that is, we could not link particular genders or ages to particular participants who completed the substantive task for the study.