Investigating the N400-Component as an Implicit Measure of Supernatural Belief

Brousianou, Athina

Investigating the N400-Component as an Implicit Measure of Supernatural Belief

Athina V. Brousianou

Department of Cognitive Psychology, Universiteit Leiden

Research Master Thesis Cognitive Neuroscience

Dr. Michiel van Elk

Word count: 8429

Abstract

Despite the increase of secularism, implicit religious and spiritual (jointly termed "supernatural") reasoning persists even in individuals who identify as nonbelievers; however, previous studies have used behavioral implicit measures to examine differences between implicit and explicit belief. The aim of this research was to investigate these differences in atheist and spiritual-but-not-religious (SBNR) individuals using event-related potentials, and explore the N400-effect as an implicit marker of (un)belief. We conducted two studies: Study 1A (N = 101) examined supernatural attitudes and study 1B (N = 109) examined supernatural beliefs. Participants were presented with positive/negative (1A) and belief/disbelief (1B) religious, spiritual, and control statements, and were asked to provide their explicit evaluation (i.e., agree/disagree) of the statements while ERPs were being recorded. The results showed inconclusive evidence of the utility of the N400 as an implicit measure of belief. Specifically, we did not find significant differences between atheists and SBNR individuals in the N400-amplitude between positive and negative (1A), or belief and disbelief (1B) religious, spiritual, and control statements. In light of the lack of statistically significant results, exploratory analyses were performed, and implications and limitations of our study, as well as suggestions for future research, were discussed.

Keywords: N400, event-related potentials, EEG, unbelief, spirituality, religiosity, secularism, naturalness of religion hypothesis, dual-process account

Investigating the N400-Component as an Implicit Measure of Supernatural Belief

Although religion has historically been ubiquitous, we have been observing a gradual shift towards "unbelief" (i.e., an umbrella term used to describe lack of belief in a God; Lindeman et al., 2019), as the number of people who identify as unbelievers is rapidly growing worldwide (Geertz & Markússon, 2010; Zuckerman, 2007). However, there is also evidence that religious and spiritual (jointly termed "supernatural") reasoning persists among unbelievers, as demonstrated by research that has examined (un)belief on an implicit level (e.g., Bering, 2002; Kelemen et al., 2013; Lindeman et al., 2014). Religiosity refers to a set of beliefs and practices prescribed by organized institutions (Mytko & Knight, 1999), whereas spirituality refers to subjective perceptions of meaning-making and universal (inter)connectedness (de Jager Meezenbroek et al., 2012). The conflict between people's explicit denial of supernatural beliefs and implicit evidence of supernatural thinking poses an enigma that this study sought to elucidate. In the following section we will provide an overview of factors contributing to religiosity and secularism, as well as of research on explicit and implicit supernatural beliefs. Lastly, we will introduce event-related potential (ERP) research as a way to implicitly measure belief. In this study, we aimed to investigate the extent to which explicit supernatural attitudes and beliefs correspond to implicit ones, as measured by electroencephalography (EEG), exploring the N400-effect as a marker of implicit belief.

What drives the rise of religious unbelief? A sociocultural explanation relates to increased existential security (Norenzayan & Gervais, 2013; Zuckerman, 2007), as historically, and experimentally, existential threats have increased religious belief (Gray & Wegner, 2010). Furthermore, countries with a high average socioeconomic status and a strong welfare state, such as Denmark, Finland, and the Netherlands, are exhibiting ever-decreasing

religious levels (Pew Research Center, 2018), further supporting the thesis that existential security might be linked to increased unbelief.

The lack of religious exposure and credibility enhancing displays (CREDs) could also be potential factors decreasing religious belief (Gervais et al., 2011). CREDs are sociocultural indicators that encourage theistic belief by increasing the credibility of religious expressions (Henrich, 2009). Individuals who are exposed to religion, compared to ones that are not, are more likely to be religious. Similarly, beliefs that are supported by CREDs are more likely to be transmitted than unsupported beliefs (Henrich, 2009; Hitzeman & Wastell, 2017). Therefore, religion does not effectively propagate in an environment where individuals are not exposed to religion and/or where CREDs are minimal or nonexistent (Gervais et al., 2011; see also, Maij et al., 2017).

In addition to cultural factors, the cognitive science of religion (Lawson, 2000) sees unbelief as the product of effortful cognitive processes. In specific, the dual-process theory (Evans, 2008) posits that religious belief is related to an intuitive processing system, which is automatic, reflexive, and fast. Conversely, unbelief is suggested to be the result of a system which is effortful, analytical, and reflective (Barrett, 2000; Boyer, 2003; Kalkman, 2014). Overall, proponents of the dual-process theory of religion have suggested that religion is the "default" cognitive state, predisposing people towards supernatural reasoning. In this view, unreligion originates from the effortful overriding of the intuitive system and the active engagement of the analytical system (Barrett, 2000; Kalkman, 2014). It is, therefore, hypothesized that unbelievers might have learned to suppress or bypass these default cognitions, enabling them to critically reject convictions that are not rationally supported (Dawkins, 2006; see also, Lindeman et al., 2013).

In support of the dual-process theory, multiple studies have demonstrated that a more intuitive cognitive style was predictive of stronger religious beliefs, while a more analytical

processing style was predictive of religious disbelief (Gervais & Norenzayan, 2012; Shenhav et al., 2012). Furthermore, experimentally inducing an intuitive compared to an analytic processing style increased self-reported religious belief (Shenhav et al., 2012), while experimentally inducing an analytic cognitive style promoted religious disbelief, even in religious participants (Gervais & Norenzayan, 2012; see also, Pennycook et al., 2012; cf. Razmyar & Reeve, 2013; Yonker et al., 2016).[1]

From an evolutionary standpoint, religiosity has evolved not only because it is adaptative in and of itself (Bulbulia, 2004; see also, Johnson, 2005; Sedikides & Gebauer, 2010; Sosis, 2003; Wilson, 2002), but because our brains are developed in such a way that predisposes people towards cognitive biases and, subsequently, supernatural reasoning (Bering, 2006; Bloom, 2005; Boyer, 2001, 2003). For example, it is hypothesized that people have developed a hyperactive agency detection device which biases them towards detecting agency in their environment, even in the absence of one (Barrett, 2004; Guthrie, 1993; cf. Van Leeuwen & van Elk, 2019). This would have been highly advantageous in our evolutionary past, as not being sensitive to the detection of a potentially dangerous agent had grave consequences.

Cognitive biases tend to emerge with the processing style of the intuitive system (Evans, 2008); in fact, it has been suggested that supernatural beliefs emerge *through* cognitive biases (Svedholm & Lindeman, 2013; Willard & Norenzayan, 2013). For instance, manipulations that induced intuitive processing were shown to increase teleological thinking (i.e., viewing things as existing for a purpose; (Kelemen & Rosset, 2009; see also, Kelemen,

[1] One of the standing criticisms of such manipulations relate to the measurement order. Authors have suggested that measuring one's processing style prior to measuring religious belief (e.g., Gervais & Norenzayan, 2012; Shenhav et al., 2012) inflates the negative relationship between analytic thinking and religiosity. For example, Finley and colleagues (2015) replicated the study by Gervais and Norenzayan (2012), but measured religious belief *before* they measured participants' processing style, finding evidence for measurement order effects. Furthermore, other studies have *not* found a negative or significant relationship between an analytical processing style and religiosity (Sanchez et al., 2017). Nonetheless, a recent meta-analysis has demonstrated that a reflective cognitive style is more associated with atheists than with religious believers (Pennycook et al., 2016).

1999). Furthermore, mind-body dualism (i.e., the view of the mind as separate from the body), teleological thinking and ontological confusions (i.e., confusing core knowledge properties, for example, thinking that an inanimate entity is an intentional agent) have all been associated with higher rather than lower religiosity (e.g., Lindeman et al., 2015; Riekki et al., 2013; Willard & Norenzayan, 2013; see also, Heywood & Bering, 2014).

Furthermore, studies have found a positive association between mentalizing abilities (e.g., being able to reason about intentional agents) and theistic belief (Kapogiannis et al., 2009; Norenzayan et al., 2012), and developmental researchers have argued that mentalizing abilities play a central role in the development and maintenance of religious beliefs (Bering, 2002; Gervais, 2013). Similarly, theory of mind (i.e., attributing mental states to others) has been imperative to our evolutionary success and such mental processes cannot easily be overridden when, for instance, we reason about the subjective experiences of the dead (for a review, see Bering, 2006). Supernatural reasoning and beliefs seem to be easily adopted and are, therefore, considered a product of the human evolution (Bering & Bjorklund, 2004; for a review, see Gervais, 2013). Thus, theorists have suggested the naturalness of religion hypothesis (NRH): The premise that religion is natural, intuitive, and automatic (Barrett, 2000; Barrett & Lanman, 2008; Bloom, 2007; Boyer, 1994). Does that mean that atheists might be more religious or spiritual than what they claim to be?

Some argue that supernatural reasoning is a stable and typical characteristic of the human cognition (Legare et al., 2012) and hypothesize that "regardless of their explicit, reflective disavowal of belief in supernatural agents, at a non-reflective level of processing, people enduringly remain 'intuitive theists'" (Järnefelt et al., 2015, p. 75; see also, Hitzeman & Wastell, 2017; Kelemen, 2004). There is ample evidence showing that non-believers frequently display implicit patterns of cognition that contrast their explicit beliefs. For instance, even physical scientists engage, under speeded conditions, in teleological

explanations of natural phenomena (Kelemen et al., 2013; see also, Banerjee & Bloom, 2014; Heywood & Bering, 2014; Järnefelt et al., 2015). Furthermore, research has shown that people who believe that deceased people are no longer conscious sometimes still spoke about a dead person in a manner that conflicted with their explicit beliefs (e.g., argued that dead persons demonstrate knowledge of being deceased; Bering, 2002; see also, Bek & Lock, 2011; Georgiadou & Pnevmatikos, 2019; Pereira et al., 2012). Moreover, Finnish atheists showed similar levels of physiological arousal to believers (as measured by skin conductance) when they challenged God to harm their beloved ones (Lindeman et al., 2014). Finally, there is evidence that certain non-believers endorse at least some supernatural beliefs (Lindeman et al., 2019, 2020). For example, 'analytic atheists' were shown to endorse almost no supernatural beliefs, whereas spiritual-but-not-religious (SBNR) individuals endorsed all examined supernatural beliefs besides the belief in God's existence (Lindeman et al., 2019).

Is religion, therefore, an ingrained feature of human cognition, as suggested by the NRH (Barrett, 2000; Bloom, 2007), or have we evolved past a religious society (Cohen, 2009; Norris & Inglehart, 2004)? Previous research has demonstrated relationships between explicit and implicit religiosity (LaBouff et al., 2010; Ross et al., 2020; see also, Jong et al., 2017), but these studies have largely used *behavioral* implicit measures, such as the implicit association test (Greenwald et al., 1998; LaBouff et al., 2010) and the affect misattribution procedure (Payne & Lundberg, 2014; Ross et al., 2020). A major advantage of utilizing event-related potential (ERP) research, compared to other implicit measures, is that ERPs are elicited automatically and do not require an overt response from the participants; for that reason, they are considered a prime candidate for implicit measurement (Frühholz et al., 2011).

The ERP of specific interest to this study is the N400-component, which is negative-going ERP—compared to the baseline—and it is elicited in response to reading words, with

its peak amplitude observed around 400 ms post-stimulus presentation. The N400 has been implicated, among others, in language processing and meaning making (for a review, see Kutas & Federmeier, 2011). Specifically, the N400-effect is linked to meaning comprehension of a word in its context, with larger N400-effects related to more difficult or amplified processing (Kutas & Federmeier, 2011). For example, the classic study by Hagoort and colleagues (2004) demonstrated larger N400-effects for sentences that violated not only semantic integration, such as "Dutch trains are *sour*", but also ones that violated world knowledge, such as "Dutch trains are *white*" (Dutch trains are famously yellow; see also, van Elk et al., 2008, 2010).

Additionally, the N400-component has recently been conceptualized as an implicit marker of valuation and worldviews, as research has demonstrated that an N400-effect is observed when a statement clashes with one's explicit beliefs or attitudes (Galli et al., 2017; Morris et al., 2003; Van Berkum et al., 2009). For example, Van Berkum and colleagues (2009) showed that statements that conflicted with one's personal (moral) values produced a larger N400-effect than value-congruent statements. Similarly, Galli and colleagues (2017) demonstrated that the N400 was observed when a pro- or against-EU statement was incongruent with one's explicit view regarding the topic.

Based on these studies, we could hypothesize that a statement which is inconsistent compared to consistent with one's worldviews or values induces effortful processing, producing a larger N400-effect (Van Berkum et al., 2009). Correspondingly, such an effect might also be observed when a statement clashes with one's supernatural attitudes or beliefs.

This study aimed to examine the NRH by investigating whether atheist and SBNR individuals endorse implicit spiritual attitudes (study 1A) and beliefs (study 1B) to the same or different extent than they explicitly state. Additionally, we examined the utility of the N400 as a measure of implicit belief. Specifically, in study 1A we presented participants with

positive and negative religious, spiritual and control *attitude* statements while ERPs where being recorded, and we asked them to provide their explicit evaluation, that is, their agreement or disagreement with the statement. In study 1B, we presented participants with positively-phrased (i.e., belief) and negatively-phrased (i.e., disbelief) religious, spiritual, and control *belief* statements while ERPs where being recorded, and we asked them to provide their explicit evaluation (i.e., agree/disagree).

We hypothesized that the N400-response to attitude and belief statements would differ between atheist and SBNR participants in a way that is consistent with their explicit endorsements. Specifically, the null hypothesis for study 1A was that ($H0_A$) the N400-difference between positive/belief and negative/disbelief evaluations of spiritual attitude statements would be the same between atheists and SBNR individuals. Our alternative hypotheses stated that (H_A) the N400-effect of spiritual attitude statements would differ between atheist and SBNR individuals; specifically, we predicted that ($H1_A$) the N400-effect would be observed in *SBNR* individuals for negative rather than positive evaluations of spiritual statements; ($H2_A$) the N400-effect would be observed in *atheists* for positive rather than negative evaluations of spiritual statements; and ($H3_A$) the size of the N400-effect would be correlated with the explicit evaluation of the spiritual attitudes. For study 1B, the null hypothesis stated that ($H0_B$) the N400-difference between belief and disbelief evaluations of spiritual attitude statements would be the same between atheists and SBNR individuals. Our alternative hypotheses stated that (H_B) the N400-effect of spiritual belief statements would differ between atheist and SBNR individuals; specifically, we predicted that ($H1_B$) the N400-effect would be observed in *SBNR* individuals for disbelief rather than belief evaluations of spiritual statements; ($H2_B$) the N400-effect would be observed in *atheists* for belief rather than disbelief evaluations of spiritual statements; and ($H3_B$) the size of the N400-effect would be correlated with the explicit evaluation of the spiritual beliefs.

Methods

The study design, hypotheses, and analysis plan were preregistered on the Open Science Framework prior to conducting this study. Due to oversight or necessity during the setting up of the study certain changes were made to some preregistered elements. Important changes are mentioned in-text (for an overview, see https://osf.io/g5fmq/).

Materials

All stimuli were presented to the participants in Dutch. For the generation of the stimuli, two pretests were conducted (one for study 1A and one for study 1B) in which 40 attitude (1A) and belief (1B) statements were generated per category (i.e., religious, spiritual, control). Participants were asked to rate the statements on a number of dimensions, including their agreement with the statement, their (positive or negative) evaluation, and whether it related to a religious, spiritual, or neutral topic. Word length and lexical frequency of the target words were matched for each category by using Celex (http://celex.mpi.nl/). In the pretest for study 1A, 60 participants were recruited (females n = 44, mode age 18–24 [n = 42]). The majority had obtained a higher education degree (n = 25), followed no religious denomination (n = 40), and self-identified as non-believers (n = 22), atheists (n = 11), and agnostics (n = 10). The pretest for study 1B contained 65 participants (females n = 53, mode age 18–24 [n = 50]). The majority had obtained a higher education degree (n = 33), followed no religious denomination (n = 46), and self-identified as non-believers (n = 24) and atheists (n = 12).

From these pretests, 25 attitude (study 1A) and 20 belief (study 1B) statements were selected per category. The criteria for selecting a statement included sufficient rater agreement that the statement belonged to a certain category, sufficient variation between the statement's category and the other categories, and sufficient variability in the raters' evaluation of the statement. See Table 1 for example stimuli used in studies 1A and 1B.

Table 1

Example Stimuli Used in Studies 1A and 1B

	Study 1A	
Statement	Positive	Negative
Religious	Praying to God daily is **recommended**.	Praying to God daily is **discouraged**.
	Believing that God can hear our prayers is **normal**.	Believing that God can hear our prayers is **strange**.
Spiritual	Meditating daily is **recommended**.	Meditating daily is **discouraged**.
	Belief in a spiritual energy is **advantageous**.	Belief in a spiritual energy is **disadvantageous**.
Control	Going to the movies daily is **advantageous**.	Going to the movies daily is **disadvantageous**.
	Believing politicians is **sensical**.	Believing politicians is **nonsensical**.
	Study 1B	
Statement	Belief	Disbelief
Religious	I do believe in the existence of **God**.	I do not believe in the existence of **God**.
	I do believe in the existence of **Satan**.	I do not believe in the existence of **Satan**.
Spiritual	I do believe in the existence of **spirits**.	I do not believe in the existence of **spirits**.
	I do believe in the existence of **karma**.	I do not believe in the existence of **karma**.
Control	I do believe in the existence of **Pikachu**.	I do not believe in the existence of **Pikachu**.
	I do believe in the existence of **Superman**.	I do not believe in the existence of **Superman**.

Note. The phrasing of the statements in both studies was positive (or belief) in 50% of the trials, and negative (or disbelief) in the rest. In study 1B, the positively phrased items referred to belief statements, while the negatively phrased items referred to disbelief statements. Critical words to which the ERPs were time-locked are marked in bold.

Alongside providing their explicit agreement/disagreement with the attitude (1A) and belief (1B) statements, participants were additionally asked to complete the explicit belief measure (EBM; Lindeman et al., 2019). This measure is a 9-item scale measuring participants' agreement to the following supernatural belief statements on a 5-point Likert scale (1 = *completely disagree;* 5 = *completely agree*): "I believe in God"; "I believe that there is life after death"; "The universe originated from intelligent design"; "The universe has

an ultimate purpose"; "I believe in fate"; "There is spiritual energy in the universe"; "In the universe, everything is connected in a way that cannot be explained scientifically"; "Telepathic mind reading is possible"; and "I believe in angels". A higher score signifies a stronger belief in supernatural entities and phenomena. Participants were also asked to indicate their age, gender, primary occupation, level of education, and their subjective SES (MacArthur ladder; Adler et al., 2000).

Participants

Participants were recruited from the online study pool of the University of Amsterdam (UvA), as well as via advertisements in local newspapers. The advertisements requested Dutch speakers that are interested to participate in research related to supernatural phenomena. Prior to the study, we ran a correlation power analysis (G*Power version 3.1; Erdfelder et al., 2009) with the purpose of identifying an optimal sample size. The analysis yielded a sample size of 67 in order to achieve a modest correlation of $\rho = 0.30$ (power = .80, $\alpha = .05$) between the explicit participant evaluations and the N400-effect. Data collection stopped when data from 125 participants was collected; most ($n = 76$) were locals. The aim was to have usable data from 100 participants; this not only allowed for adequate power of detecting group differences, but it also accounted for data that was too noisy to be used in the analysis. Upon completion of the study, participants were compensated with €20; UvA students could choose study credit instead.

Participants were excluded prior to the analysis based on visual inspection of the raw EEG data. Participants were further excluded when more than five electrodes needed to be interpolated, or when more than 20% of all trials were rejected based on an automated artifact rejection procedure. In study 1A, 24 participants were excluded, and 16 participants in study 1B (part of the excluded/remaining participants overlapped between the two studies).

Participant Demographics Based on Study

Study 1A – Supernatural Attitudes. After exclusions, 101 participants were used in the analysis (females *n* = 76, mean age 36.61 years, *SD* = 15.90, range = 18–73). Forty-four participants were in employment, 38 were students, while the rest (*n* = 19) did not specify their primary occupation. Forty-nine participants had completed a university or college degree. The majority (*n* = 33) assessed their socioeconomic status (SES) as being on the 7th step of the MacArthur ladder, followed by the 8th step (*n* = 30). Finally, participants selected their religious or spiritual identity from a drop-down list containing the following identities: atheist, unbeliever, agnostic, non-believer, secular, religious, SBNR, spiritual seeker, and other. Most participants self-identified as SBNR (*n* = 33), religious (*n* = 14), and non-believer (*n* = 12).

Study 1B – Supernatural Beliefs. After exclusions, 109 participants were used in the analysis (females *n* = 81, mean age 35.37 years, *SD* = 15.65, range = 18–73). Forty-five were in employment, 45 were students, and the rest did not specify their primary occupation. Fifty participants had completed a university or college degree. The majority (*n* = 38) assessed their socioeconomic status (SES) as being on the 7th step of the MacArthur ladder, followed by the 8th step (*n* = 30). Finally, participants selected their religious or spiritual identity from the aforementioned drop-down list. Most participants self-identified as SBNR (*n* = 36), religious (*n* = 15), and non-believer (*n* = 14). Table 2 presents the selected spiritual identities in studies 1A and 1B.

Table 2

Spiritual Identity Frequencies in Studies 1A and 1B

	Study 1A	Study 1B
Selected spiritual identity	Frequency (%)	Frequency (%)
Atheist	6 (5.90)	7 (6.40)

	Study 1A	Study 1B
Selected spiritual identity	Frequency (%)	Frequency (%)
Unbeliever[a]	7 (6.90)	7 (6.40)
Agnostic	10 (9.90)	11 (10.10)
Non-believer[b]	12 (11.90)	14 (12.80)
Secular	1 (1)	1 (0.90)
Religious	14 (13.90)	15 (13.80)
SBNR	33 (32.70)	36 (33)
Spiritual seeker	11 (10.90)	11 (10.10)
Other	7 (6.90)	7 (6.40)
Total	101	109

Note. Note that most of the participants in studies 1A and 1B are overlapping. SBNR = spiritual but not religious.

[a]Against religion.

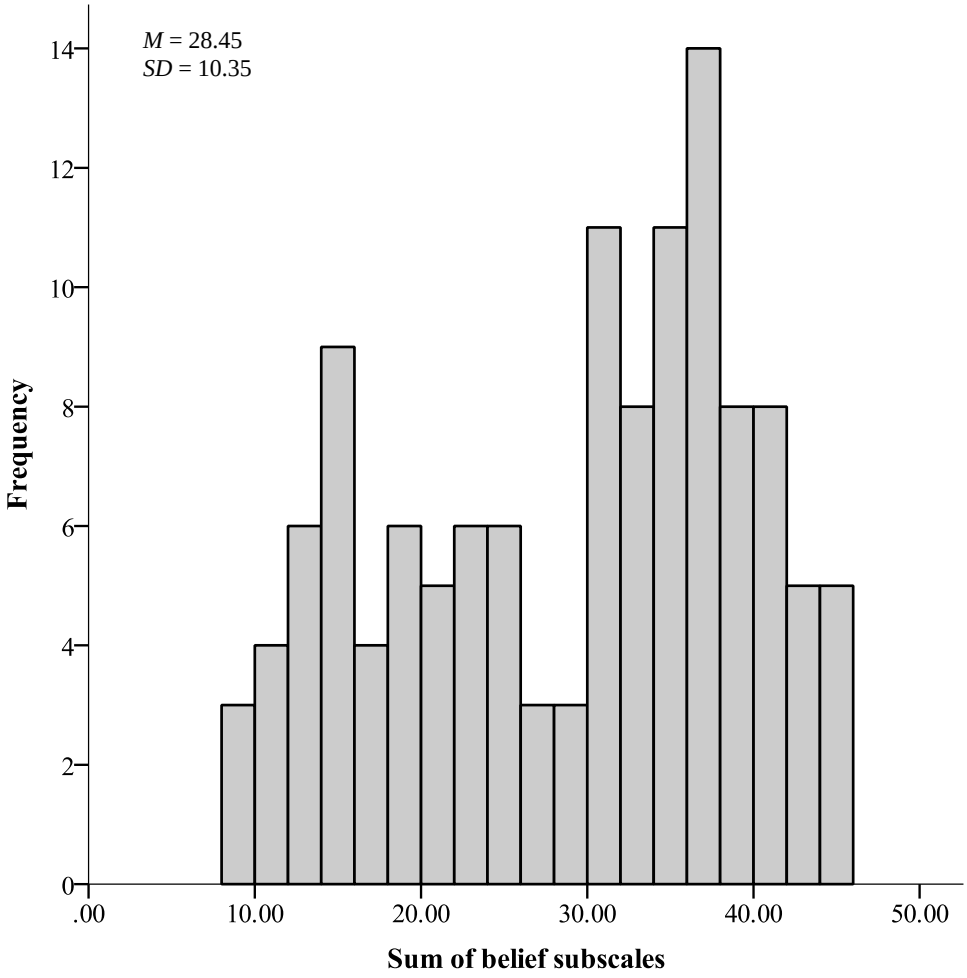
[b]Not necessarily against religion, but lacking beliefs.

Participant Allocation

In a previous study, participants were assigned an atheist or SBNR identity based on their responses on the EBM (Lindeman et al., 2019). Subsequently, the approximate scores of atheists and SBNR on the different belief subscales are known: Atheists are defined as scoring ≤ 2 on all of the belief variables; SBNR are defined as scoring > 2 on these scales. However, this distinction turned out to be too stringent for our study, as only seven participants in the whole dataset scored ≤ 2 on all of the subscales, resulting in too few atheists to perform meaningful comparisons. For that reason, we decided to utilize the bimodal distribution of the belief scores (see Figure 1) in order to split our sample into atheists and SBNR. To that end, participants for whom the sum of the subscales was lower than the mean (M = 28.45) were assigned to the atheist group, and participants scoring higher than the mean were assigned to the SBNR group. This resulted in 46 atheists and 55 SBNR in study 1A, and 50 atheists and 59 SBNR in study 1B.

Figure 1

Distribution of the Sum of the Belief Subscales



Note. N = 125. The distribution of the belief subscales is bimodal, naturally dividing the total sample into "low(er)" and "high(er)" believers.

Experimental Setup and Procedure

Participants were first asked to provide an informed consent of participation in the study. Study 1A was completed before 1B; at the end of study 1A, participants completed the EBM and relevant demographic questionnaires. Experimental blocks were separated by self-paced breaks. Study 1A consisted of five experimental blocks, with 30 trials/block, for a total of 150 trials: 25 positive and 25 negative evaluative statements x 3 conditions (religious, spiritual, control). Study 1B consisted of six experimental blocks, with 20 trials/block, for a

total of 120 trials: 20 belief (i.e., phrased in a positive way) and 20 disbelief statements (i.e., phrased in a negative way) x 3 conditions (religious, spiritual, control). In both studies, 50% of the statements were phrased in a positive (or belief) and 50% in a negative (or disbelief) way. Study instructions were provided before each experimental block began. The participants were further asked to sit in a comfortable position and avoid blinking during word presentation. Both studies begun with 10 practice trials to get participants acquainted with the task.

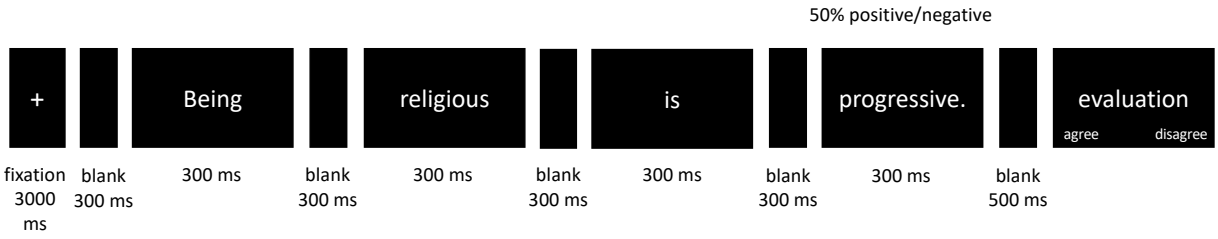
The experiment was run on a computer screen (1920 x 1080 pixels; 60cm diagonal; 60 Hz refresh rate), which was placed at approximately 40cm from the participants. The task was programmed using Presentation Software (version 19.0; Neurobehavioral Systems Inc., CA, USA, www.neurobs.com). All material was presented on a black background in white font (size 24) at the center of the screen.

Study 1A – Supernatural Attitudes

In study 1A, the EEG signal was measured while attitude statements were presented word-by-word (see Figure 2 for an overview of the experimental procedure). Trials started with a fixation cross for 3000 ms, followed by a blank screen for 300 ms. Each word was presented one-by-one for 300 ms, with a pause of 300 ms between each word. The target word, to which the ERP was time-locked, was always the last word (50% positive/negative), indicated by a full stop. A 500 ms blank screen followed. Participants were then asked to rate their agreement with the statement by pressing with their right index or middle finger the left or right response button (agree/disagree; counterbalanced across participants) on the button box attached to the EEG.

Figure 2

Experimental Procedure of Study 1A



Note. Each trial begun with a 3000 ms fixation cross, followed by a 300 ms blank screen. The statement was presented word-by-word; each word was on the screen for 300 ms, and each was followed by a blank screen for 300 ms, until the target, and final, word appeared for 300 ms (indicated by a full stop). The target word was positive in 50% of the trials (e.g., "progressive"), and negative in the rest (e.g., "obsolete"). A 500 ms blank screen followed. At the end of each trial, participants were asked to evaluate their agreement with the statement by pressing the left or right button (agree/disagree). This figure is not drawn to scale.

Study 1B – Supernatural Beliefs

In study 1B, the EEG signal was measured while belief statements were presented at the center of the screen (see Figure 3 for an overview of the experimental procedure). Each trial begun with a fixation cross for 1000 ms, followed by a 300 ms blank screen. Then, the statement "I do believe/do not believe in the existence of" (50% belief/disbelief) was presented for 1500 ms, followed by a fixation point underneath the statement for 1000 ms. Afterwards, the fixation disappeared and 300 ms later the target (and final) word, to which the ERP was time-locked, appeared below the statement for 1000 ms, indicated by a full stop. A 1000 ms blank screen followed, after which the participants were asked to rate their agreement with the statement by pressing with their right index or middle finger the left or right response button (agree/disagree; counterbalanced across participants) on the button box attached to the EEG.

Figure 3

Experimental Procedure of Study 1B

Note. Each trial begun with a 1000 ms fixation cross, followed by a 300 ms blank screen. After this, the belief statement was presented for 1500 ms; in 50% of the trials a belief statement was presented (i.e., phrased in a positive way: "I do believe in the existence of"), while in the rest a disbelief statement was presented (i.e., phrased in a negative way: "I do not believe in the existence of"). Then, a fixation cross appeared underneath the statement for 1000 ms. After 300 ms, the target and final word appeared below the statement for 1000 ms (indicated by a full stop). A 1000 ms blank screen followed. After each trial, participants were asked to evaluate their agreement with the statement by pressing the left or right button (agree/disagree). This figure is not drawn to scale.

EEG Recording and Analysis

The Actiview system (BioSemi, Amsterdam, The Netherlands) was used for data acquisition. The EEG was recorded at 2048 Hz using 64 active electrodes placed according to the standard 10/20 system. The horizontal and vertical EOG was measured by the electrodes on the outer canthi and above and below the participant's left eye. Referencing was done offline to the average of all electrodes. The data was filtered with a low-cutoff of 1 Hz and a high-cutoff of 30 Hz. We excluded trials in which the amplitude of one of the electrodes exceeded ±150 μV, and those that were contaminated by eye-blinks, based on an automated artifact rejection procedure.

The averaged ERP amplitude per condition was the main focus of the EEG analysis. A −100 to 0 ms baseline correction was applied and data was segmented from −200 to 800 ms relative to target word onset. For the analysis of the N400-effect, we focused on a 3 x 3 cluster of electrodes around Cz, which is the area that typically displays the maximum N400-amplitude. The cluster contained electrodes FC1, C1, CP1, CPz, FC2, FCz, Cz, C2, and CP2. The averaged ERP amplitude was extracted for analysis from 350–450 ms following stimulus onset.

In both studies, the data was analyzed using a repeated measures analysis of variance (RM ANOVA). First, we analyzed the behavioral data (i.e., explicit evaluations of attitudes/beliefs) and then the implicit data (i.e., N400-effect). In both studies, the factors were Statement (religious/spiritual/control) and Evaluation (positive/negative; belief/disbelief). The factor Group (atheist/SBNR) was additionally added to explore group differences. The dependent variable in the explicit analysis was the proportion of agreement with the statement, and in the ERP analysis the N400-amplitude. The factor Electrode was included as an additional factor to investigate differences in the spatial distribution of the N400 between conditions.

In both studies, we expected a three-way interaction between the statement category, evaluation, and group. In order to investigate the directionality of the critical interactions, we conducted post-hoc RM ANOVAs and t-tests comparing group differences in the proportions of agreement with the statements (explicit analysis), or the N400-effect (i.e., the difference between negative/disbelief and positive/belief evaluations; ERP analysis), for each of the statement categories. Furthermore, we used a Pearson's correlation analysis to assess the relation between the N400-effect and the explicit evaluation of the attitude and belief statements. To this end, we calculated for each participant the N400-difference between negative/disbelief and positive/belief evaluations and then correlated this to the explicit

proportion of agreement with the spiritual attitudes (1A) and spiritual beliefs (1B). Finally, we ran a correlation analysis between the EBM and the proportions of agreement with supernatural beliefs (1B), in order to validate the relationship between the variables.

Results

For both studies, ANOVA assumptions were examined and, when necessary, corrected. A Greenhouse-Geisser (for $\varepsilon < .75$) or Huynh-Feldt (for $\varepsilon > .75$) correction was applied where needed to account for sphericity violations. This correction did not alter the results in any meaningful way.

Study 1A: Supernatural Attitudes

Explicit Attitudes

The explicit data on supernatural attitudes was examined for differences between groups and conditions. Figure 4 presents the mean proportion of agreement with each condition (Statement x Evaluation) for atheist and SBNR individuals.

A 3 x 2 x 2 RM ANOVA was conducted with the factors Statement (religious/spiritual/control), and Evaluation (positive/negative; referring to proportions of agreement with positive and negative statements) entered as within-subject factors, and Group (atheist/SBNR) as the between-subject factor. The analysis showed that all main and interaction effects were significant (see Table 3 for all RM ANOVA results). Specifically, we found significant main effects of Statement, $F(2, 198) = 12.11$, $\eta_p^2 = .11$, $p < .001$, and Evaluation, $F(1, 99) = 71.71$, $\eta_p^2 = .42$, $p < .001$, as well as an interaction between Statement and Evaluation, $F(1.83, 180.93) = 59.45$, $\eta_p^2 = .38$, $p < .001$. The between-subject effect of Group was not significant, $F(1, 99) = 0.46$, $\eta_p^2 = .01$, $p = .50$.

Furthermore, we found a significant interaction of Statement x Group, $F(2, 198) = 4.54$, $\eta_p^2 = .04$, $p = .01$, as well a significant interaction of Evaluation x Group, $F(1, 99) = 118.74$, $\eta_p^2 = .55$, $p < .001$. Finally, the critical three-way interaction between Statement,

Evaluation, and Group was also significant, $F(2, 198) = 52.58$, $\eta_p^2 = .35$, $p < .001$, meaning that the two groups differed in their agreement with the statements depending on the statements' category and evaluation. To elucidate these differences, we ran three post-hoc RM ANOVAs testing the factors Evaluation and Group separately for the religious, spiritual, and control statements (all corrected for multiple comparisons to $\alpha = .025$).

Regarding the *religious statements*, the main effect of Evaluation was not significant, $F(1, 99) = 3.55$, $\eta_p^2 = .04$, $p = .06$. However, we observed a significant interaction between Evaluation and Group, $F(1, 99) = 60.83$, $\eta_p^2 = .38$, $p < .001$. Post-hoc t-tests confirmed significant differences between the two groups on both levels of the statements' evaluation: $t_{pos}(99)$ (i.e., positive) $= -7.57$, $p < .001$; $t_{neg}(99)$ (i.e., negative) $= 7.09$, $p < .001$. The results indicated that SBNR people agreed significantly more with the positive religious statements, $M = .50$, $SD = .21$, than atheists did, $M = .22$, $SD = .16$. Conversely, atheists agreed significantly more with the negative religious statements, $M = .59$, $SD = .25$, than the SBNR, $M = .28$, $SD = .18$.

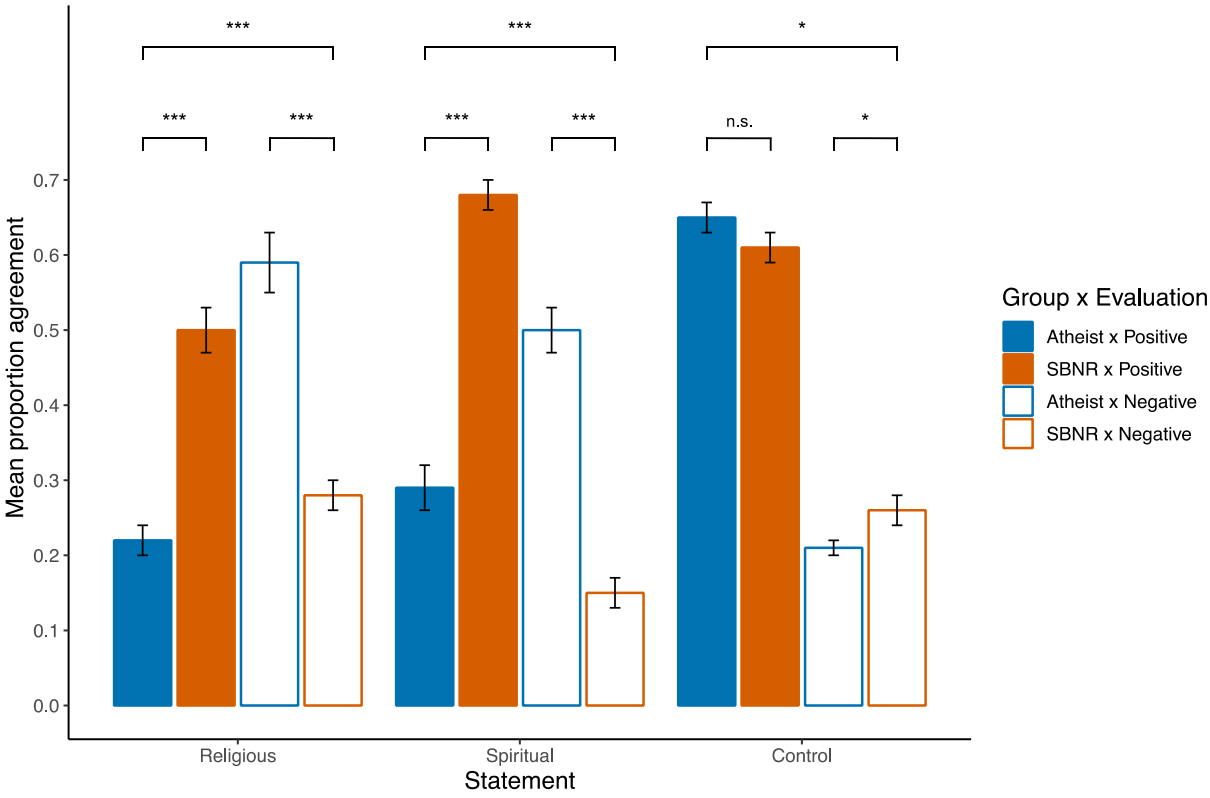
Regarding the *spiritual statements*, the results indicated a significant main effect of Evaluation, $F(1, 99) = 21.71$, $\eta_p^2 = .18$, $p < .001$, and a significant interaction of Evaluation x Group, $F(1, 99) = 124.89$, $\eta_p^2 = .56$, $p < .001$. Post-hoc t-tests showed significant differences between the groups on both levels of Evaluation: $t_{pos}(99) = -11.40$, $p < .001$; $t_{neg}(99) = 9.35$, $p < .001$. Unsurprisingly, the SBNR people agreed significantly more with the positive spiritual statements, $M = .68$, $SD = .17$, than the atheists, $M = .29$, $SD = .17$. Conversely, atheists agreed significantly more with the negative spiritual statements, $M = .50$, $SD = .22$, than SBNR individuals, $M = .16$, $SD = .14$.

Finally, regarding the *control statements*, the main effect of Evaluation was significant, $F(1, 99) = 393.12$, $\eta_p^2 = .80$, $p < .001$, as well as the interaction between Evaluation and Group, $F(1, 99) = 5.35$, $\eta_p^2 = .05$, $p = .02$. The post-hoc t-tests confirmed a

significant difference between the groups only for the negative control statements, $t(99) = -2.41$, $p = .02$. This indicated that SBNR people agreed significantly more with the negative control statements, $M = .26$, $SD = .12$, than atheists did, $M = .21$, $SD = .09$.

Figure 4

Mean Proportion of Agreement with the Different Statement Categories in Study 1A



Note. Atheists $n = 46$; SBNR $n = 55$. Positive and negative evaluation refer to the participants' explicit agreement with positive and negative statements, respectively. The error bars represent the standard error of the mean. The significance levels (Bonferroni corrected) indicate the statistical significance of the post-hoc repeated measures analysis of variance (top), and the t-tests (bottom). SBNR = spiritual but not religious; n.s. = not significant. * $p < .025$. *** $p < .001$.

Table 3

Repeated Measures ANOVA Results on Differences in Explicit Supernatural Attitudes Between Different Levels of Statement, Evaluation and Group in Study 1A

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Statement	0.11	0.05	2 (198)	12.11	< .001	.11
Evaluation	3.75	3.75	1 (99)	71.71	< .001	.42
Group[a]	0.001	0.001	1 (99)	0.46	.50	.01
Statement x Evaluation[b]	5.41	2.96	1.83 (180.93)	59.45	< .001	.38
Statement x Group	0.04	0.02	2 (198)	4.54	.01	.04
Evaluation x Group	6.21	6.21	1 (99)	118.74	< .001	.55
Statement x Evaluation x Group	4.79	2.39	2 (198)	52.58	< .001	.35

Note. Type III Sum of Squares. Statement had 3 levels: religious, spiritual, and control; Evaluation had 2 levels: positive and negative (referring to the participants' explicit agreement with positive and negative statements, respectively); Group had 2 levels: atheist and SBNR. SBNR = spiritual but not religious.

[a]Between-subject effect.

[b]Huynh-Feldt sphericity correction.

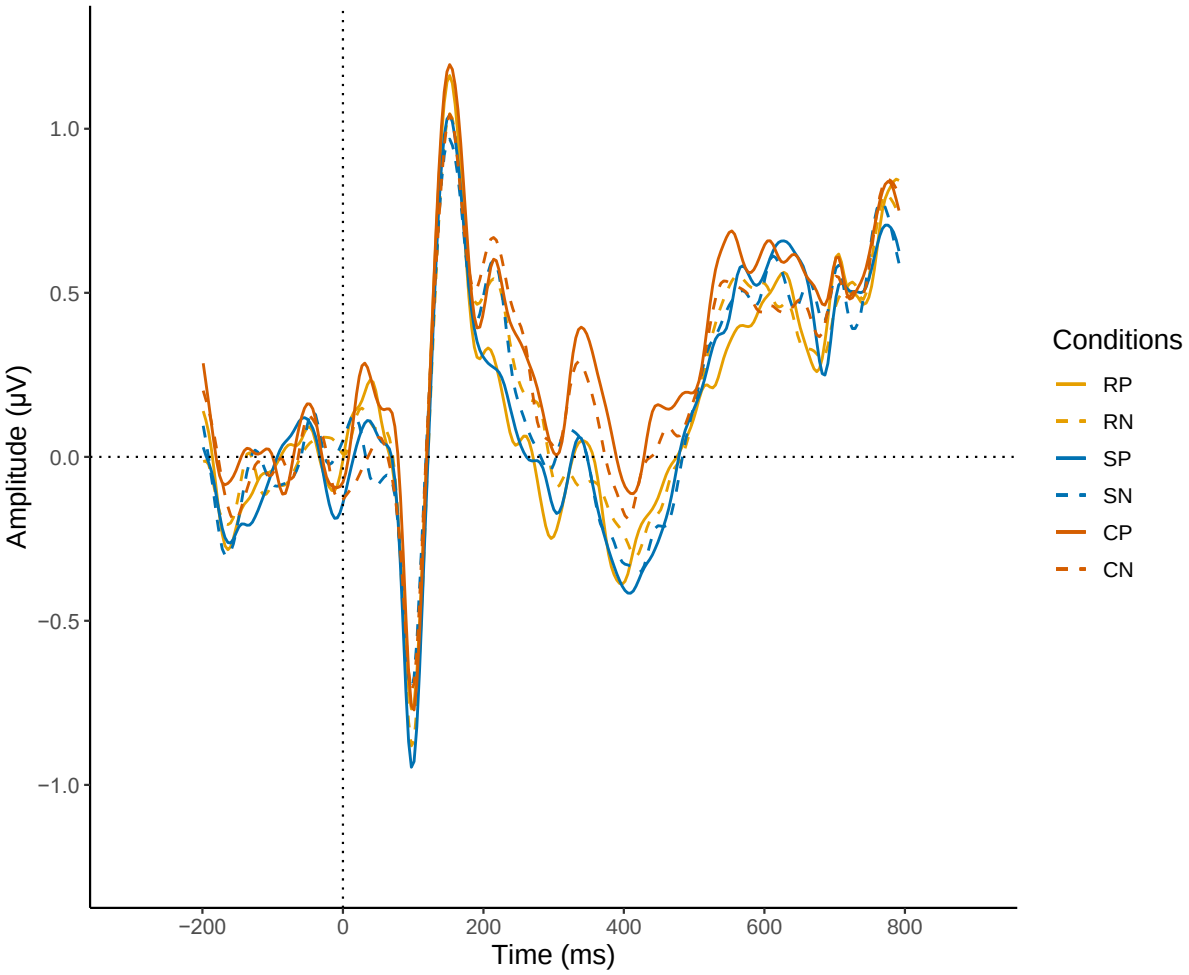
ERP Analysis

In order to examine the N400-effect, the average ERP amplitude (μV) was extracted in a window of 350–450 ms after stimulus onset. Figure 5 depicts the averaged ERPs of each condition, and Figure 6 shows the averaged ERPs per condition for each group. In Figure 7, the mean amplitude per condition is depicted.

A 9 x 3 x 2 x 2 RM ANOVA was conducted with the factors Electrode (reflecting the 3 x 3 electrode cluster used to analyze the N400-component), Statement (religious/spiritual/and control), and Evaluation (positive/negative) entered as within-subject factors, and Group (atheist/SBNR) as the between-subject factor. Regarding the within-subject effects, there was a significant main effect of Electrode, $F(2.10, 207.96) = 18.44$, $\eta_p^2 = .16$, $p < .001$, and a significant main effect of Statement, $F(2, 198) = 16.82$, $\eta_p^2 = .15$, $p < .001$. This indicated that the N400-amplitude of the spiritual statements was significantly more negative, $M = -0.29$, 95% confidence interval (CI) [−0.46, −0.13], than that of the religious ones, $M = -0.22$, [−0.39, −0.06], whereas the N400-amplitude of control statements was relatively positive, $M = 0.01$, [−0.16, 0.17]. Furthermore, there were significant interaction effects of Electrode x Statement, $F(7.24, 717.03) = 6.22$, $\eta_p^2 = .06$, $p < .001$, and Electrode x Evaluation, $F(4.40, 435.74) = 4.60$, $\eta_p^2 = .04$, $p = .001$. Contrary to our hypotheses, no other main effects and interactions reached statistical significance. The between-subjects effect of Group was also not significant, $F(1, 99) = 3.58$, $\eta_p^2 = .04$, $p = .06$. Based on these results, we cannot reject the $H0_A$: We found no statistical difference in the average N400-amplitude between positive and negative evaluations of spiritual statements between atheist and SBNR individuals (see Table 4 for the ANOVA results pertaining to the hypotheses of study 1A).

Figure 5

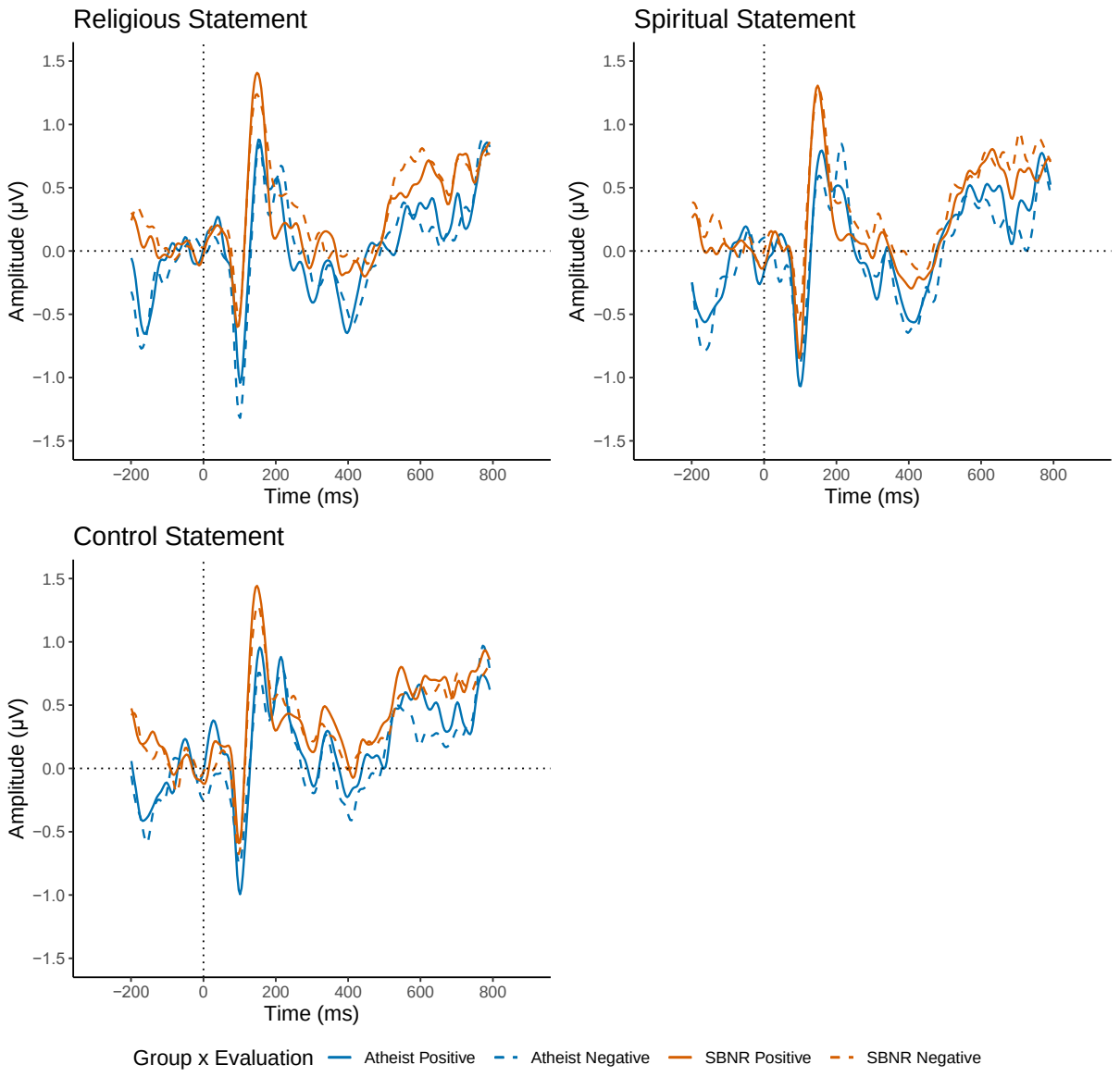
Averaged ERPs Across the Different Conditions of Study 1A



Note. All participants (after exclusions; *N* = 101) were included. ERPs = event-related potentials; RP = religious statement, positive evaluation; RN = religious statement, negative evaluation; SP = spiritual statement, positive evaluation; SN = spiritual statement, negative evaluation; CP = control statement, positive evaluation; CN = control statement, negative evaluation.

Figure 6

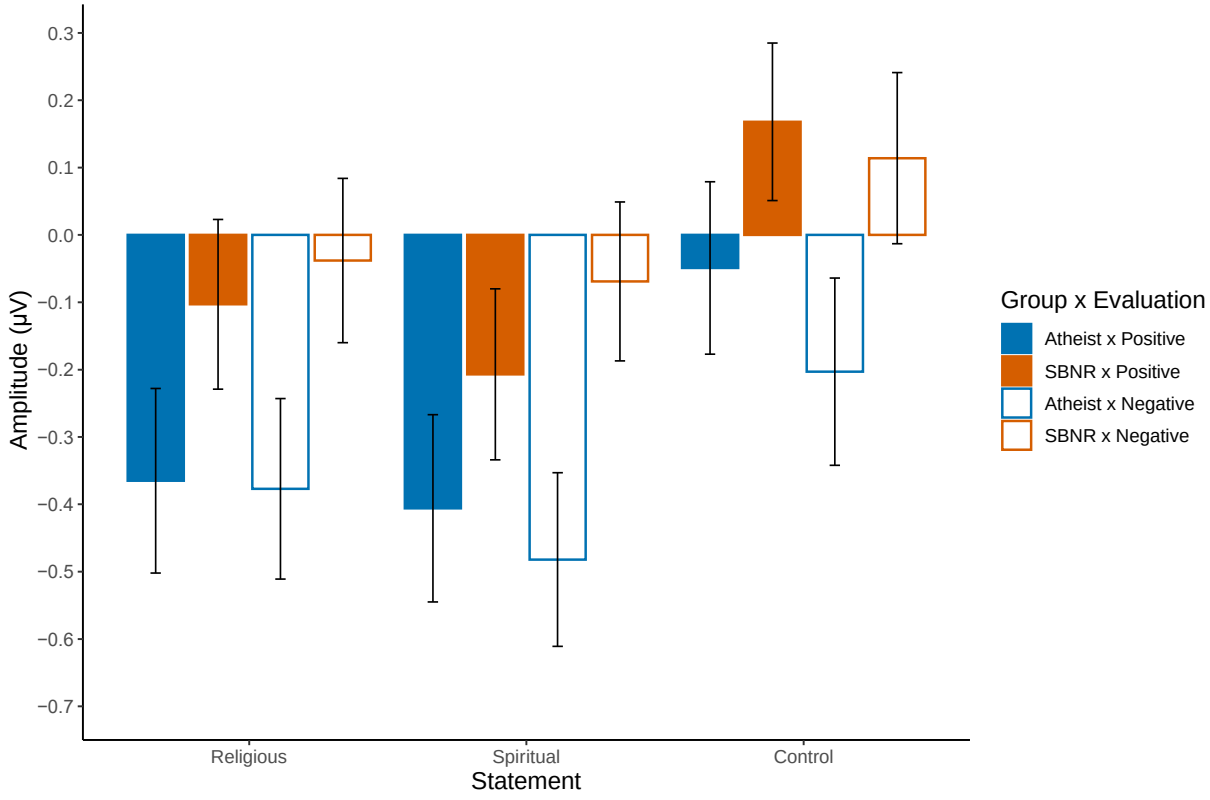
Averaged ERPs per Condition for the Atheist and SBNR Groups



Note. Mean voltage amplitude (μV) as a function of time (ms) for the religious (top left), spiritual (top right), and control (bottom left) statements. The blue lines represent the atheist individuals and the red lines the SBNR; the solid lines represent the positive evaluations, while the dashed lines represent the negative evaluations. ERPs = event-related potentials; SBNR = spiritual but not religious.

Figure 7

Mean N400-Amplitude per Condition in Study 1A



Note. The bars represent the standard error of the mean. SBNR = spiritual but not religious.

Table 4

Repeated Measures ANOVA Results on Differences in the N400-Amplitude Between Different Levels of Statement, Evaluation and Group in Study 1A

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Statement	87.73	43.86	2 (198)	16.82	< .001	.15
Evaluation	0.33	0.33	1 (99)	0.11	.74	.001
Group[a]	2.12	2.12	1 (99)	3.58	.06	.04
Statement x Evaluation	5.30	2.65	2 (198)	1.11	.33	.01
Statement x Group	0.41	0.21	2 (198)	0.08	.92	.001

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Evaluation x Group	5.75	5.75	1 (99)	1.96	.16	.02
Statement x Evaluation x Group	1.21	0.61	2 (198)	0.25	.78	.003

Note. Type III Sum of Squares. Statement had 3 levels: religious, spiritual, and control; Evaluation had 2 levels: positive and negative; Group had 2 levels: atheist and SBNR. SBNR = spiritual but not religious.

[a]Between-subject effect.

Correlation Analysis

In order to investigate the relationship between the N400-effect and the explicit evaluation of supernatural attitudes, a Pearson's correlation analysis was conducted. For this, we used the N400-effect of each condition (i.e., the difference between negative and positive statements for each category), and the proportion of agreement with each condition (i.e., difference of proportion of agreement between negative and positive statements per category). Contrary to expectations, no correlations between the variables of interest were significant at $p = .05$ (see Table 5 for the correlation analysis results).

Table 5

Correlations Between the N400-Effect of each Condition and the Proportion of Agreement with each Condition in Study 1A

	N400-religious	N400-spiritual	N400-control	Religious agreement	Spiritual agreement	Control agreement
N400-religious	—					
N400-spiritual	.09	—				

	N400-religious	N400-spiritual	N400-control	Religious agreement	Spiritual agreement	Control agreement
N400-control	.07	.06	—			
Religious agreement	.08	−.11	−.10	—		
Spiritual agreement	−.12	−.04	−.08	.47**	—	
Control agreement	.14	.04	−.06	.08	−.26**	—

Note. Two-tailed correlations. The N400-variables denote the N400-effect for each statement category (i.e., the difference between negative and positive statements). The agreement variables denote the proportion of agreement with each statement category (i.e., the difference of the proportion of agreement between negative and positive statements).

** $p < .01$.

Study 1B: Supernatural Beliefs

Explicit Beliefs

The explicit data on supernatural beliefs was examined for differences between groups and conditions in a similar manner to the explicit attitude analysis of study 1A. The only difference involved the factor Evaluation, which contained two different, albeit related, levels: belief and disbelief. This factor referred to the proportion of belief and disbelief statements that participants agreed with. Figure 8 represents the mean proportion of agreement with each condition (Statement x Evaluation) for both groups, and Table 6 depicts the results of this RM ANOVA analysis.

The analysis showed that the main effect of Statement was not significant, $F(2, 214) = 1.71$, $\eta_p^2 = .02$, $p = .18$, but, the main effect of Evaluation was, $F(1, 107) = 191.62$, $\eta_p^2 = .64$, $p < .001$. The interaction between Statement and Evaluation was also significant, $F(2, 214) =$

129.04, $\eta_p^2 = .55$, $p < .001$. The between-subject effect of Group was found to not be significant, $F(1, 107) = 1.29$, $\eta_p^2 = .01$, $p = .26$.

Furthermore, we found significant interactions between Statement and Group, $F(2, 214) = 6.77$, $\eta_p^2 = .06$, $p = .001$, as well as between Evaluation and Group, $F(1, 107) = 198.27$, $\eta_p^2 = .65$, $p < .001$. Finally, the critical three-way interaction between Statement, Evaluation, and Group was also found to be significant, $F(2, 214) = 28.82$, $\eta_p^2 = .21$, $p < .001$, meaning that the two groups differed in their agreement with the statements depending on the statements' category and evaluation. In order to clarify the direction of these differences, we ran three post-hoc RM ANOVAs testing the factors Evaluation and Group for the religious, spiritual, and control statements separately (all corrected for multiple comparisons to $\alpha = .025$).

Regarding the *religious statements*, the results showed a significant main effect of Evaluation, $F(1, 107) = 47.83$, $\eta_p^2 = .31$, $p < .001$, and a significant interaction between Evaluation and Group, $F(1, 107) = 85.17$, $\eta_p^2 = .44$, $p < .001$. Post-hoc t-tests confirmed significant differences between the two groups on both the belief and the disbelief religious statements: $t_B(107)$ (i.e., belief) $= -8.72$, $p < .001$; $t_D(107)$ (i.e., disbelief) $= 9.28$, $p < .001$. The results indicated that the SBNR group agreed significantly more with the belief religious statements, $M = .52$, $SD = .31$, than atheists did, $M = .09$, $SD = .19$. Conversely, atheists agreed significantly more with the disbelief religious statements, $M = .87$, $SD = .20$, than the SBNR, $M = .41$, $SD = .29$.

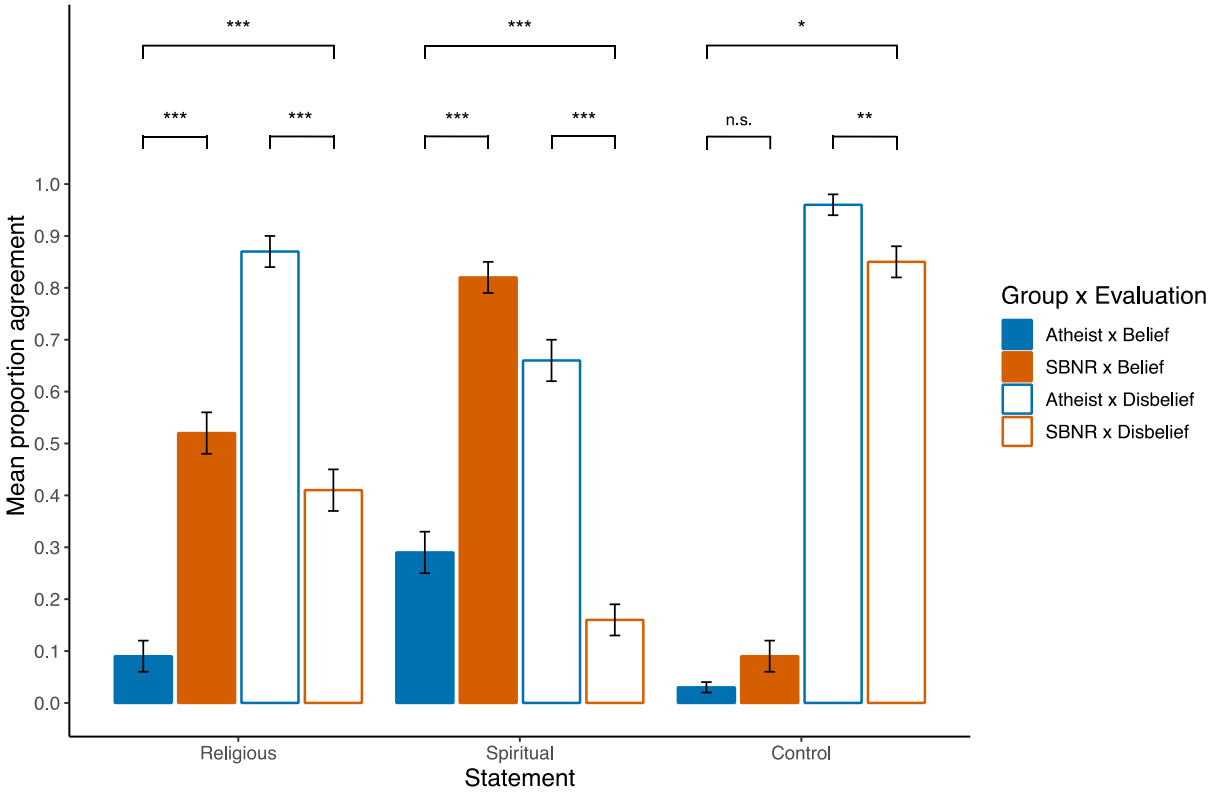
Regarding the *spiritual statements*, we observed a significant main effect of Evaluation, $F(1, 107) = 9.71$, $\eta_p^2 = .08$, $p = .002$, as well as a significant interaction between Evaluation and Group, $F(1, 107) = 127.43$, $\eta_p^2 = .54$, $p < .001$. Post-hoc t-tests confirmed these differences for both levels of the spiritual statements, $t_B(107) = -11.52$, $p < .001$; $t_D(107) = 10.86$, $p < .001$. Unsurprisingly, SBNR individuals agreed significantly more with

the belief spiritual statements, *M* = .82, *SD* = .21, than atheists, *M* = .29, *SD* = .27. Conversely, atheists agreed significantly more with the disbelief spiritual statements, *M* = .66, *SD* = .28, than the SBNR, *M* = .16, *SD* = .20.

Finally, regarding the *control statements*, the main effect of Evaluation was significant, $F(1, 107) = 608.91$, $\eta_p^2 = .85$, $p < .001$, as well as the interaction between Evaluation and Group, $F(1, 107) = 6.15$, $\eta_p^2 = .05$, $p = .02$. The post-hoc t-tests demonstrated a significant difference between the groups only for the disbelief control statements, $t(107) = 2.97$, $p = .004$. This indicated that atheists agreed significantly more with the disbelief control statements, *M* = .96, *SD* = .12, than SBNR individuals, *M* = .85, *SD* = .24.

Figure 8

Mean Proportion of Agreement with the Different Statement Categories in Study 1B



Note. Atheist *n* = 50; SBNR *n* = 59. Belief and disbelief evaluation refers to the participants' explicit agreement with belief and disbelief statements, respectively. The error bars represent

the standard error of the mean. The significance levels (Bonferroni corrected) indicate the statistical significance of the post-hoc repeated measures analysis of variance (top), and the t-tests (bottom). SBNR = spiritual but not religious; n.s. = not significant.

* $p < .025$. ** $p < .01$. *** $p < .001$.

Table 6

Repeated Measures ANOVA Results on Differences in Explicit Supernatural Attitudes Between Different Levels of Statement, Evaluation and Group in Study 1B

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Statement	0.01	0.01	2 (214)	1.71	.18	.02
Evaluation	19.29	19.29	1 (107)	191.62	< .001	.64
Group[a]	0.001	0.001	1 (107)	1.29	.26	.01
Statement x Evaluation	26.37	13.19	2 (214)	129.04	< .001	.55
Statement x Group	0.05	0.03	2 (214)	6.77	.001	.02
Evaluation x Group	19.96	19.96	1 (107)	198.27	< .001	.65
Statement x Evaluation x Group	5.89	2.94	2 (214)	28.82	< .001	.21

Note. Type III Sum of Squares. Statement had 3 levels: religious, spiritual, and control; Evaluation had 2 levels: belief and disbelief (referring to the participants' explicit agreement with belief and disbelief statements, respectively); Group had 2 levels: atheist and SBNR. SBNR = spiritual but not religious.

[a]Between-subject effect.

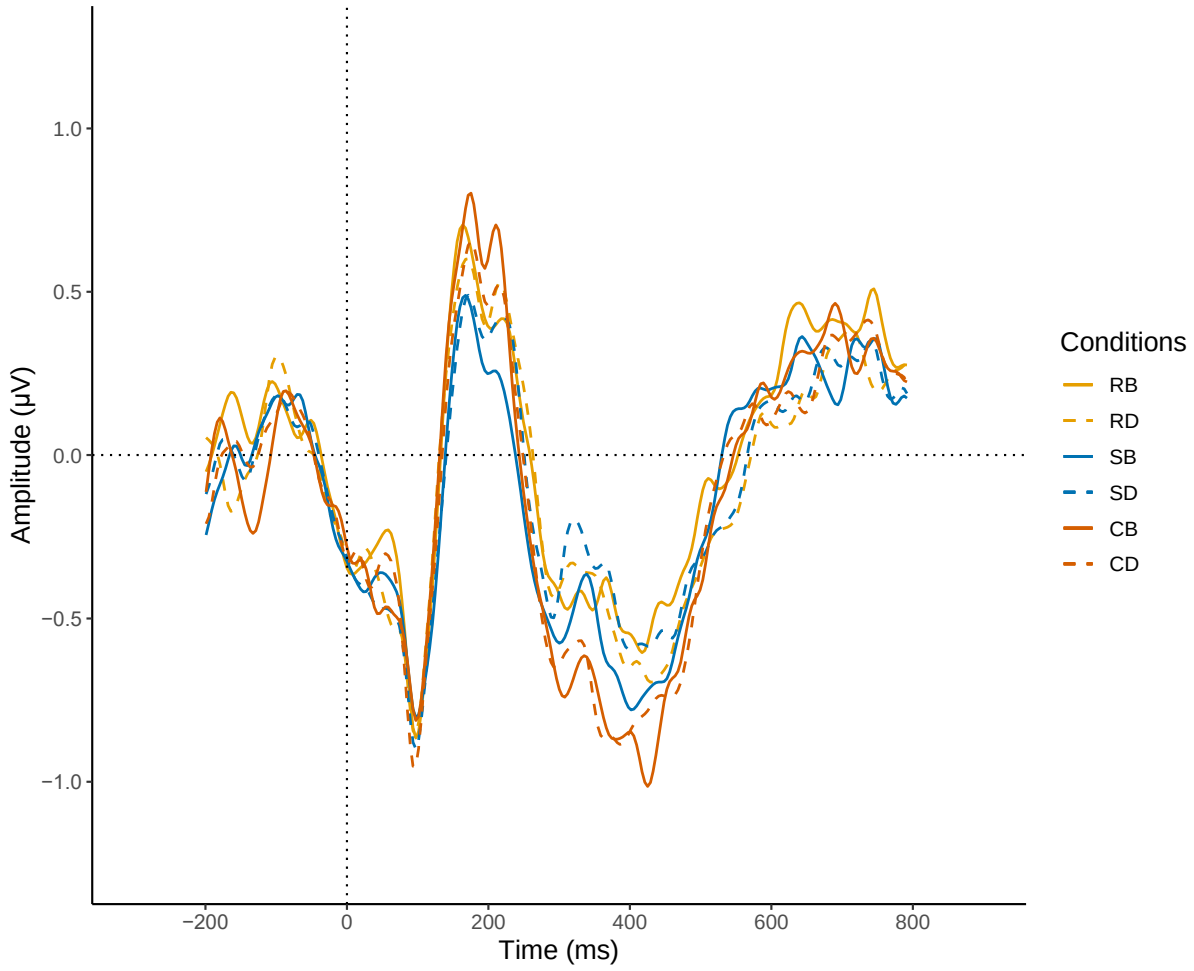
ERP Analysis

For study 1B, the same specifications for data extraction and the same RM ANOVA factors were used as in the ERP analysis of study 1A. The only difference in the ANOVA was that the factor Evaluation was composed of different, yet related, levels: belief (i.e., positive evaluation) vs. disbelief (i.e., negative evaluation). Figure 9 shows the averaged ERPs of each condition, and Figure 10 depicts the averaged ERPs per condition for each group. In Figure 11, the mean amplitude per condition is depicted.

The RM ANOVA results showed significant main effects of Electrode, $F(1.99, 212.90) = 14.83$, $\eta_p^2 = .12$, $p < .001$, and of Statement, $F(2, 214) = 11.62$, $\eta_p^2 = .10$, $p < .001$. The N400-amplitude of the spiritual statements was significantly more negative, $M = -0.61$, 95% CI [−0.79, −0.43], than that of the religious ones, M = −0.55, [−0.71, −0.40]; however, the control statements displayed the largest N400-amplitude of all statement categories, $M = -0.85$, [−1.03, −0.67]. Moreover, there was a significant interaction between Electrode and Statement, $F(6.09, 651.10) = 9.11$, $\eta_p^2 = .08$, $p < .001$, and an interaction between Statement and Evaluation that approached significance, $F(2, 214) = 2.91$, $\eta_p^2 = .03$, $p = .057$. Contrary to expectations, no other main effects and interactions reached statistical significance. The between-subjects effect of Group was also not significant, $F(1, 107) = 2.12$, $\eta_p^2 = .02$, $p = .15$. Based on these results, we cannot reject the $H0_B$: We found no statistical difference in the N400-amplitude between belief and disbelief spiritual statements between atheist and SBNR individuals (see Table 7 for the ANOVA results that pertain to the hypotheses of study 1B).

Figure 9

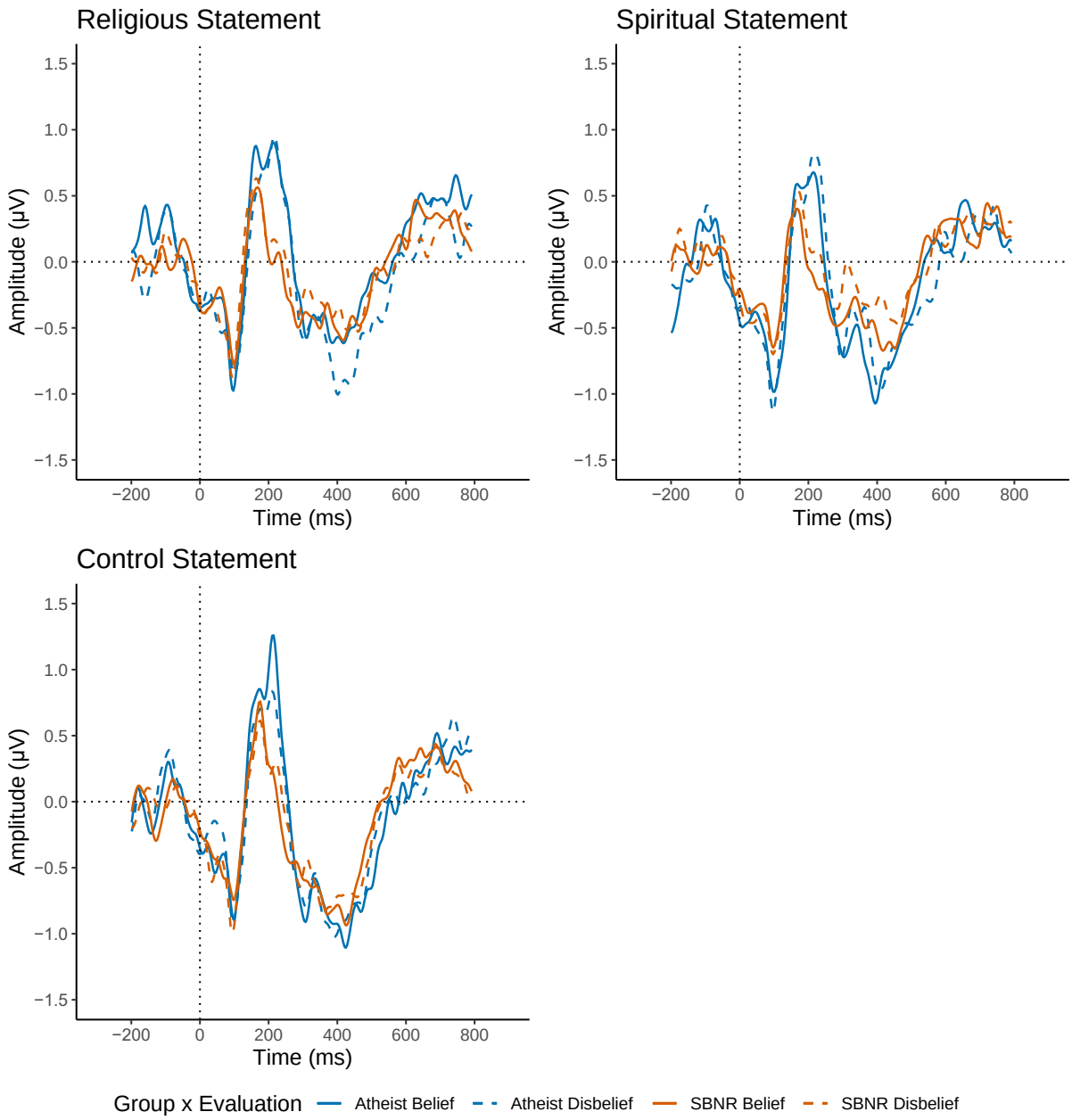
Averaged ERPs Across the Different Conditions of Study 1B



Note. All participants (after exclusions; *N* = 109) were included. ERPs = event-related potentials; RB = religious belief statement; RD = religious disbelief statement; SB = spiritual belief statement; SD = spiritual disbelief statement; CB = control belief statement; CD = control disbelief statement.

Figure 10

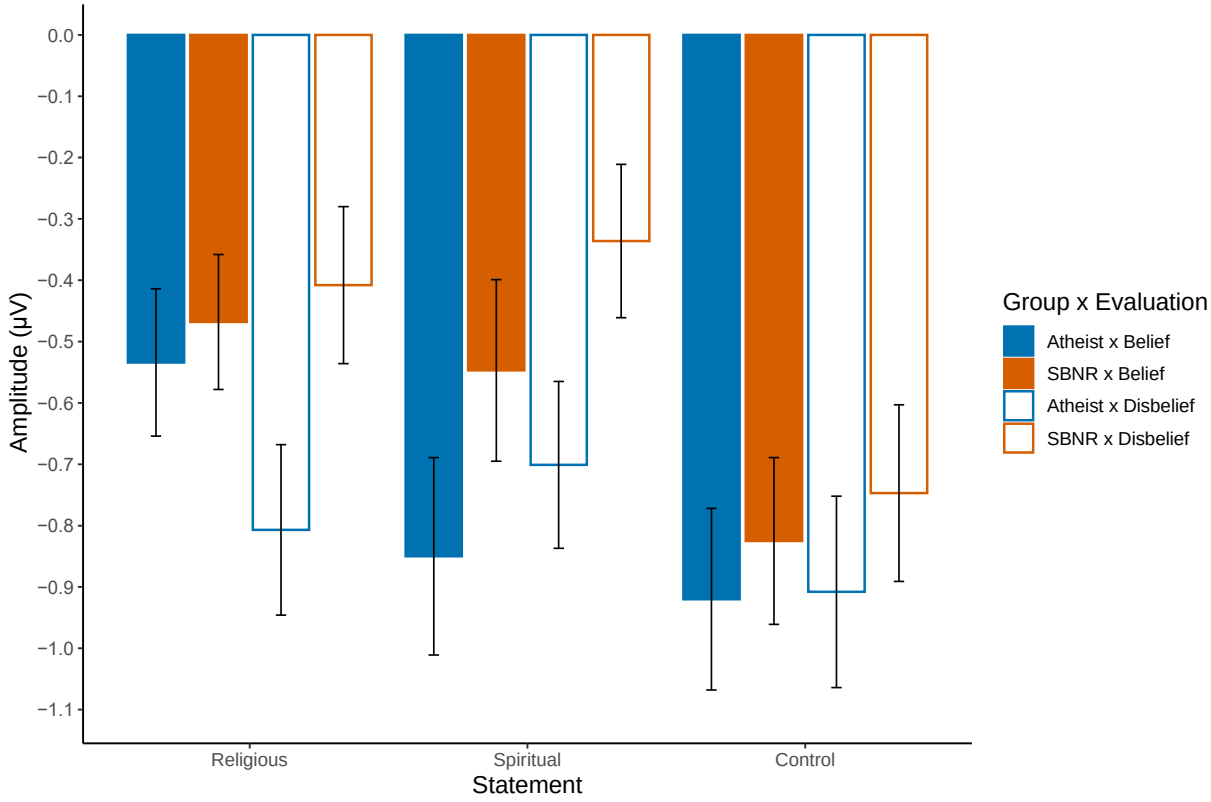
Averaged ERPs per Condition for the Atheist and SBNR Groups



Note. Mean voltage amplitude (μV) as a function of time (ms) for the religious (top left), spiritual (top right), and control (bottom left) statements. The blue lines represent the atheist individuals and the red lines the SBNR; the solid lines represent the belief evaluation (i.e., positively phrased), while the dashed lines represent the disbelief evaluations (i.e., negatively phrased). ERPs = event-related potentials; SBNR = spiritual but not religious.

Figure 11

Mean N400-Amplitude per Condition in Study 1B



Note. The bars represent the standard error of the mean. SBNR = spiritual but not religious.

Table 7

Repeated Measures ANOVA Results on Differences in the N400-Amplitude Between Different Levels of Statement, Evaluation and Group in Study 1B

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Statement	96.66	48.33	2 (214)	11.62	< .001	.10
Evaluation	2.31	2.31	1 (107)	0.59	.44	.01
Group[a]	1.45	1.45	1 (107)	2.12	.15	.02
Statement x Evaluation	19.97	9.98	2 (214)	2.91	.057	.03
Statement x Group	10.27	5.14	2 (214)	1.24	.29	.01

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Evaluation x Group	8.68	8.68	1 (107)	2.23	.14	.02
Statement x Evaluation x Group	5.86	2.93	2 (214)	0.85	.43	.01

Note. Type III Sum of Squares. Statement had 3 levels: religious, spiritual, and control; Evaluation had 2 levels: belief and disbelief; Group had 2 levels: atheist and SBNR. SBNR = spiritual but not religious.

[a]Between-subject effect.

Correlation Analysis

In order to validate that the proportions of agreement with each condition reflect the participants' true beliefs, we run a Pearson's correlation analysis between the EBM and the proportions of agreement with the statements, expressed as a difference between disbelief and belief per statement category (see Table 8 for the results). Expectedly, we observed strong, significant correlations between the explicit belief score and both the religious, $\rho = -.73$, $p < .01$, and the spiritual agreement variables, $\rho = -.80$, $p < .01$. This suggested that the higher the belief score, the more one agreed with the religious and spiritual belief statements, and vice versa. We also found a strong correlation between the religious and spiritual agreement, $\rho = .52$, $p < .01$, which signified that the more one agreed with the religious belief statements, the more they agreed with the spiritual belief statements. The same holds true for the disbelief statements.

Furthermore, we conducted a correlation analysis to examine the relationship between the N400-effect and the explicit evaluations of supernatural beliefs. To this end, we used the N400-effect of each condition (i.e., the difference between disbelief and belief statements per statement category), and the aforementioned proportions of agreement variable. Only the

correlation between religious agreement and the N400-effect for religious statements was shown to be significant, $\rho = -.20$, $p < .05$ (see Figure 12 for a scatterplot depicting this relationship). This reflected that the larger the proportion of agreement with religious *disbelief* statements, the larger the N400-effect appeared to be for the *disbelief* compared to belief statements, and vice versa (see Table 9 for all results).

Table 8

Correlations Between the Belief score of Each Participant and the Proportion of Agreement with each Condition in Study 1B

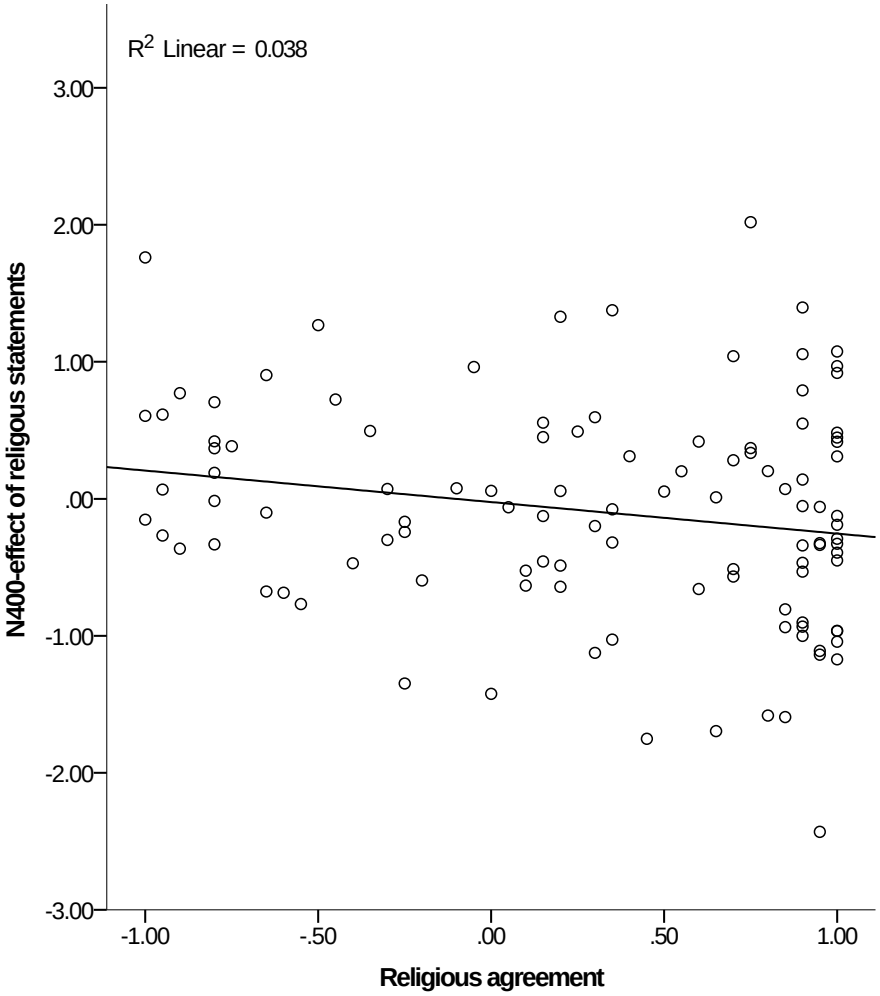
	Belief score	Religious agreement	Spiritual agreement	Control agreement
Belief score	—			
Religious agreement	−.73**	—		
Spiritual agreement	−.80**	.52**	—	
Control agreement	−.17	.11	.15	—

Note. Two-tailed correlations. The belief score is a summed scored of the EBM (Lindeman et al., 2019); the higher the score, the more one believes in supernatural phenomena. The agreement variables denote the proportion of agreement with each statement category (i.e., the difference of the proportion of agreement between disbelief and belief statements).

** $p < .01$.

Figure 12

Scatterplot Depicting the Relationship Between the N400-Effect of and the Proportion of Agreement with Religious Statements



Note. Religious agreement signifies the proportion of agreement with religious statements, expressed as a difference between disbelief and belief statements. The N400-effect of religious statements signifies a difference between the N400-amplitude of disbelief compared to belief statements. The more *positive* the value of the religious agreement variable, the higher the proportion of agreement with the *disbelief* compared to belief statements. The more *negative* the N400-effect value, the larger the effect was for *disbelief* compared to belief statements.

Table 9

Correlations Between the N400-Effect of Each Condition and the Proportion of Agreement with each Condition in Study 1B

	N400-religious	N400-spiritual	N400-control	Religious agreement	Spiritual agreement	Control agreement
N400-religious	—					
N400-spiritual	.06	—				
N400-control	.02	.06	—			
Religious agreement	−.20*	.14	−.02	—		
Spiritual agreement	−.10	−.16	−.11	.52**	—	
Control agreement	−.21*	.11	−.10	.11	.15	—

Note. Two-tailed correlations. The N400-variables denote the N400-effect for each statement category (i.e., the difference between disbelief and belief statements). The agreement variables denote the proportion of agreement with each statement category (i.e., the difference of the proportion of agreement between disbelief and belief statements).

* $p < .05$. ** $p < .01$.

Discussion

The purpose of this study was to examine the evidence for the naturalness of religion hypothesis (NRH), which posits that religion is natural, intuitive and automatic (Barrett, 2000; Bloom, 2007), and whether the N400-component is an effective implicit measure of belief. We did this by elucidating differences in the implicit evaluation of attitude (study 1A) and belief (study 1B) statements by atheist and SBNR individuals. Participants were asked to provide their agreement or disagreement to positive/negative (1A) and belief/disbelief (1B)

religious, spiritual, and control statements while ERPs were being recorded. Participants were presented with statements such as "Being religious is *progressive/obsolete*", "Reading your chakras is *desirable/undesirable*", and "Discussions about sports are *relevant/irrelevant*" (for religious, spiritual, and control, respectively); the N400-effect was measured in response to the final word of the sentences.

Discussion of Results

The analysis of the explicit participant evaluations of both studies showed significant three-way interactions between statement category (religious/spiritual/control), evaluation (positive/negative; belief/disbelief), and group (atheist/SBNR). Specifically, the SBNR agreed more with the positive compared to the negative (1A), and with the belief compared to disbelief supernatural statements (1B), especially the spiritual ones. In contrast, atheists agreed more with the negative compared to the positive (1A), and with the disbelief compared to belief supernatural statements (1B), especially the religious ones.

However, these differences were not apparent in the ERP analysis of the studies. Specifically, SBNR individuals did not display a significant N400-effect for the positive compared to the negative ($H1_A$), or for the belief compared to disbelief spiritual statements ($H1_B$). Additionally, atheists did not display a significant N400-effect for the negative compared to the positive ($H2_A$), or for the disbelief compared to belief spiritual statements ($H2_B$).

Furthermore, the correlation between the N400-effect and the explicit participant evaluations of spiritual attitudes ($H3_A$) and beliefs ($H3_B$) was also not statistically significant. However, in study 1B, we observed a significant, albeit weak, association between the N400-effect and the explicit evaluation of religious statements: The stronger the agreement with religious disbelief statements, the larger the N400-effect for *disbelief* compared to belief evaluations of these statements. This direction was confirmed in a post-hoc analysis where we

examined the belief and disbelief statements separately (see Appendix A for results). Specifically, a larger N400-effect was observed in people who explicitly agreed to a larger extent with the disbelief compared to belief religious statements. This finding is in line with the premise that the N400 is related to the processing of meaningful information related to one's worldviews or values (Kutas & Federmeier, 2011). Therefore, this result could be constituted as a preliminary indication of the N400-effect reflecting supernatural belief violations, in line with previous studies that demonstrated that the N400-effect was related to worldview or value violations (Galli et al., 2017; Hagoort et al., 2004; Van Berkum et al., 2009).

In order to be able to make inferences about group differences, we conducted some further post-hoc correlation analyses (see Appendix A). The aforementioned relationship between the N400-effect and the explicit evaluation of religious statements was apparent only in the SBNR group, indicating a conflict between implicit and explicit beliefs. On the contrary, the more atheists agreed with the religious disbelief statements, the smaller the N400-amplitude was, reflecting a parallel between their implicit and explicit beliefs. Even though these results provide an indication that the N400 could, indeed, reflect an implicit measure of belief, we should interpret these results with caution, as none of the exploratory analyses reached statistical significance, and the effect sizes were fairly small.

We additionally ran a correlation analysis to validate the relationship between the EBM (Lindeman et al., 2019) and the explicit evaluation of the statements, for which we found a strong association. Specifically, the more one believed in supernatural phenomena, the more they agreed with the belief compared to disbelief statements, and vice versa. Additionally, we observed a strong relationship between the explicit agreement with religious and spiritual statements: The more one agreed with the religious belief/disbelief statements, the more they also agreed with the spiritual belief/disbelief statements.

Overall, these results, in combination with the explicit participant evaluations, are in line with previous research that has demonstrated qualitative differences between unbeliever groups in their religious and spiritual attitudes and beliefs (Lindeman et al., 2019, 2020; see also, Lindeman et al., 2016; Lindeman & Lipsanen, 2016). Furthermore, we observed mixed evidence for the NRH (Barrett, 2000). Specifically, there is some indication that religion is, indeed, natural, as we observed a larger N400-amplitude for the disbelief religious statements when participants explicitly agreed with them. In contrast, however, to the expectation that such an effect would be observed for the atheist individuals, this study indicated that it is the SBNR individuals who might implicitly hold *more* religious beliefs than they explicitly state. However, despite the aforementioned relationships, we cannot draw the desired conclusions based on the confirmatory analyses of studies 1A and 1B due to the lack of significant results.

Limitations

There are certain methodological considerations that could account for this lack of significant results. First, we sought to make comparisons between atheist and SBNR individuals. However, the constructed groups did not adequately reflect the intended spiritual differences, even though the current sample size was large for the standards of ERP research (Clayson et al., 2019). The prespecified cut-off criterion for creating the atheist group was too strict for this study, as only seven participants in the whole sample had ≤ 2 in all belief subscales. Such a group would have been too small to make meaningful comparisons, therefore, we utilized the sample mean as the cut-off for group construction. Furthermore, the total sample (i.e., before exclusions) contained, among others, 18 self-proclaimed religious individuals, and 48 individuals who explicitly reported moderate or strong belief in God in the EBM (see Appendix B for the self-reported spiritual identities). In order to keep the analysis in accordance to the preregistration, none of these individuals were excluded from the

analysis. For these reasons, the atheist group inevitably contained—broadly speaking—spiritual individuals and the SBNR group was not strictly areligious.

In order to explore whether our findings could be attributed to these group characteristics, we reallocated participants utilizing strict criteria according to the preregistration and reran the RM ANOVA with the new 'extreme' groups (see Appendix C for the results). The critical three-way interaction between statement, evaluation and group for both studies remained not significant. Although the samples were perhaps too small (N_{1A} = 21; N_{1B} = 22) to make meaningful comparisons, this analysis indicated that the lack of significant findings were perhaps not due to the way the groups were constructed.

A second limitation is related to the presented material. Based on discussions among the research team and revisions of participant feedback (see Appendix D), it became apparent that the choice of stimuli, especially for study 1A, might not have been optimally constructed to test our hypotheses. The large majority of N400-studies deal with lexical and semantic conflicts, to which the N400 is particularly sensitive (Kutas & Federmeier, 2011), such as "I take my coffee with cream and *socks*" (Kutas & Hillyard, 1980). The stimuli used in this study do not reflect such conflicts; in fact, in order to control for lower-level language differences we matched the final word of the statements for various lexical and semantic variables (e.g., word-length and frequency, association with religion and spirituality). This resulted in somewhat unusual expressions and statement structures that potentially did not adequately capture group differences.

However, the statements of study 1B did not have an unusual structure. An alternative explanation for our results could be that for differences in the N400 to become apparent, one might need to use material that is particularly polarized. For example, Van Berkum and colleagues (2009) presented to their participants statements that can be considered fairly outspoken and categorical, such as "A society that condones abortion is a *good/bad* society".

One could suggest that such statements are expected to elicit strong emotions, especially in strongly opinionated people, as was the participant sample the authors recruited. As another example, the study by Galli et al. (2017) included statements with a strong pro- or against-EU orientation, such as "If Britain leaves Europe our quality of life will be *reduced/enhanced*". These authors also tested individuals with a strong predisposition towards these statements, during a time of particular sociopolitical turmoil in the UK. Therefore, it is possible that for differences in the N400 to become apparent, one might need to utilize such strongly-phrased statements, or even perhaps recruit a sample that holds strong opinions on the topic under investigation.

Suggestions for Improvement

In order to be able to investigate the group differences most optimally, one should run the analysis with "true" atheist and SBNR individuals. That is, participants should be allocated in the atheist group when they score ≤ 2 in the subscales of the belief measure, and in the SBNR group when they score > 2; both should strictly not believe in God. This will ensure that the recruited sample has strong opinions about the topics and there are clear boundaries between the groups.

Moreover, the statements should be adapted to better reflect the intended differences. Overall, certain target words were unsuitable for conveying (strong) attitude statements, or did not provide sufficient discriminatory power between groups, such as the pairs "relevant/irrelevant", "correct/incorrect", "imaginable/unimaginable", and "thinkable/unthinkable". Furthermore, certain sentences did not adequately reflect attitudes or worldviews, for instance, "Believing that Mary was a virgin is *imaginable/unimaginable*". Finding something (un)imaginable does not constitute an attitude, value, or worldview about the topic. A lot of people, irrespective of their supernatural attitudes and beliefs, might consider it imaginable that Mary was a virgin; besides, such medical accounts have actually

been reported (Herring et al., 2013). Other sentences might have used inappropriate target words in the context of the statement and failed to capture the desired group differences, such as "Reading your chakras is *desirable/undesirable*". It is conceivable that there are atheists who might find this desirable, and SBNR who might find this undesirable. Finally, previous studies that did find strong N400-effects utilized statements that contained self-referential items (i.e., statements that were directly or indirectly self-relevant; see for example Van Berkum et al., 2009). Self-relevance might decrease ambiguity in interpretation by making the statement readily identifiable and relatable.

Rephrasing the statements as, for example, "Believing that Mary was a virgin makes you *smart/stupid*" and "For a productive society, the reading of chakras is *necessary/unnecessary*" might provide a stronger test of group differences, as they are phrased in a more unambiguous way and convey attitudes/worldviews more clearly. See Table 10 for more suggestions on how to adapt the stimuli, and Appendix E for further elaboration on the rationale of the adaptations.

Table 10

Original and Adapted Stimuli

Original statement	Adapted statement
Believing that God punishes you for your sins is **correct/incorrect**. *There are no inherently correct/incorrect beliefs, which might lead to potential ambiguity in people's understanding of the statement.*	Believing that God punishes you for your sins is **beneficial/harmful**.
Being baptized is **desirable/undesirable**. *Not self-referential.*	For me, being baptized is **desirable/undesirable**.
Being religious is **progressive/obsolete**. T*he target words are not relevant for religious attitudes and beliefs.*	In our society, being religious is **necessary/unnecessary**.

Original statement	Adapted statement
Obeying the ten commandments is **progressive/obsolete**. *Same rationale as above.*	Obeying the 10 commandments is **important/unimportant**.
Praying before eating is **sensical/nonsensical**. *Independent of belief, people can agree/disagree with either statement as anything can be interpreted as sensical.*	For the good of our society, praying before eating is **necessary/unnecessary**.
Worshiping nature must be **recommended/discouraged**. *This is one of the better statements, because it contains the word "must", but it still does not convey a strong attitude. Additionally, "recommended" has a weaker valence than "discouraged".*	Worshipping nature must be **encouraged/discouraged**.
Suggesting that we have an ethereal body is **correct/incorrect**. *There are no inherently correct/incorrect attitudes/beliefs, which might lead to potential ambiguity in people's understanding of the statement.*	For me, believing that we have an ethereal body is **normal/weird**.
Occult phenomena are **harmless/harmful**. *Presupposes that occult phenomena exist and that they are either harmless or harmful. Additionally, "harmless" (neutral) is not an appropriate antonym of "harmful" (negative).*	Believing in the existence of occult phenomena is **beneficial/harmful**.
Reading your horoscope is **relevant/irrelevant**. *Independent of belief, people can agree/disagree with either statement, as anything can be considered relevant.*	I believe that horoscopes are **meaningful/meaningless**.

Note. The target words are marked in bold; the rationale why these statements were adapted is provided in italics. The first five statements refer to religious topics, and the last four to spiritual ones. See Appendix E for further elaboration on the rational of the adaptations.

Finally, a multilevel analysis of the individual statements might be useful in identifying whether there were any statements that did elicit an N400-effect. By including the different statements as individual factors we could ascertain whether they had a direct effect on the N400 rather than solely examining the effect of the condition (i.e., Statement x

Evaluation). If there are any statements that did elicit an N400-effect, we could use them as templates on how to adapt or adjust the rest of the statements. For example, we could adopt the same or similar target pairs, or phrasing of the statements.

Conclusion

In summary, due to the lack of statistically significant relationships, we were not able to conclusively validate the N400 as an implicit measure of (un)belief. Despite this, our study has provided preliminary insights into the study of unbelief; however, the directionality of the results was more mixed than theorized. The effects of secularization (Bruce, 2002) might have provided a social norm that encourages people to respond in a secular way while their true beliefs might be different. Interestingly, however, SBNR individuals seemed to be more spiritual or religious than what they claimed to be, while atheists appeared to be as atheist as they claimed, suggesting inconclusive support for the NRH (Barrett, 2000). For that reason, future research should not only try to improve this study, but should also try to further clarify differences in unbeliever groups. Specifically, as societies grow increasingly secular, it is important to understand the new avenues through which people understand meaning, existence, and their relationships with each other and the world.

References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy white women. *Health Psychology*, *19*(6), 586–592. https://doi.org/10.1037/0278-6133.19.6.586

Banerjee, K., & Bloom, P. (2014). Why did this happen to me? Religious believers' and non-believers' teleological reasoning about life events. *Cognition*, *133*(1), 277–303. https://doi.org/10.1016/j.cognition.2014.06.017

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences*, *4*(1), 29–34. https://doi.org/10.1016/S1364-6613(99)01419-9

Barrett, J. L. (2004). *Why would anyone believe in God?* AltaMira Press.

Barrett, J. L., & Lanman, J. A. (2008). The science of religious beliefs. *Religion*, *38*(2), 109–124. https://doi.org/10.1016/j.religion.2008.01.007

Bek, J., & Lock, S. (2011). Afterlife beliefs: Category specificity and sensitivity to biological priming. *Religion, Brain and Behavior*, *1*(1), 5–17. https://doi.org/10.1080/2153599X.2010.550724

Bering, J. M. (2002). The existential theory of mind. *Review of General Psychology*, *6*(1), 3–24. https://doi.org/10.1037/1089-2680.6.1.3

Bering, J. M. (2006). The folk psychology of souls. *Behavioral and Brain Sciences*, *29*(5), 453–462. https://doi.org/10.1017/S0140525X06009101

Bering, J. M., & Bjorklund, D. F. (2004). The natural emergence of reasoning about the afterlife as a developmental regularity. *Developmental Psychology*, *40*(2), 217–233. https://doi.org/10.1037/0012-1649.40.2.217

Bloom, P. (2005). *Descarte's baby: How the science of child development explains what makes us human*. Basic Books.

Bloom, P. (2007). Religion is natural. *Developmental Science*, *10*(1), 147–151. https://doi.org/10.1111/j.1467-7687.2007.00577.x

Boyer, P. (1994). *The naturalness of religious ideas: A cognitive theory of religion*. University of California Press.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. Basic Books. https://doi.org/10.5860/choice.39-2732a

Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in Cognitive Sciences*, *7*(3), 119–124. https://doi.org/10.1016/S1364-6613(03)00031-7

Bruce, S. (2002). *God is Dead: Secularization in the West*. Blackwell Publishing.

Bulbulia, J. (2004). The cognitive and evolutionary psychology of religion. *Biology and Philosophy*, *19*(5), 655–686. https://doi.org/10.1007/s10539-005-5568-6

Clayson, P. E., Carbine, K. A., Baldwin, S. A., & Larson, M. J. (2019). Methodological reporting behavior, sample sizes, and statistical power in studies of event-related potentials: Barriers to reproducibility and replicability. *Psychophysiology*, *56*(11), e13437. https://doi.org/10.1111/psyp.13437

Cohen, A. B. (2009). Many forms of culture. *American Psychologist*, *64*(3), 194–204. https://doi.org/10.1037/a0015308

Dawkins, R. (2006). *The God delusion*. Houghton Mifflin.

de Jager Meezenbroek, E., Garssen, B., van den Berg, M., Tuytel, G., van Dierendonck, D., Visser, A., & Schaufeli, W. B. (2012). Measuring spirituality as a universal human experience: Development of the spiritual attitude and involvement list (SAIL). *Journal of Psychosocial Oncology*, *30*(2), 141–167. https://doi.org/10.1080/07347332.2011.651258

Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*,

41(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Finley, A. J., Tang, D., & Schmeichel, B. J. (2015). Revisiting the relationship between individual differences in analytic thinking and religious belief: Evidence that measurement order moderates their inverse correlation. *PLoS ONE*, *10*(9), 1–10. https://doi.org/10.1371/journal.pone.0138922

Frühholz, S., Jellinghaus, A., & Herrmann, M. (2011). Time course of implicit processing and explicit processing of emotional faces and emotional words. *Biological Psychology*, *87*(2), 265–274. https://doi.org/10.1016/j.biopsycho.2011.03.008

Galli, G., Sirota, M., Materassi, M., Zaninotto, F., & Terry, P. (2017). Brain indices of disagreement with one's social values predict EU referendum voting behavior. *Social Cognitive and Affective Neuroscience*, *12*(11), 1758–1765. https://doi.org/10.1093/scan/nsx105

Geertz, A. W., & Markússon, G. I. (2010). Religion is natural, atheism is not: On why everybody is both right and wrong. *Religion*, *40*(3), 152–165. https://doi.org/10.1016/j.religion.2009.11.003

Georgiadou, T., & Pnevmatikos, D. (2019). An exploration of afterlife beliefs in religiously- and secularly-oriented adults. *Journal of Beliefs and Values*, *40*(2), 159–171. https://doi.org/10.1080/13617672.2019.1583921

Gervais, W. M. (2013). Perceiving minds and gods: How mind perception enables, constrains, and is triggered by belief in gods. *Perspectives on Psychological Science*, *8*(4), 380–394. https://doi.org/10.1177/1745691613489836

Gervais, W. M., & Norenzayan, A. (2012). Analytic thinking promotes religious disbelief.

Science, *336*(6080), 493–496. https://doi.org/10.1126/science.1215647

Gervais, W. M., Willard, A. K., Norenzayan, A., & Henrich, J. (2011). The cultural transmission of faith: Why innate intuitions are necessary, but insufficient, to explain religious belief. *Religion*, *41*(3), 389–410. https://doi.org/10.1080/0048721X.2011.604510

Gray, K., & Wegner, D. M. (2010). Blaming God for our pain: Human suffering and the divine mind. *Personality and Social Psychology Review*, *14*(1), 7–16. https://doi.org/10.1177/1088868309350299

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*(6), 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Guthrie, S. (1993). *Faces in the clouds: A new theory of religion*. Oxford University Press.

Hagoort, P., Hald, L., Bastiaansen, M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*(5669), 438–441. https://doi.org/10.1126/science.1095455

Henrich, J. (2009). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, *30*(4), 244–260. https://doi.org/10.1016/j.evolhumbehav.2009.03.005

Herring, A. H., Attard, S. M., Gordon-Larsen, P., Joyner, W. H., & Halpern, C. T. (2013). Like a virgin (mother): Analysis of data from a longitudinal, US population representative sample survey. *BMJ*, *347*, f7102. https://doi.org/10.1136/bmj.f7102

Heywood, B. T., & Bering, J. M. (2014). "Meant to be": How religious beliefs and cultural religiosity affect the implicit bias to think teleologically. *Religion, Brain and Behavior*, *4*(3), 183–201. https://doi.org/10.1080/2153599X.2013.782888

Hitzeman, C., & Wastell, C. (2017). Are atheists implicit theists? *Journal of Cognition and*

Culture, *17*(1–2), 27–50. https://doi.org/10.1163/15685373-12342190

Järnefelt, E., Canfield, C. F., & Kelemen, D. (2015). The divided mind of a disbeliever: Intuitive beliefs about nature as purposefully created among different groups of non-religious adults. *Cognition*, *140*, 72–88. https://doi.org/10.1016/j.cognition.2015.02.005

Johnson, D. D. P. (2005). God's punishment and public goods: A test of the supernatural punishment hypothesis in 186 world cultures. *Human Nature*, *16*(4), 410–446. https://doi.org/10.1007/s12110-005-1017-0

Jong, J., Zahl, B. P., & Sharp, C. A. (2017). Indirect and implicit measures of religiosity. In R. Finke & C. D. Bader (Eds.), *Faithful measures: New methods in the measurement of religion* (pp. 78–107). New York University Press. https://doi.org/https://doi.org/10.18574/nyu/9781479875214.003.0004

Kalkman, D. P. (2014). Three cognitive routes to atheism: A dual-process account. *Religion*, *44*(1), 72–83. https://doi.org/10.1080/0048721X.2013.841110

Kapogiannis, D., Barbey, A. K., Su, M., Zamboni, G., Krueger, F., & Grafman, J. (2009). Cognitive and neural foundations of religious belief. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(12), 4876–4881. https://doi.org/10.1073/pnas.0811717106

Kelemen, D. (1999). Function, goals and intention: Children's teleological reasoning about objects. *Trends in Cognitive Sciences*, *3*(12), 461–468. https://doi.org/10.1016/S1364-6613(99)01402-3

Kelemen, D. (2004). Are children "intuitive theists"? Reasoning about purpose and design in nature. *Psychological Science*, *15*(5), 295–301. https://doi.org/10.1111/j.0956-7976.2004.00672.x

Kelemen, D., & Rosset, E. (2009). The human function compunction: Teleological explanation in adults. *Cognition*, *111*(1), 138–143.

https://doi.org/10.1016/j.cognition.2009.01.001

Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, *142*(4), 1074–1083. https://doi.org/10.1037/a0030399

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*(1), 621–647. https://doi.org/10.1146/annurev.psych.093008.131123

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. https://doi.org/10.1126/science.7350657

LaBouff, J. P., Rowatt, W. C., Johnson, M. K., Thedford, M., & Tsang, J. A. (2010). Development and initial validation of an implicit measure of religiousness-spirituality. *Journal for the Scientific Study of Religion*, *49*(3), 439–455. https://doi.org/10.1111/j.1468-5906.2010.01521.x

Lawson, E. T. (2000). Towards a cognitive science of religion. *Numen*, *47*(3), 338–349. https://doi.org/10.1163/156852700511586

Legare, C. H., Evans, E. M., Rosengren, K. S., & Harris, P. L. (2012). The coexistence of natural and supernatural explanations across cultures and development. *Child Development*, *83*(3), 779–793. https://doi.org/10.1111/j.1467-8624.2012.01743.x

Lindeman, M., Heywood, B., Riekki, T., & Makkonen, T. (2014). Atheists become emotionally aroused when daring God to do terrible things. *International Journal for the Psychology of Religion*, *24*(2), 124–132. https://doi.org/10.1080/10508619.2013.771991

Lindeman, M., & Lipsanen, J. (2016). Diverse cognitive profiles of religious believers and nonbelievers. *The International Journal for the Psychology of Religion*, *26*(3), 185–192.

https://doi.org/10.1080/10508619.2015.1091695

Lindeman, M., Marin, P., Schjoedt, U., & van Elk, M. (2020). Nonreligious identity in three Western European countries: A closer look at nonbelievers' self-identifications and attitudes towards religion. *International Journal for the Psychology of Religion*, *30*(4), 288–303. https://doi.org/10.1080/10508619.2020.1746984

Lindeman, M., Svedholm-Häkkinen, A. M., & Lipsanen, J. (2015). Ontological confusions but not mentalizing abilities predict religious belief, paranormal belief, and belief in supernatural purpose. *Cognition*, *134*, 63–76. https://doi.org/10.1016/j.cognition.2014.09.008

Lindeman, M., Svedholm-Häkkinen, A. M., & Riekki, T. (2016). Skepticism: Genuine unbelief or implicit beliefs in the supernatural? *Consciousness and Cognition*, *42*, 216–228. https://doi.org/10.1016/j.concog.2016.03.019

Lindeman, M., Svedholm, A. M., Riekki, T., Raij, T., & Hari, R. (2013). Is it just a brick wall or a sign from the universe? An fMRI study of supernatural believers and skeptics. *Social Cognitive and Affective Neuroscience*, *8*(8), 943–949. https://doi.org/10.1093/scan/nss096

Lindeman, M., van Elk, M., Lipsanen, J., Marin, P., & Schjoedt, U. (2019). Religious unbelief in three western European countries: Identifying and characterizing unbeliever types using latent class analysis. *International Journal for the Psychology of Religion*, *29*(3), 184–203. https://doi.org/10.1080/10508619.2019.1591140

Maij, D. L. R., van Harreveld, F., Gervais, W., Schrag, Y., Mohr, C., & van Elk, M. (2017). Mentalizing skills do not differentiate believers from non-believers, but credibility enhancing displays do. *PLoS ONE*, *12*(8). https://doi.org/10.1371/journal.pone.0182764

Morris, J. P., Squires, N. K., Taber, C. S., & Lodge, M. (2003). Activation of political attitudes: A psychophysiological examination of the hot cognition hypothesis. *Political*

Psychology, *24*(4), 727–745. https://doi.org/10.1046/j.1467-9221.2003.00349.x

Mytko, J. J., & Knight, S. J. (1999). Body, mind and spirit: Towards the integration of religiosity and spirituality in cancer quality of life research. *Psycho-Oncology*, *8*, 439–450. https://doi.org/10.1002/(sici)1099-1611(199909/10)8:5<439::aid-pon421>3.0.co;2-l

Norenzayan, A., & Gervais, W. M. (2013). The origins of religious disbelief. *Trends in Cognitive Sciences*, *17*(1), 20–25. https://doi.org/10.1016/j.tics.2012.11.006

Norenzayan, A., Gervais, W. M., & Trzesniewski, K. H. (2012). Mentalizing deficits constrain belief in a personal God. *PLoS ONE*, *7*(5), e36880. https://doi.org/10.1371/journal.pone.0036880

Norris, P., & Inglehart, R. (2004). *Sacred and secular: Religion and politics worldwide*. Cambridge University Press.

Payne, K., & Lundberg, K. (2014). The affect misattribution procedure: Ten years of evidence on reliability, validity, and mechanisms. *Social and Personality Psychology Compass*, *8*(12), 672–686. https://doi.org/10.1111/spc3.12148

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition*, *123*(3), 335–346. https://doi.org/10.1016/j.cognition.2012.03.003

Pennycook, G., Ross, R. M., Koehler, D. J., & Fugelsang, J. A. (2016). Atheists and agnostics are more reflective than religious believers: Four empirical studies and a meta-analysis. *PLoS ONE*, *11*(4), e0153039. https://doi.org/10.1371/journal.pone.0153039

Pereira, V., Faísca, L., & De Sá-Saraiva, R. (2012). Immortality of the soul as an intuitive idea: Towards a psychological explanation of the origins of afterlife beliefs. *Journal of Cognition and Culture*, *12*(1–2), 101–127. https://doi.org/10.1163/156853712X633956

Pew Research Center. (2018). *Being Christian in western Europe*. Retrieved from: https://www.pewforum.org/2018/05/29/being-christian-in-western-europe/

Razmyar, S., & Reeve, C. L. (2013). Individual differences in religiosity as a function of cognitive ability and cognitive style. *Intelligence*, *41*(5), 667–673. https://doi.org/10.1016/j.intell.2013.09.003

Riekki, T., Lindeman, M., & Lipsanen, J. (2013). Conceptions about the mind-body problem and their relations to afterlife beliefs, paranormal beliefs, religiosity, and ontological confusions. *Advances in Cognitive Psychology*, *9*(3), 112–120. https://doi.org/10.2478/v10053-008-0138-5

Ross, R. M., Brown-Iannuzzi, J. L., Gervais, W. M., Jong, J., Lanman, J. A., McKay, R., & Pennycook, G. (2020). Measuring supernatural belief implicitly using the affect misattribution procedure. *Religion, Brain and Behavior*, *10*(4), 393–406. https://doi.org/10.1080/2153599X.2019.1619620

Sanchez, C., Sundermeier, B., Gray, K., & Calin-Jageman, R. J. (2017). Direct replication of Gervais & Norenzayan (2012): No evidence that analytic thinking decreases religious belief. *PLoS ONE*, *12*(2), e0172636. https://doi.org/10.1371/journal.pone.0172636

Saroglou, V., Delpierre, V., & Dernelle, R. (2004). Values and religiosity: A meta-analysis of studies using Schwartz's model. *Personality and Individual Differences*, *37*(4), 721–734. https://doi.org/10.1016/j.paid.2003.10.005

Sedikides, C., & Gebauer, J. E. (2010). Religiosity as self-enhancement: A meta-analysis of the relation between socially desirable responding and religiosity. *Personality and Social Psychology Review*, *14*(1), 17–36. https://doi.org/10.1177/1088868309351002

Shenhav, A., Rand, D. G., & Greene, J. D. (2012). Divine intuition: Cognitive style influences belief in God. *Journal of Experimental Psychology: General*, *141*(3), 423–428. https://doi.org/10.1037/a0025391

Sosis, R. (2003). Why aren't we all Hutterites? Costly signaling theory and religious behavior. *Human Nature*, *14*(2), 91–127. https://doi.org/10.1007/s12110-003-1000-6

Svedholm, A. M., & Lindeman, M. (2013). The separate roles of the reflective mind and involuntary inhibitory control in gatekeeping paranormal beliefs and the underlying intuitive confusions. *British Journal of Psychology*, *104*(3), 303–319. https://doi.org/10.1111/j.2044-8295.2012.02118.x

Van Berkum, J. J. A., Holleman, B., Nieuwland, M., Otten, M., & Murre, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science*, *20*(9), 1092–1099. https://doi.org/10.1111/j.1467-9280.2009.02411.x

van Elk, M., van Schie, H. T., & Bekkering, H. (2008). Semantics in action: An electrophysiological study on the use of semantic knowledge for action. *Journal of Physiology Paris*, *102*(1–3), 95–100. https://doi.org/10.1016/j.jphysparis.2008.03.011

van Elk, M., van Schie, H. T., & Bekkering, H. (2010). The N400-concreteness effect reflects the retrieval of semantic information during the preparation of meaningful actions. *Biological Psychology*, *85*(1), 134–142. https://doi.org/10.1016/j.biopsycho.2010.06.004

Van Leeuwen, N., & van Elk, M. (2019). Seeking the supernatural: The interactive religious experience model. *Religion, Brain and Behavior*, *9*(3), 221–251. https://doi.org/10.1080/2153599X.2018.1453529

Willard, A. K., & Norenzayan, A. (2013). Cognitive biases explain religious belief, paranormal belief, and belief in life's purpose. *Cognition*, *129*(2), 379–391. https://doi.org/10.1016/j.cognition.2013.07.016

Wilson, D. S. (2002). *Darwin's Cathedral: Evolution, religion, and the nature of society*. The University of Chicago Press.

Yonker, J. E., Edman, L. R. O. ., Cresswell, J., & Barrett, J. L. (2016). Primed analytic thought and religiosity: The importance of individual characteristics. *Psychology of Religion and Spirituality*, *8*(4), 298–308. https://doi.org/10.1037/rel0000095

Zuckerman, P. (2007). Atheism: Contemporary numbers and patterns. In M. Martin (Ed.), *The*

Cambridge Companion to Atheism (pp. 47–66). Cambridge University Press. https://doi.org/10.1017/CCOL0521842700.004

Appendix A

Exploratory Correlation Analyses in Study 1B

In order to clarify what drove the association between the N400-effect of religious statements (i.e., difference in the N400-amplitude between disbelief and belief statements) and the proportion of agreement with these statements (i.e., difference in the proportion of agreement between disbelief and belief statements), we ran correlation analyses separately for the belief and disbelief statements. Results of this analysis can be found in Table A1; Figure A1 depicts the scatterplots of these associations. Specifically, this analysis indicated that there was a stronger relationship between the N400-amplitude of and the agreement with *disbelief* compared to belief religious statements.

Table A1

Correlations Between the N400-Amplitude of and the Explicit Agreement with Belief and Disbelief Religious Statements

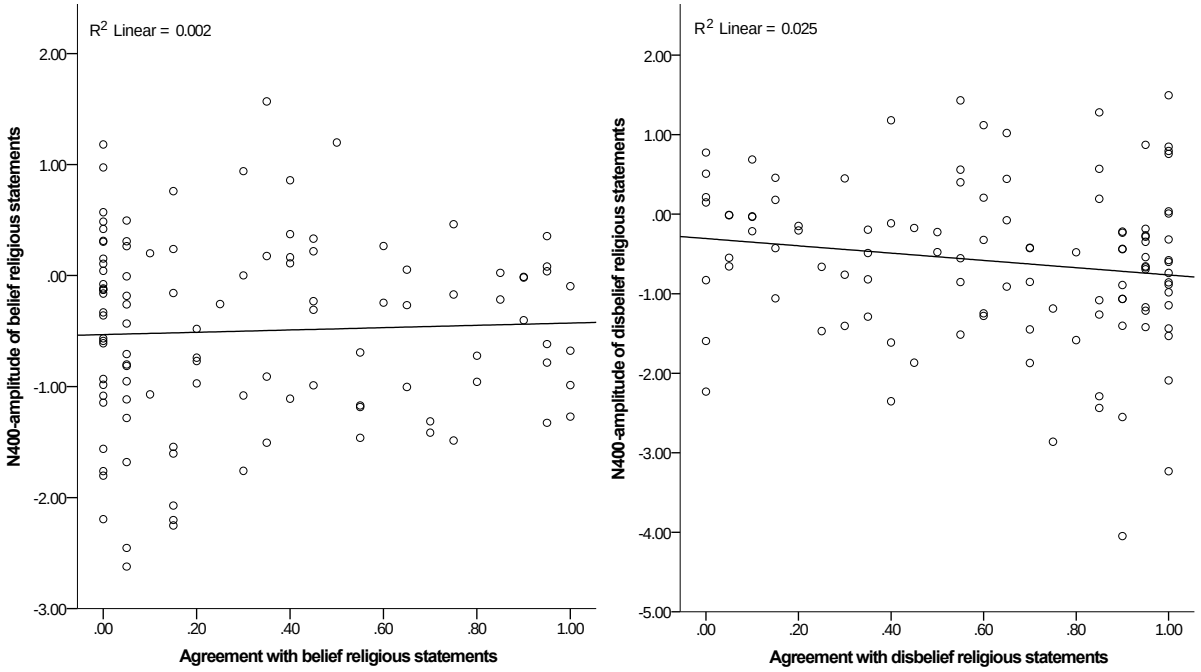
	N400-belief	N400-disbelief	Belief agreement	Disbelief agreement
N400-belief	—			
N400-disbelief	.64**	—		
Belief agreement	**.04**	.21*		
Disbelief agreement	−.03	**−.16**	−.94**	—

Note. Two-tailed correlations. The N400-variables denote the mean N400-amplitude for the belief and disbelief statements. The agreement variables denote the participants' explicit agreement with belief and disbelief statements. The relationships of interest are marked in bold.

* $p < .05$. ** $p < .01$.

Figure A1

Scatterplots Depicting the Relationship Between the N400-Amplitude and Participants' Explicit Agreement with Religious Statements



Note. The relationship between the N400-amplitude and the explicit evaluation of *belief* religious statements is presented on the left, and the relationship between the N400-amplitude and the explicit evaluation of *disbelief* religious statements is presented on the right.

In order to clarify group differences, we ran a correlation analysis examining differences between atheists and SBNR in the relationship between the N400-amplitude and the participant explicit evaluations of belief and disbelief religious statements. Table A2 presents the results of this analysis and Figures A2 and A3 depict the scatterplots of these associations.

Based on this analysis, we can extrapolate that the drive behind the association between a larger N400-amplitude of and a stronger agreement with disbelief religious statements might be the SBNR group. More specifically, it seems that the implicit and explicit beliefs of atheists align, while the implicit and explicit beliefs of SBNR individuals do not

align for the disbelief statements. We should interpret the results of these analysis with caution, as none of the correlations were statistically significant and the effect sizes were fairly small.

Table A2

Correlations Between the N400-Amplitude of and the Explicit Agreement with Belief and Disbelief Religious Statements for Atheist and SBNR Individuals

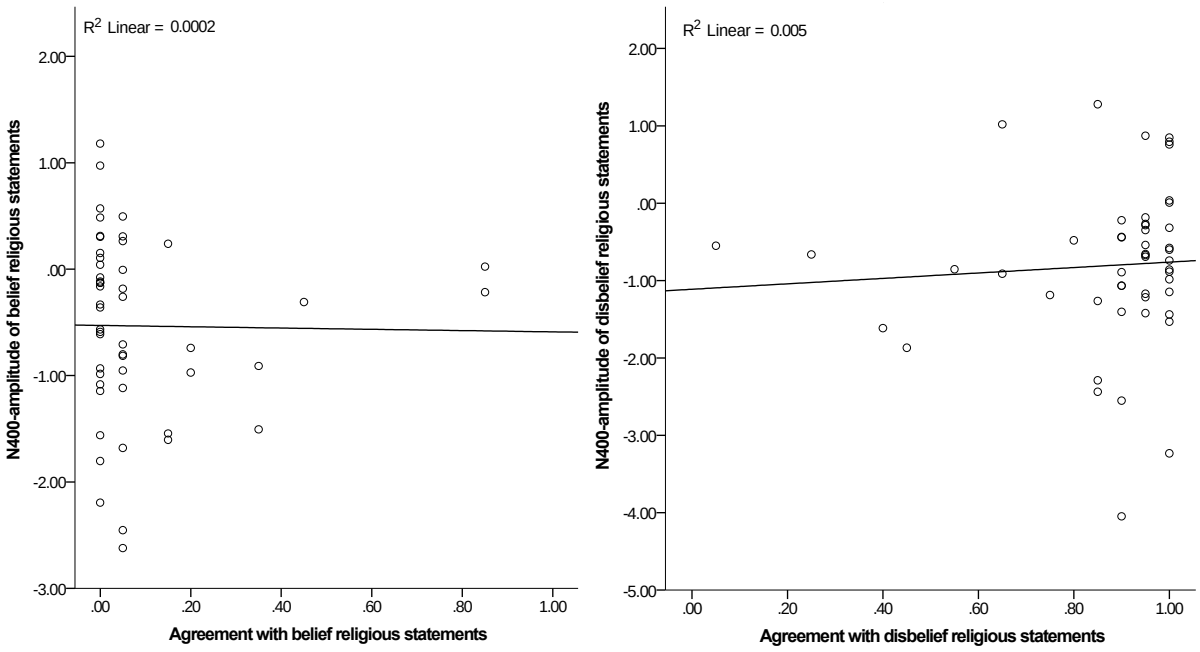
Group		N400-belief	N400-disbelief	Belief agreement	Disbelief agreement
Atheist	N400-belief	—			
	N400-disbelief	.58**	—		
	Belief agreement	**−.01**	.06		
	Disbelief agreement	.05	**.07**	−.86**	—
SBNR	N400-belief	—			
	N400-disbelief	.70**	—		
	Belief agreement	**.04**	.14	—	
	Disbelief agreement	−.04	**−.10**	−.92**	—

Note. Two-tailed correlations. The N400-variables denote the mean N400-amplitude for the belief and disbelief statements. The agreement variables denote the participants' explicit agreement with belief and disbelief statements. The relationships of interest are marked in bold. SBNR = spiritual but not religious.

** $p < .01$.

Figure A2

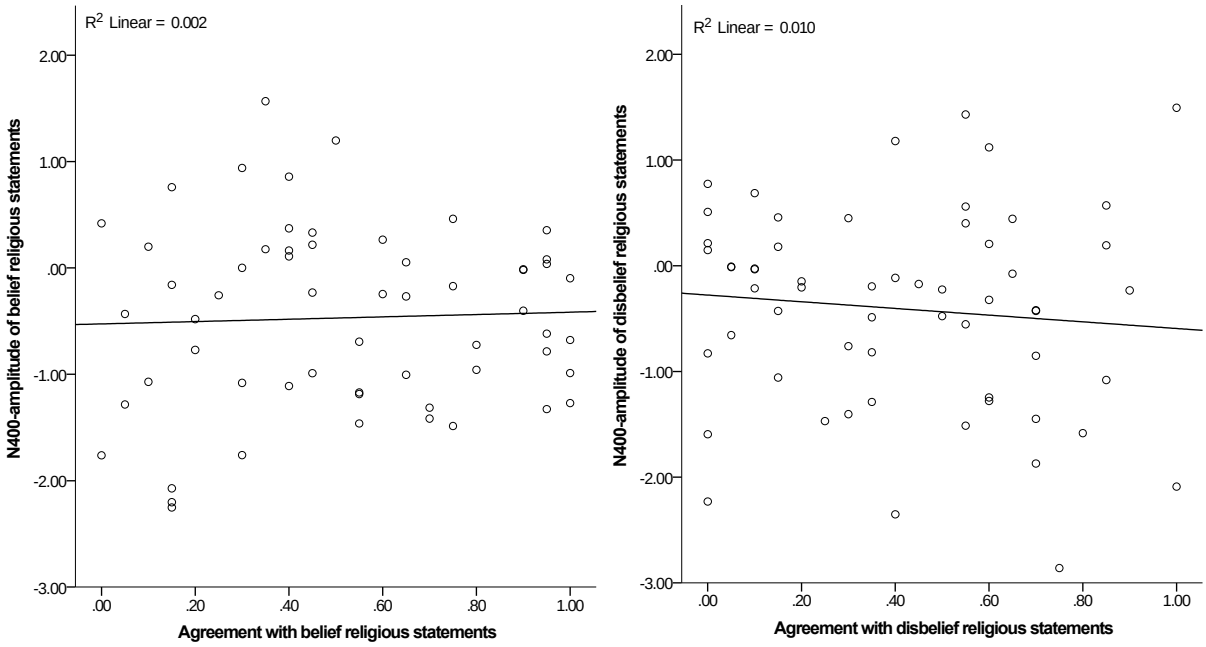
Scatterplots Depicting the Relationship Between the N400-Amplitude of and Atheists' Explicit Agreement with Belief and Disbelief Religious Statements



Note. The relationship between the N400-amplitude and the explicit evaluation of *belief* religious statements by atheist individuals is presented on the left, while that between the N400-amplitude and the explicit evaluation of *disbelief* religious statements is presented on the right.

Figure A3

Scatterplots Depicting the Relationship Between the N400-Amplitude of and SBNR Individuals' Explicit Agreement with Belief and Disbelief Religious Statements



Note. The relationship between the N400-amplitude and the explicit evaluation of *belief* religious statements by SBNR individuals is presented on the left, while that between the N400-amplitude and the explicit evaluation of *disbelief* religious statements is presented on the right.

Appendix B

Self-Reported Spiritual Identities

Table B1 presents the self-reported spiritual identities of participants in the atheist and SBNR groups, and Table B2 presents the spiritual identity descriptions of people who chose the 'other' category. Table B3 depicts the reported agreement with the explicit belief measure question "I believe in God" (Lindeman et al., 2019).

Table B1

Spiritual Identities of Participants in each Group

	Atheist	SBNR
Selected spiritual identity	Frequency (%)	Frequency (%)
Atheist	7 (13)	1 (1.40)
Unbeliever	7 (13)	1 (1.40)
Agnostic	8 (14.80)	3 (4.20)
Non-believer	15 (27.80)	1 (1.40)
Secular	1 (1.90)	—
Religious	1 (1.90)	17 (23.90)
SBNR	7 (13)	34 (47.90)
Spiritual seeker	6 (11.10)	7 (9.90)
Other unbeliever	1 (1.90)	2 (2.80)
Other believer	1 (1.90)	5 (7)
Total	54	71

Note. $N = 125$. The 'other unbeliever' category denotes self-reported spiritual identities that could be considered largely atheist or otherwise lacking religious or spiritual belief; the 'other believer' category denotes self-reported identities that could be considered largely religious.

Table B2

Participant Descriptions of the 'Other' Spiritual Identity

#	Summaries
1	Agnostic, and I believe that the word 'God' can also mean love and energy.
2	Animist.
3	Atheist, because the workings of human behavior do not align, in my opinion, with the concept of 'god'. Maybe agnostic, too. I am against religion as a form of organized

#	Summaries
	power-structure, but I am also respectful towards feelings of piety and I am searching for meaning in/of life.
4	Believer due to being raised in diverse religions and spiritualities, actively spiritual.
5	I am convinced that belief brings certain psychological and social benefits, however it also brings a lot of violence. I don't believe, therefore, in a religion, and I would rather see everybody having the same convictions so that these negative effects of religion disappear. But I would not want to forbid religion because there are still positive effects.
6, 7[a]	Religious and spiritual.
8	Spiritual, and with the heart and soul religious, not with the mind.
9	Spiritual and raised Catholic, but not practicing.

Note. The summaries are translations from Dutch to English. Some participants (e.g., participants 2, 6 and 7) only provided a single or a couple words as their identity; others (e.g., participants 3 and 5) provided a more elaborate explanation. The short descriptions are translated verbatim; the longer ones are summarized in a way that we saw fitting based on the participant's whole description.

[a]There were two participants that reported 'religious and spiritual' as their description of their spiritual identity.

Table B3

Agreement Frequencies of the Question 'I Believe in God'

Agreement	Frequency (%)
Completely disagree	46 (36.80)
Moderately disagree	11 (8.80)
Neutral	20 (16)
Moderately agree	19 (15.20)
Completely agree	29 (23.20)

Note. N = 125. This question is a subscale of the explicit belief measure (Lindeman et al., 2019).

Appendix C

ERP Analysis of 'Extreme' Groups

Due to the fact that the created groups did not optimally reflect the desired spiritual differences (i.e., atheists were not 'pure' atheists, and SBNR were not entirely areligious), we decided to run the ERP analysis of studies 1A and 1B creating the groups as initially conceptualized in a previous study (Lindeman et al., 2019) and in the preregistration. To that end, atheists were defined as scoring ≤ 2 in all belief subscales, and SBNR were defined as scoring > 2 in the belief scales (Lindeman et al., 2019). Furthermore, we excluded all participants that self-identified as religious, participants that chose the 'other' category, and participants who scored more than 1 in the explicit belief measure question "I believe in God" (1 = *completely disagree*; 5 = *completely agree*). Specifically for the creation of the SBNR group, all participants that self-identified as anything but SBNR or spiritual seekers were additionally excluded in order to safeguard against any potential self-identified unbelievers who might have scored > 2 in the belief subscales. Therefore, based on these criteria, in the whole sample (N = 125) there were 7 atheists and 16 SBNR individuals.

Study 1A – Supernatural Attitudes

Based on new group construction criteria, Study 1A had a total of 21 participants (atheist n = 7; SBNR n = 14). We conducted a RM ANOVA with the same factors as discussed in the Results section of study 1A. The critical three-way interaction between Statement, Evaluation and Group was not significant (see Table C1 for all relevant results of this analysis). Based on these results we can infer that the lack of statistically significant results found in study 1A were probably not driven by the way the groups were constructed.

Table C1

Repeated Measures ANOVA Results of Study 1A on Differences in the N400-Amplitude Between Different Levels of Statement and Evaluation Using the 'Extreme' Groups

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Electrode[a]	55.89	26.98	2.07 (39.35)	4.21	.02	.18
Electrode x Evaluation[a]	5.46	1.47	3.71 (70.48)	3.25	.02	.15
Statement	33.61	16.80	2 (38)	5.40	.01	.22
Evaluation	5.62	5.62	1 (19)	2.51	.13	.42
Group[b]	2.21	2.21	1 (19)	0.08	.78	.004
Statement x Evaluation	1.33	0.66	2 (38)	0.24	.79	.01
Statement x Group	0.81	0.41	2 (38)	0.13	.88	.01
Evaluation x Group	2.31	2.31	1 (19)	1.03	.32	.05
Statement x Evaluation x Group	1.40	0.70	2 (38)	0.25	.78	.01

Note. $N = 21$. Type III Sum of Squares. Statement had 3 levels: religious, spiritual, and control; Evaluation had 2 levels: positive and negative (referring to the participants' explicit agreement with positive and negative statements, respectively); Group had 2 levels: atheist and SBNR.

[a]Greenhouse-Geisser correction.

[b]Between-subject effect.

Study 1B – Supernatural Beliefs

Based on the new group construction criteria, study 1B had 22 participants (atheist $n = 7$; SBNR $n = 15$). We conducted an RM ANOVA with the same factors as discussed in the

Results section of study 1B. The critical three-way interaction between Statement, Evaluation and Group was not significant (see Table C2 for all relevant results of this analysis). Based on these results we can infer that the lack of statistically significant results found in this study were probably not driven by the way the groups were constructed.

Table C2

Repeated Measures ANOVA Results of Study 1B on Differences in the N400-Amplitude Between Different Levels of Statement and Evaluation Using the 'Extreme' Groups

Factor	Sum of Squares	Mean Square	*df* (error)	*F*	*p*	η_p^2
Electrode[a]	81.39	37.83	2.15 (43.03)	5.28	.01	.21
Statement	4.54	2.27	2 (40)	0.38	.69	.02
Evaluation	0.85	0.85	1 (20)	0.21	.65	.01
Group[b]	21.96	21.96	1 (20)	0.70	.41	.03
Statement x Evaluation[c]	0.96	0.58	1.67 (33.32)	0.18	.80	.01
Statement x Group	22	11	2 (40)	1.85	.17	.09
Evaluation x Group	0.15	0.15	1 (20)	0.04	.85	.002
Statement x Evaluation x Group	7.24	3.62	2 (40)	1.33	.28	.06

Note. $N = 21$. Type III Sum of Squares. Statement had 3 levels: religious, spiritual, and control; Evaluation had 2 levels: positive and negative (referring to the participants' explicit agreement with positive and negative statements, respectively); Group had 2 levels: atheist and SBNR.

[a]Greenhouse-Geisser correction.

[b]Between-subject effect.

[c]Huynh-Feldt correction.

Appendix D

Participant Feedback

The following table presents participant feedback related to the content or structure of the statements in study 1A. The participants were given the opportunity to give feedback after they completed study 1A, but not after study 1B; therefore, we do not have any feedback that is specifically related to study 1B.

Table D1

Participant feedback related to the content of the statements in study 1A

#	Participant feedback
1	Het was vaak moeilijk om een keuze te maken, of ik voor of tegen de aangegeven kwalificatie, onzinnig/zinnig, correct/incorrect, van een bepaalde gedachte of religieuze opvatting ben, omdat het daarbij afhangt van de vraag, of je meent, dat een gedachte objectief juist of onjuist is, of dat een bepaald geloof een zinvolle gevoelswaarde kan hebben voor (andere) personen, zoals troost biedende rituelen in een andere cultuur, maar daartegenover staat weer het sociale gevaar van heksenvervolging. Bidden en mediteren kan rust geven en persoonlijke zingeving in het leven, al geloof ik niet, dat God de gebeden verhoort. Je gehoord te voelen kan echter toch steun geven.
2	Er kwamen een aantal vragen over eutherisch lichaam of iets en ik had geen idee wat dat precies inhield. Ik denk dat als ik dit aan mensen in mijn omgeving zou vragen, zij het ook niet zouden weten. Misschien een ander woord kiezen?
3	Tijdens de EEG taak vond ik de stelling zeer ambigu. Bij stellingen met 'het is denkbaar' koos ik altijd voor eens, omdat alles denkbaar is, maar niet alles is aannemelijk. Ook stellingen zoals 'bidden voor het eten is zinvol' vond ik onlogisch, want waarvoor bedoelt u dat het zinvol is? Voor de spijsvertering is bidden voor het eten namelijk zinvol omdat aandacht naar het eten gaat en de speeksel productie al opgang komt. Maar bidden voor het eten als een verplichting voor het geloof tegen de zin van het individu in is niet zinvol omdat dat onnodig negatieve gevoelens opwekt. Ik kreeg de indruk dat u met de stellingen de mening van participanten over religieuze overtuigingen en handelingen maar ook niet religiueze onderwerpen wilde meten, maar door de binaire keuze mogelijkheden denk ik dat u de plank mis slaat.
4	Bij een aantal stellingen had ik geen mening. Daarom heb ik bij de stelling als deze voor de 2e keer gesteld was het andere antwoord gegeven.
5	Misschien iets duidelijker het woord 'God' definiëren. Is het de Christelijke God, of meer het goddelijke(energie) of eigen interpretatie?
6	De vragen zijn erg absoluut gesteld.

#	Participant feedback
7	Sommige van de stellingen zijn lastig te beantwoorden. Bijv een stelling als: Geloven dat Jezus kon lopen op water is slecht. Het feit dat je het gelooft is niet slecht, maar het is naar mijn idee niet waar. Daardoor wordt het antwoorden op sommige stellingen wat lastig.
8	Sommige vragen hebben geen zwart/wit antwoord en hangen van de definitie van bijvoorbeeld 'God' af. Of er met 'horoscopen' een paar regels achter in de krant bedoeld worden of een radix chart door een competent astroloog gemaakt. De een is onzinnig, de andere niet.
9	Bij de test evt. meer keuzes zoals bv 'neutraal'.
10	Een begrip als 'lot' wordt door u wellicht anders begrepen als door mij. Als gelovige noem ik wat u lot (destiny) noemt: 'voorzienigheid' (providence).
11	Taalgebruik iets meer definieren.
12	Zelf heb ik moeite met de benaming God. Is voor mij verwarrend merkte ik tijdens de test. Ik gebruik liever De Bron. God is voor mij religieus en bijbels. Daar kan ik mij niet zo in vinden.
13	Ik geloof dat wij allen god zijn. Allemaal een beetje en samen zijn wij een. Iedereen moet mogen geloven wat hij of zij wilt. Bidden voor het eten is goed omdat het belangrijk is om dankbaar te zijn, ook is bidden in het algemeen een vorm van mediteren en dus zinvol. De verhalen in de bijbel zijn symbolische verhalen die te letterlijk zijn genomen. De zoon van god bestaat niet want iedereen is god. In de bijbel wordt er over de zoon van god gesproken, hiermee wordt de zon bedoeld. Spirituele energie bestaat niet, paranormale energie wel. Spiritueel betekend zelf bewust.
14	Ook de vragen zijn voor meerdere intepretaties opvatbaar, wat wordt bedoelt met bepaalde kernbegrippen waardes. Ik bedoel dat als ik het ergens mee eens of oneens ben, wil dat niet zeggen dat ik dat altijd zo vind, maar wel in bepaalde omstandigheden.
15	De vragen tijdens de eerste sessie waren gebaseeerd op het Christendom, volgens mij omdat ik dat had ingevuld maar ik heb daar alleen een achtergrond in en ben verder spirutueel georienteerd. Sommige termen begreep ik niet of vond ik niet geheel van toepassing. Allicht relevant voor bias.

Note. All relevant participant feedback is included verbatim; any feedback or comments unrelated to the statements of study 1A were deleted for brevity and clarity. From the 125 participants, 41 left some feedback, of which 22 simply stated that they have no comments/feedback.

Appendix E

Stimuli Recommendations

In order to create stimuli that will optimally capture the intended group differences, one has to consider three important elements. First, the statements should be phrased in such a way that we would expect the atheists to agree more with the negative rather than the positive evaluations (e.g., "Telepathy must be *ignored*"), while we expect the SBNR individuals to agree more with the positive rather than the negative evaluations (e.g., "Telepathy must be *recognized*"). Statements or target words that could allow for multiple interpretations *irrespective of supernatural attitudes/beliefs* should either get adapted or removed from the stimulus set, such as the statement "Praying before eating is *sensical/nonsensical*", as one might find sense in praying before eating irrespective of their supernatural beliefs. Moreover, target pairs such as "recommended/discouraged" (does not convey a strong distinction), "thinkable/unthinkable" (anything could be thinkable), "imaginable/unimaginable" (anything could be imaginable), "correct/incorrect" (attitudes and beliefs are not inherently correct or incorrect as facts are), and "relevant/irrelevant" (anything could be considered relevant) should also be deleted or adapted. For example, describing something as relevant or irrelevant is unrelated to belief, as is the case in the statement "Discussions about religious topics are *relevant/irrelevant*". Here, one could argue that as long as religious people exist, discussions about religious topics will always be relevant, even if only on a philosophical level. Furthermore, the chosen target pairs should be appropriate opposites of each other in order to convey stronger attitudes and subsequently elicit stronger responses. For example, the antonym of "harmful" (negative valence) is not "harmless" (neutral valence); instead, it might be preferable to use the word "beneficial" (positive valence). Similarly, using "recommended" has a weaker valence than "encouraged" when used as the antonym of "discouraged".

Second, based on previous studies that have studied the N400-effect utilizing worldview or value statements (e.g., Galli et al., 2017; Van Berkum et al., 2009), it might be preferable that the statements reflect a direct (i.e., I think/believe; for me; in my opinion) or indirect (i.e., for our society) relevance to the person. This has the potential of increasing the strength of the conveyed attitude, consequently enhancing any potential worldview or value violations. Therefore, adding such direct or indirect self-references in the statements might make them more personal and readily identifiable, such as "For me, being baptized is *desirable/undesirable*" (the original statement presents a vaguer attitude). Besides, belief in supernatural phenomena (or the lack thereof) is a highly personal matter, which might require a personal perspective in order to be better understood.

Third, the factors that elicit the N400 should be carefully considered, such as valence and word frequency. For instance, some statements might have elicited an N400-effect with either target word, due to for example, unpredictability (Kutas & Federmeier, 2011). The statement "Being religious is *progressive/obsolete*" might have elicited a semantic surprise, as "progressive" and "obsolete" might not reflect how people readily conceptualize religiosity, irrespective of their beliefs. Therefore, such words might have globally been perceived as unexpected. Furthermore, religion is typically associated with conservatism (Saroglou et al., 2004), and as such even religious people might disagree with the word "progressive" seen in such a context. Our suggested recommendations (see Table 10) take into account these factors by incorporating words that are frequently used and associated with the topic under investigation, as well as embody the intended valence in order to capture the intended group differences.

Therefore, we suggest that we should adapt the stimuli in such a way to unambiguously convey a strong attitude for which specific group differences are expected. For that purpose, appropriate antonyms should be used that are relevant and related to

supernatural phenomena, while taking into account word frequency and valence, so that we do not artificially elicit N400-effects.

Finally, there are certain sentences that seem to have adequately captured the aforementioned elements in order to identify the hypothesized group differences. For example, statements such as "Religious authorities must be *acknowledged/ignored*", and "Alternative medicine must be *allowed/forbidden*" utilized appropriate antonyms, clearly conveyed strong attitudes (by utilizing the word "must") that will elicit group differences, and expressed a personal relevance (i.e., something that affects society will have an effect on us, too). In order to ascertain the sentences that might have produced an N400-effect, we could, in a follow-up study, run a multilevel analysis in order to see the effect of the specific sentences on the N400 rather than the effect of the condition (Statement x Evaluation).