UNIVERSITY OF CALIFORNIA

Santa Barbara

Core intuitions about persons co-exist and interfere with acquired Christian beliefs about God

A thesis submitted in partial satisfaction of the

requirements for the degree Master of Arts

in Psychology

by

Michael Barlev

Committee in charge:

Professor Tamsin German, Chair

Professor Leda Cosmides

Professor Ann Taves

December 2015

The thesis of Michael Barlev is approved.

_____

Leda Cosmides

_____

Ann Taves

_____

Tamsin German, Committee Chair

Core intuitions about persons co-exist and interfere with acquired Christian beliefs about God

ABSTRACT

Core intuitions about persons co-exist and interfere with acquired Christian beliefs about God

By

Michael Barlev

In three experiments, using a novel sentence verification paradigm, we tested the hypothesis that acquired Christian beliefs about God which are inconsistent with core intuitions about persons co-exist with, rather than replace, those intuitions in the minds of religious believers. Participants were asked to evaluate a series of statements for which core intuitions and acquired religious beliefs were consistent (i.e. true according to both [e.g. "God has beliefs that are true"] or false according to both [e.g. "all beliefs God has are false"]) or inconsistent (i.e. true on intuition but false theologically [e.g. "God has beliefs that are false"] or false on intuition but true theologically [e.g. "all beliefs God has are true"]). Participants (1) were less accurate and took longer to respond to the inconsistent statements, suggesting that core intuitions both co-exist alongside and interfere with acquired religious beliefs (Experiments 1 and 2), (2) were disproportionately more likely to make errors on the inconsistent statements when responding under time pressure than when responding with no time pressure, suggesting that the resolution of conflicts between inconsistent co-existing beliefs requires cognitive resources (Experiment 2). Experiment 3 ruled-out a plausible alternative interpretation of these results.

1 Introduction

The tendency to attribute supernatural entities (e.g. gods, spirits, ancestor spirits, and divine beings) with person-like characteristics is widespread among present and past human cultures; indeed, it is noted in writings dating as far back as ancient Greece (Boyer, 1994a, b; 2001). However, it was only with relatively recent theoretical advances in evolutionary psychology and cognitive science that this tendency could be explained via the evolved, universal information processing architecture of the human mind: supernatural entities are attributed with person-like characteristics because just like representations of persons they are formed by co-opting the evolved person concept (also referred to as a "person template"; e.g. Boyer, 2001; Boyer & Ramble, 2001). The person concept consists of default inferences about persons, such as about their physicality, biology, and psychology which reliably develop from a skeletal set of inferences about persons present in infancy and from associated learning adaptations (e.g. Carey, 2009; Baillargeon, Scott, & Bian, 2015 for recent reviews).

However, supernatural entities are also believed to have characteristics inconsistent with default inferences about persons. The exact characteristics depend on the supernatural entity and the theological tradition; for example, while persons are intuitively believed to have limited perceptual and mental abilities, in all mainstream Christian denominations God is believed to be omniscient (Boyer, 1994a,b, 2001; Boyer & Ramble, 2001). In this study we investigate the hypothesis by Sperber and colleagues (Sperber, 1985, 1996, 1997, 2000; Mercier & Sperber, 2009; also see Boyer 1994a, b, 2001; Barrett & Keil, 1996; Barrett, 1998, 1999) according to which characteristics attributed to supernatural entities which are

inconsistent with default inferences about persons (1) do not replace these inferences, but (2) co-exist with them in the minds of religious believers.[1]

1.1 Previous Research

A variety of studies utilizing self-report methodologies have explored the psychological, physiological, and physical attributions children and adults make about God, concluding that both are quite willing to attribute certain human characteristics, specifically mental states, to God (e.g. Gray, Gray & Wegner, 2007; Shtulman, 2008; Shtulman & Lindeman, 2015; Lane, Wellman, & Evans, 2010, 2012, 2014) – indeed, often their own mental states (Epley, Converse, Delbosc, Monteleone, & Cacioppo, 2009) – and that in adults individual difference variables such as attachment style are associated with different mental state attributions (e.g. Granqvist & Kirkpatrick, 2008; Kirkpatrick, 2005).

Although these findings are consistent with the co-existence hypothesis they are also consistent with two alternatives: First, many adults might not have acquired the relevant theology and might not fully know which characteristics God is assumed to have. For example, Lane, Wellman, & Evans (2014) find a developmental progression in children's differentiation of God and persons supportive of this interpretation: elementary school

[1] Sperber and colleagues further hypothesize that (3) acquired beliefs that are inconsistent with core intuitions exist in a specialized meta-representational "bubble" which isolates them from core concepts which exist in a mental data-base of beliefs; in contrast to the data-base of beliefs which can be accessed unconsciously and spontaneously, meta-representations can only be accessed consciously (e.g. Sperber, 1997, 2000; Mercier & Sperber, 2009; for a related discussion of dual-process theory see Evans, 2003, 2008; Evans & Stanovich, 2013; also see Mercier & Sperber, 2011). We do not investigate this issue which, while interesting, is beyond the scope of the experiments reported here.

children but not preschoolers differentiate between the mental and perceptual capacities of persons and God, but nonetheless attribute only partial omniscience to God, and it is not until late adolescence that God is attributed with full omniscience. Additionally, Lane, Wellman, & Evans (2012) find that religiously schooled preschoolers differentiate between the capacities of persons and God earlier than secularly schooled preschoolers. In order to control for this alternative, the current study includes only those characteristics that are intuitively attributed to persons but theologically withheld from God which most religious believers from the population recruited from are likely to be familiar with. For example, the findings by Lane, Wellman, & Evans (2014) suggest that by late adolescence to early adulthood religious believers are familiar with theological notions of God's omniscience; the findings by Barrett & Keil (1996) suggest that by the same age religious believers and non-believers are familiar with a variety of theological notions pertaining to Gods psychology, perception, and physicality.

Second, many adults might have acquired the relevant theology but nonetheless intentionally deviate from it by attributing certain characteristics to God. For example, in her ethnographic study of Evangelical Christians in Chicago and the Bay Area anthropologist Tanya Luhrmann found that many of her informants attributed person-like mental states to God so as to, they reported, be able to experience God more closely and intimately (Luhrmann, 2012). But, when pressed, the informants would acknowledge that their attributions deviated from the theology of their group. In order to control for intentional deviations from theology that can occur in self-reported beliefs, the current study includes measures of beliefs not susceptible to intentional deviation.

The primary line of research satisfying the above requirements to have investigated the co-existence hypothesis (which is there termed "theological incorrectness") is the studies of memory confusions in religious adults by Barrett and colleagues (Christian adults in Barrett & Keil, 1996; Hindu adults in Barrett, 1998). For example, in Barrett & Keil (1996) participants asked to recall narratives such as about God intervening to answer a prayer were shown to mistakenly add physical or psychological limitations to God's actions not present in the original narratives (e.g. that to intervene God has to finish answering another prayer or stop another action such as helping an angel work on a crossword puzzle) and which are inconsistent with the participants' self-reported theological beliefs. The authors interpreted these memory confusions as showing that participants represented God via the agent concept, and in recalling the narrative were mistakenly relying on default intuitions about persons (e.g. sequential action).

However, the studies by Barrett and colleagues have themselves been critiqued on other grounds by other scholars. One major objection is that a person-like representation of God was implied in the narratives themselves (e.g. Shtulman, 2008)[2]. As Shtulman notes:

"God was described in other stories as pushing a large stone, looking at the rock, listening to the birds, enjoying the smell, and helping an angel work on a crossword puzzle. Any participants who might have disagreed with the anthropomorphic implications of these statements were still required to reason on their basis. To these participants, stories about a looking, listening, helping God would be as incongruent with their personal beliefs as stories

[2] A similar objection whereby the task implies that God is person-like – or, specifically, has a physical body – has been raised by Hyde (1990) toward studies in which children are asked to draw an image of God.

about a looking, listening, helping teapot, yet one could hardly fault them for drawing anthropomorphic inferences consistent with the stories' premises."

Shtulman (2008) considers it plausible that this language may have contributed to the person concept based responses seen in the recall of the narrative in the Barrett & Keil (1996) and Barrett (1998) studies.[3]

1.2 The Current Study

The primary goal of the current study is to provide a novel test of the co-existence hypothesis in religious beliefs, specifically beliefs about God among Christian religious believers. The methodology used is the sentence verification task employed by Shtulman & Valcarcel (2012)[4] which consists of an overt measure of response accuracy, and a covert measure of response time, with the latter providing the necessary control for intentional deviations from theology mentioned in section 1.1. In the task participants are required to

[3] A second objection raised by Shtulman concerns the overall rate of anthropomorphic responding in the studies by Barrett and colleagues for narratives involving religious agents, which he points out was (1) lower than 100%, and (2) not substantially different from the rate of anthropomorphic responding observed for narratives involving a supercomputer (a non-religious agents with full information access; ibid. p. 1125). However, a problem with evaluating this objection is that there is no consensus on what amount of anthropomorphism – beyond demonstrating that it exists – is required to support the co-existence hypothesis, or how minimal such responding needs to be to refute it.

[4] Shtulman & Valcarcel used this methodology to study a parallel case of the co-existence of early- and later-acquired scientific beliefs (see also: Bloom & Skolnick-Weisberg, 2007; Goldberg & Thompson-Schill, 2009; Kelemen & Rossett, 2009; Kelemen, Rottman, & Seston, 2012; Legare & Gelman, 2008; Shtulman, 2006; Shtulman & Valcarel, 2012).

endorse or reject statements of two broad kinds: consistent statements that were true or false according to both intuition and theology (e.g. “God has beliefs that are true”; “All beliefs God has are false”), and inconsistent statements that were either true intuitively but false theologically (e.g. “God has beliefs that are false”) or false intuitively but true theologically (e.g. “All beliefs God has are true”).

As per the outline of the person concept in the introduction, intuitive beliefs about both the physicality and psychology of persons were derived from reliably developing default inferences about persons that have been well established by extensive research with infants and toddlers (e.g. Carey, 2009; Baillargeon, Scott, & Bian, 2015). Acquired religious beliefs were derived from formal Christian theology that is shared between the Christian religious denominations with which the participants recruited for this study identified (primarily Catholics and Charismatic Christians in Experiment 1, and Catholics and Protestants in Experiments 2 and 3), and that previous surveys of Christian religious believers suggested our participants would be familiar with. The statements used in the study were additionally reviewed by theological officials at the churches from which participants were recruited, who confirmed that participants would have been exposed to and plausibly acquired the theological beliefs from which the statements were derived.

The current study examined four predictions which follow from the co-existence hypothesis:

(1) If core intuitions co-exist with acquired beliefs which are inconsistent with them then they might also interfere with those beliefs, and this interference might cause inconsistent statements to be responded to less accurately and more slowly than consistent statements. In principle co-existence is possible without interference, but since interference

necessarily requires co-existence, to demonstrate interference would also be to demonstrate co-existence. The first prediction therefore aims to support and extend the co-existence hypothesis as demonstrated for scientific beliefs (e.g. Shtulman & Valcarcel, 2012) to the domain of religious beliefs.

(2) If acquired religious and scientific beliefs and inconsistent core intuitions are in conflict (e.g. Barrett & Keil, 1996; Barrett, 1998, 1999; Kelemen & Rossett, 2009; Kelemen, Rottman, & Seston, 2012) then it is plausible that there are cognitive mechanisms involved in resolving this conflict. Previous findings suggest that when participants are put under time pressure their tendency to make erroneous intuitive teleological attributions is increased (e.g. Kelemen & Rosset, 2009; Kelemen, Rottman, & Seston, 2012). Therefore, we predicted that when put under time pressure participants would show a decrease in accuracy on inconsistent statements more so than on consistent statements. In both Experiments 1 and 2 a subset of participants were therefore pseudo-randomly assigned to conditions utilizing speeding manipulations. In Experiment 1 a subset of participants were assigned to a speeded instructions condition where they were encouraged to respond quickly, while the remaining subset received the task with no such encouragement (unspeeded instructions condition). In Experiment 2 a subset of participants were assigned to a speeded responding condition where a time limit was set on responses, while the remaining subset received the task with no such time limit (unspeeded responding condition).

(3) Kelemen & Rossett (2009) propose that inhibitory mechanisms might be employed to suppress certain types of teleological attributions, and found that a measure of inhibitory control (the behavioral Stroop task), was one predictor of scientific accuracy on

such responses.[5] Similarly, Lindeman & Aarnio (2007) argue that ontological confusions based on intuition are central to paranormal beliefs, and there is recent evidence that the tendency to make such confusions is related to individual differences in the efficiency of inhibitory mechanisms (Svedholm & Lindeman, 2013; see also Lindeman, Reikki, & Hood, 2011). If inhibitory resources are the mechanism which inhibits inconsistent intuitions from interfering with acquired beliefs then it is plausible that individual differences in inhibition, as indexed by the behavioral Stroop task (Kelemen & Rosset, 2009; Kelemen, Rottman, & Seston, 2012; Lindeman, Reikki, & Hood, 2011; Svedholm & Lindeman, 2013), rather than other executive functions such as working memory (Broadway & Engel, 2010; Redick et al., 2012), would be related to performance on the sentence verification task.

(4) We consider it unlikely that any amount of practice with acquired beliefs could replace inconsistent intuitions (e.g. see Goldberg & Thompson-Schill, 2009, for evidence of "childhood animism" with biology professors, and Kelemen, Rottman, & Seston, 2012, for evidence of teleological intuitions science professors), but it is plausible that practice could attenuate interference from inconsistent intuitions. We therefore predicted that individual differences in practice with acquired religious and scientific beliefs (indexed with measures of religion and science education) would be related to performance on the sentence verification task.

A secondary goal of the current study was to conduct a full replication of the Shtulman & Valcarel (2012) experiment assessing the co-existence of early- and later-

[5] Kelemen & Rossett (2009) distinguish between 'implicitly' held teleological biological beliefs, where participants would need to inhibit a tendency to endorse statements of such beliefs to be correct, from 'explicitly' held teleological biological beliefs, where (incorrect) endorsements would happen because people explicitly hold those beliefs, with no inhibition being deployed (Experiment 2, pp 141-142).

acquired scientific beliefs. In the current climate of concern over the replicability of findings across science, but especially in psychological science (Ioannidis, 2005; Nosek & Lakens, 2014; see also other papers in that special issue), undertaking replications of existing findings alongside attempted extensions is one valuable additional tool available to the psychological scientific community that might eventually offset the problem of (lack of) replication. Accordingly, the new religion statements were intermixed in the current study with the entire set of science statements used by Shtuman & Valcarel (2012) so as to determine if that result replicated, albeit in a different design context. It was predicted that the same pattern of findings would emerge.

2.0 Experiment 1

2.1 Method

2.1.1 Participants

Participants were 44 university students (56% female), ranging in age from 18 to 24 (M = 20), and drawn from two different samples: (1) two local churches, one Catholic and one Charismatic, that serve an almost exclusively college-aged population (these participants were paid for their time), and (2) the psychology participant pool at a university in Southern California (these participants received class credit). Fifty four percent of participants identified as White, 25% identified as Hispanic or Latino, and 20% identified as Asian.

All participants were pre-selected according to the following criteria, meant to ensure primary exposure to and belief in only one religion: they needed to have been brought up within the same Christian denomination with which they currently identify, and to identify as religious believers (indexed as a minimum rating of "slightly religious" on a one-item

religiosity question). Five participants did not match these criteria, despite the initial pre-selection, and were not included in the final sample (N=39), who were assigned pseudo-randomly to receive the task under speeded instructions (N=19) or unspeeded instructions (N=20).

Of the final sample fifty percent of participants identified as Roman Catholic, 41% identified as simply “Christian”, and 9% identified as one of a number of Protestant Christian denominations (e.g. Baptist, Lutheran). Since Charismatic Christians don’t usually associate with a specific Christian denomination it is likely that many of the participants who identified themselves as "Christian” in this sample were Charismatic Christians. In support of this, many participants who identified as “Christian” in this sample also reported being affiliated with the local Charismatic church. On a 4-point Likert scale (range 0 to 3; Not at all, Slightly, Moderately, Very), participants on average reported being moderately religious ($M = 2.20$, $SD = .83$) and moderately spiritual ($M = 2.01$, $SD = .87$), and the two were highly correlated ($r = .522$, $p = .001$).

2.1.2 Design

The primary dependent variable was response accuracy, and with respect to this dependent variable, the design was a 2 (Domain: Religion versus Science) x 2 (Consistency: Consistent versus Inconsistent) x 2 (Instructions: Speeded versus Unspeeded) factorial with within-subjects repeated measures on the first two factors.

Response time data were also collected for participants in the speeded instructions condition. For this dependent variable the design was a 2 (Domain: Religion versus Science)

x 2 (Consistency: Consistent versus Inconsistent) factorial with within-subjects repeated measures.

2.1.3 Materials

The religion statements were constructed in groups of four statements, with each group targeting a particular physical or psychological limitation that is typically attributed to persons but withheld from God in reflective Christian theology. For example, one group of statements concerned God's beliefs (assumed to be always true), while another concerned the necessity of communication for God to obtain information (God is assumed to know the contents of prayers without those needing to be communicated), and another concerned the necessity for God to have a physical body to act on the world (God is assumed not to have a physical body and not to require a physical body to act on the world). For each group, a pair of consistent and inconsistent statements was constructed such that there was one that was true on both intuition and theology, one that was true on neither, one that was true only on intuition, and one that was true only on theology. There were 12 groups of 4 statements each, for a total of 48 statements. Example statements appear in Table 1, and a full list of all religion statements can be found in the Appendix.

Following Shtulman & Valcarcel (2012) there were an equal number of statements that were objectively true and false to deter response strategies, and the four statements within each group were balanced in terms of overall sentence structure, complexity, and length in words.

The science statements used were the same 200 statements used by Shtulman & Valcarcel, covering 10 areas of mathematics and science (astronomy, evolution, fractions,

genetics, germs, matter, mechanics, physiology, thermodynamics, and waves). Example statements appear in Table 1, and a full list of all science statements can be found in the Appendix.

Table 1

Sample Statements from the Domains of Religion and Science.

Intuition	Reflection	Religion Statements	Science Statements
Consistent			
T	T	God has beliefs that are true.	Rocks are composed of matter.
F	F	All beliefs God has are false.	Numbers are composed of matter.
Inconsistent			
T	F	God has beliefs that are false.	Fire is composed of matter.
F	T	All beliefs God has are true.	Air is composed of matter.
Consistent			
T	T	God can hear what I say out loud.	People turn food into energy.
F	F	God can't hear what I say out loud.	Rocks turn food into energy.
Inconsistent			
T	F	God can't hear what I say to myself.	Plants turn food into energy.
F	T	God can hear what I say to myself.	Bacteria turn food into energy.
Consistent			
T	T	God knows what I want and what I pray for.	Humans are descended from tree-dwelling creatures.
F	F	Even if I pray for it, God won't know what I want.	Humans are descended from plants.
Inconsistent			
T	F	God won't know what I want unless I pray for it.	Humans are descended from chimpanzees.
F	T	God will know what I want even if I don't pray for it.	Humans are descended from sea-dwelling creatures.
Consistent			
T	T	God can see people's actions.	A moving bullet loses speed.
F	F	God can't see people's actions.	A moving bullet loses weight.
Inconsistent			
T	F	God needs eyes to see.	A moving bullet loses force.
F	T	God can see without eyes.	A moving bullet loses height.

Note. Statements that are true both intuitively and reflectively are considered consistent; statements that are true on one and false on the other are considered inconsistent. Science statements are from Shtulman & Valcarcel (2012).

Additional materials included (1) a 144-item behavioral Stroop task which included the following three conditions (48 items per condition): Congruent (the words RED, BLUE, GREEN, and YELLOW appearing in red, blue, green, and yellow color, respectively), Incongruent (the words RED, BLUE, GREEN, and YELLOW appearing in a color different

than the one they spell), and Neutral (the words LOT, SHIP, KNIFE, FLOWER – length-matched and frequency-matched to the color words, appearing in colors); a Stroop response time interference score is then calculated by subtracting the mean response time on the neutral condition from the mean response time on the incongruent condition, (2) a running span working memory task (Broadway & Engel, 2010), and (3) a short survey seeking demographic information, measures of both self-reported religiosity and spirituality, extent of participants' religious education (measured as frequency of monthly attendance at Church and as years of formal religious education), and extent of math and science education (measured by asking participants to list all math and science courses taken in college).

2.1.4 Procedure

In a quiet testing room, groups of up to 6 participants took the experiment at semi-private computer testing stations. Participants in both the speeded and unspeeded instructions conditions completed, in this order, the sentence verification task, the behavioral Stroop task, the working memory task, and the survey. In the speeded instructions condition the instructions to the sentence verification task emphasized both response accuracy and speed (in multiple parts of the instructions participants were told to "respond as quickly as you can, while making as few mistakes as you can" and that "speed and accuracy are both very important"), and responses were collected via key presses to facilitate faster and less deliberate responding (presented via E-Prime software). In the unspeeded instructions condition the instructions emphasized accuracy only, and responses were presented in survey form (presented via Qualtrics software) to facilitate slower and more deliberate responding. In both instructions conditions the sentence verification task items were presented one-by-

one and in a randomized order, and whether the right or left hand were used to respond "true" or "false" was randomized between participants.

2.2 Results

2.2.1 Sentence response accuracy

The primary hypothesis under test was that participants will be more accurate responding to items in which core intuitions are consistent with acquired beliefs in the domains of religion and science. The sentence response accuracy data were subjected to a 2 (Domain: religion versus science) X 2 (Consistency: consistent versus inconsistent) X 2 (Instructions: unspeeded versus speeded) mixed analysis of variance (ANOVA) with repeated measures on the first two factors, revealing main effects of Domain ($F_{1, 37} = 348.7$, $p < .001$, partial $\eta^2 = .90$) and Consistency ($F_{1, 37} = 226.2$, $p < .001$, partial $\eta^2 = .86$) qualified by an interaction between Domain and Consistency ($F_{1, 37} = 50.5$, $p < .001$, partial $\eta^2 = .58$). There was no main effect of Instructions, and the Instructions factor did not enter into any two- or three-way interactions (all Fs < 1.3, all ps = n.s.). With no main effect of Instructions and no interactions involving the Instructions factor, no further analyses involving it are reported. The interaction between Domain and Consistency is shown for both instructions conditions in Figure 1.

Simple main effect analyses confirmed that participants performed better on the religion items than on the science items for both the consistent ($M_{religion} = 98.10\%$, $SD_{religion} = 4.26\%$ versus $M_{science} = 84.28\%$, $SD_{science} = 5.70\%$; t (38) = 12.42, p < .001, $d = 2.75$) and inconsistent items ($M_{religion} = 89.32\%$, $SD_{religion} = 9.97\%$ versus $M_{science} = 62.23\%$, $SD_{science} = 9.51\%$; t (38) = 15.93, p < .001, $d = 2.78$), and that the interaction resulted from the size of

the effect for consistency being more than twice as large for the science items ($M_{consistent}$ = 84.28%, $SD_{consistent}$ = 5.70%; versus $M_{inconsistent}$ = 62.23%, $SD_{inconsistent}$ = 9.51%; t (38) = 16.90, $p < .001$, $d = 2.81$) than for the religion items ($M_{consistent}$ = 98.10%, $SD_{consistent}$ = 4.26%; versus $M_{inconsistent}$ = 89.32%, $SD_{inconsistent}$ = 9.97%; t (38) = 6.06, $p < .001$, $d = 1.15$). [6]

These results support our first prediction by replicating and extending the findings of Shtulman & Valcarel (2012) to show that there is a similar conflict between core intuitions and acquired beliefs in the domain of religion. However, our second prediction that the difference in response accuracies between consistent and inconsistent items would be greater under instruction to respond quickly was not supported.

[6] Shapiro-Wilk tests revealed that several of the response accuracy distributions were not normal (consistent religion items in the unspeeded instructions condition: SW = .455, df = 20, $p < .001$; consistent and inconsistent religion items in the speeded instructions condition: SW = .599, df = 19, $p < .001$, SW = .681, df = 19, $p < .001$, respectively). The simple main effects were therefore analyzed with non-parametric tests to supplement the parametric tests reported here. A series of planned comparisons using the Wilcoxon Signed-Rank Test confirmed all findings reported here (all $Zs < -4.83$, all $ps < .001$).

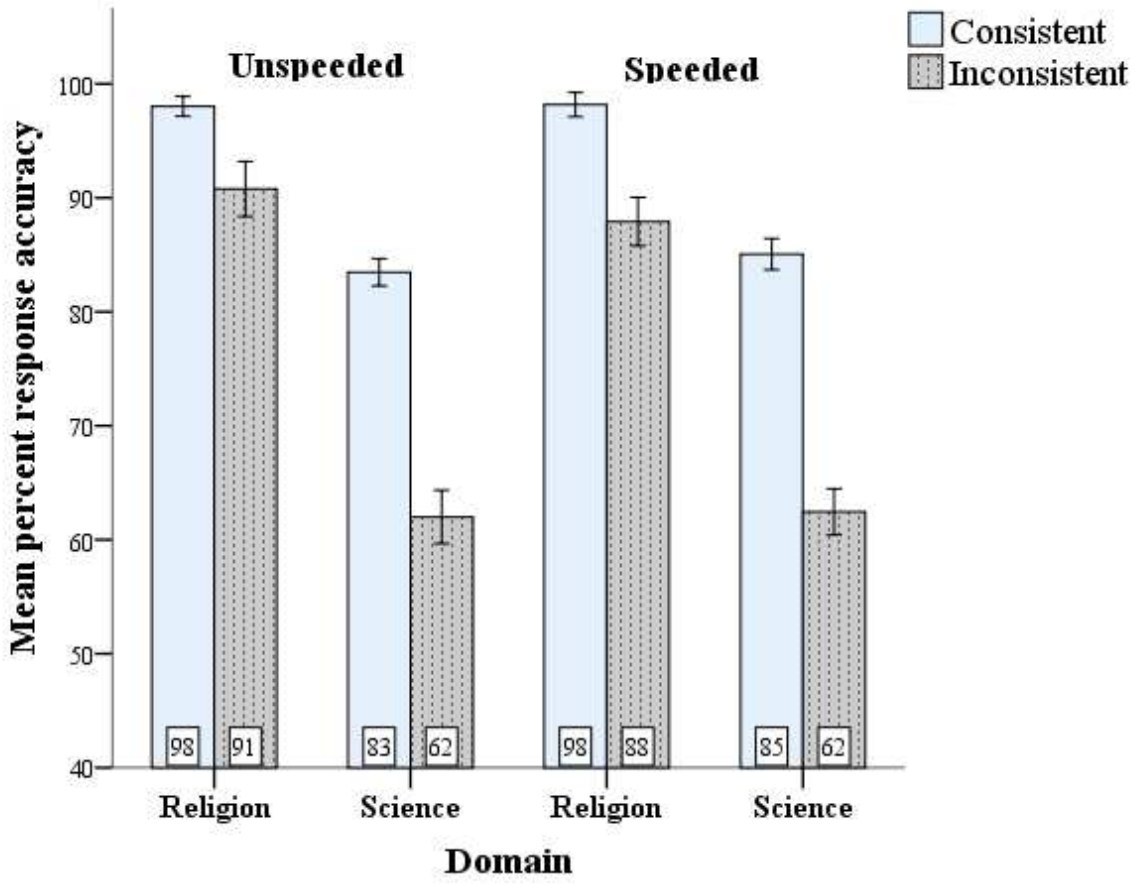

Figure 1. Mean percent response accuracy on consistent and inconsistent items in the domains of religion and science, for both the speeded and unspeeded instructions conditions. Error Bars: +/- 1 SE.

2.2.2 Sentence response time[7]

Sentence response time data were collected for those participants who received speeded instructions. Response time data for correct responses were entered into a 2 (Domain; Religion versus Science) X 2 (Consistency; Consistent versus Inconsistent) repeated measures ANOVA, revealing a main effect of Consistency only ($F_{1, 19} = 40.88$, $p < .001$, partial $\eta^2 = .68$). There was no main effect of Domain and no interaction (both Fs < 1.0,

[7] A very small number of response time data points (<1%) were removed for being more than 3SD beyond the mean response time of any given participant.

ps = n.s.). Simple main effect analyses confirmed that participants were faster on the consistent items than they were on the inconsistent items for both religion items ($M_{\text{consistent}}$ = 2972ms, $SD_{\text{consistent}}$ = 703ms versus $M_{\text{inconsistent}}$ = 3219ms, $SD_{\text{consistent}}$ = 765ms), $F_{1, 19}$ = 11.58, $p < .005$, $d = 0.34$) and science items ($M_{\text{consistent}}$ = 2971ms, $SD_{\text{consistent}}$ = 505ms versus $M_{\text{inconsistent}}$ = 3298ms, $SD_{\text{consistent}}$ = 563ms), $F_{1, 19} = 29.55$, $p < .001$, $d = 0.61$), and that there were no differences in response times between science and religion items for either consistent (M_{religion} = 2972ms, SD_{religion} = 703ms versus M_{science} = 2971ms, $SD_{\text{consistent}}$ = 505ms), or inconsistent items (M_{religion} = 3219ms, SD_{religion} = 765ms versus M_{science} = 3298ms, $SD_{\text{consistent}}$ = 563ms), both $Fs < 1.0$, ps = n.s.).

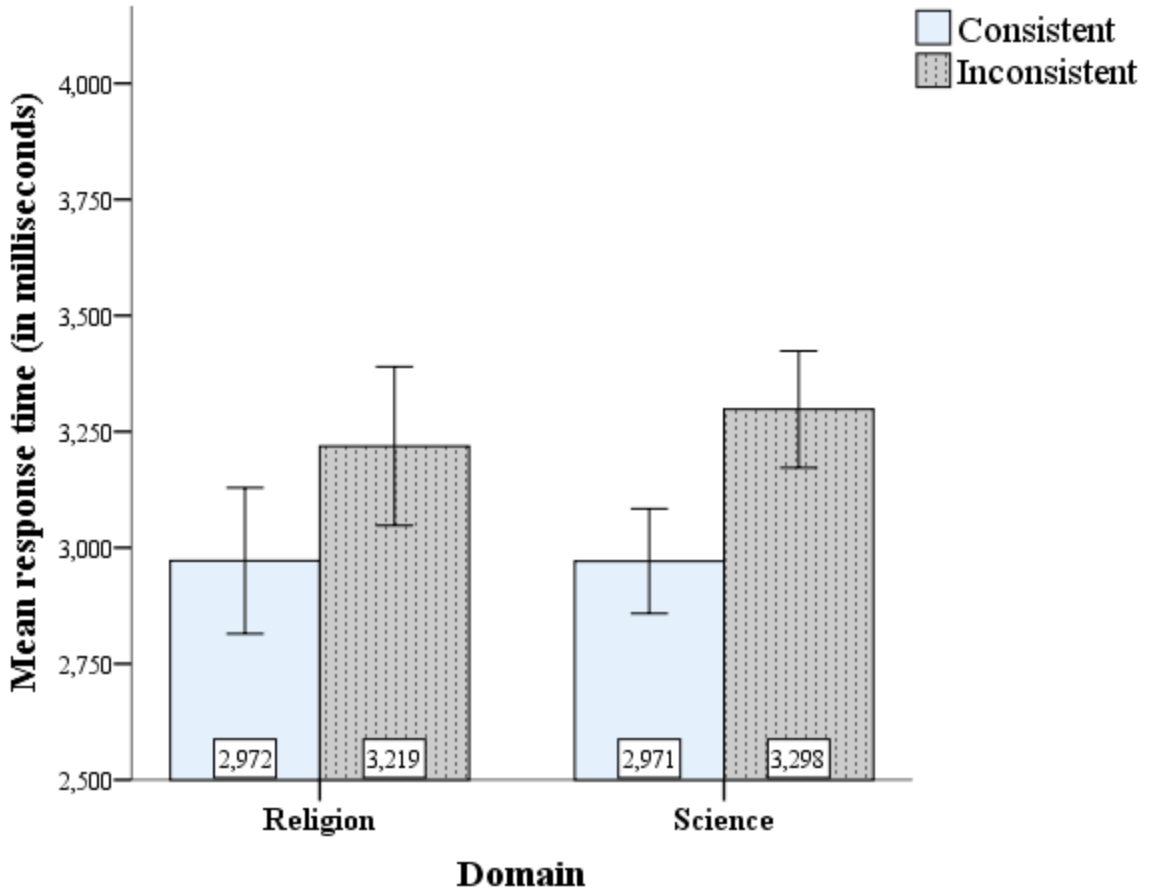

Figure 2. Mean response time (in milliseconds) on consistent and inconsistent items in the domains of religion and science, for the speeded instructions condition only. Error Bars: +/- 1 SE.

2.2.3 Associations with measures of executive functions

The inhibition measure (Stroop response time interference scores: M = 81ms, SD = 76ms, for correct responses only) along with the working memory measure (M = 25.62, SD = 12.26) were entered into a correlational analysis with interference scores on accuracy (both speeded and unspeeded instructions) and response time (speeded instructions only) on the sentence verification task for both the religion and science items, calculated as the difference between the consistent and inconsistent conditions. In this way accuracy on the consistent condition serves as a baseline index of familiarity with the domain under consideration and

the response time serves as a baseline index of response rate in the domain. Lower scores for response accuracy and higher scores for response time on the inconsistent condition serve as the index of the degree of interference of intuitions on acquired beliefs. Neither of the two executive functions measures were correlated with either the accuracy or response time interference scores in either the religion or science domains (all ps = n.s.).

2.2.4 Associations with measures of education

Participants were assigned an education composite score, calculated by summing the number of content areas – mathematics, biology, chemistry, and physics – they have taken at least one course in. The composite education variable ranged from 1 to 4 (M = 2.69, SD = 1.00), and did not correlate with either the accuracy (both speeded and unspeeded instructions) or response time (speeded instructions only) interference scores for science items (both ps = n.s.). Shtulman & Valcarcel similarly collected data on the math and science courses their participants have taken, and using a slightly different variable (total number of courses taken, rather than the composite score for content areas used here) similarly did not find that it predicted any of the effects reported in their study.

Two dichotomous measures of religious education were considered: regular monthly attendance at Church (66% of participants reported attending church at least once a month), and theology study (42% of participant reported having studied theology). A one-way analysis of variance (ANOVA) revealed that regular monthly attendance at church did not predict either the accuracy (both speeded and unspeeded instructions) or response time (speeded instructions only) interference scores for religion items (both Fs < 2.5, both ps = n.s.). However, participants who regularly attended church were overall more accurate on

both consistent (M = 99.17%, SD = 2.30% versus M = 94.55%, SD = 6.97%; $F_{1, 37}$ = 10.01, p < .01, partial η^2 = .22) and inconsistent ($M_{consistent}$ = 90.83%, $SD_{consistent}$ = 7.61% versus $M_{inconsistent}$ = 84.26%, $SD_{inconsistent}$ = 15.00%; $F_{1, 37}$ = 3.18, p = .08, partial η^2 = .08) religion items.

A one-way analysis of variance (ANOVA) additionally revealed that having studied theology predicted a lower accuracy interference score (both speeded and unspeeded instructions) for religion items ($M_{religion}$ = 6.29%, $SD_{religion}$ = 5.59% versus $M_{religion}$ = 12.01%, $SD_{religion}$ = 11.59%; $F_{1, 37}$ = 4.13, p < .05, partial η^2 = .10); it did not predict the response time interference score (speeded instructions only) for religion items (F < 2.5, p = n.s.).

3.0 Experiment 2

Experiment 1 failed to find an effect of instructions to respond quickly on the sentence verification task: participants responded with almost identical levels of accuracy under speeded and unspeeded instructions. The lack of an effect for the instructions manipulation in the current study is not critical for evaluating the co-existence hypothesis, particularly given the strong effects of consistency on response accuracy and speed. Indeed, in Kelemen & Rosset (2009) and Kelemen, Rottman, & Seston (2012) there is evidence of co-existence of teleological thinking alongside scientific thinking at the slowest response rates imposed (e.g. in the unspeeded condition in the two experiments reported in Kelemen & Rosset endorsements of teleological explanations range from 29 to 42% versus 47 to 54% in their speeded condition).

The most likely explanation for this outcome is that instructions to respond quickly alone were not a strong enough manipulation to put participants under time pressure in the speeded instructions condition. In some previous studies that used a speeding manipulation, instructions to respond quickly were accompanied by time limits on participants' response windows (e.g. Kelemen & Rossett, 2009, Kelemen, Rottman, & Seston, 2012; Lindeman, Reikki & Hood, 2011; Svedholm & Lindeman, 2013). In Experiment 2, therefore, response time limits were added to the speeding manipulation to further investigate the prediction concerning stressing executive resources derived from the co-existence hypothesis.

3.1 Method

3.1.1 Participants

Participants were 75 university students (80% female), ranging in age from 18 to 24 (M = 19), and drawn from the psychology participant pool at a university in Southern California. All participants received class credit for their time. Thirty eight percent of participants identified as White, 29% identified as Hispanic or Latino, and 29% identified as Asian, and 4% identified as "Other".

As in experiment 1 all participants were pre-selected according to criteria meant to ensure primary exposure to and belief in only one religious tradition (see Experiment 1 for details). Four participants did not match these criteria, despite the initial pre-selection, and were not included in the final sample (N = 71), who were assigned pseudo-randomly to receive the task under time limit (N = 32) or no time limit (N = 39).

Of the final sample forty percent of participants identified as Roman Catholic, 39% identified as non-denominational Christian[8], and 21% identified as one of a number of Protestant Christian denominations (e.g. Presbyterian, Baptist). On a 4-point Likert scale (range 0 to 3) participants on average reported being moderately religious ($M = 1.93$, $SD = .64$) and moderately spiritual ($M = 1.83$, $SD = .83$), and the two were highly correlated ($r = .571$, $p < .001$).

3.1.2 Design

As in Experiment 1 the primary dependent variable was response accuracy, and a secondary dependent variable was response time which was collected for participants in both conditions. The design was a 2 (Domain: Religion versus Science) x 2 (Consistency: Consistent versus Inconsistent) x 2 (Condition: time limit versus no time limit on responding) factorial design with within-subjects repeated measures on the first two factors.

3.1.3 Materials

The materials used were identical to those used in Experiment 1 with the exception of the behavioral Stroop task which was modified to resemble the task used by Lindeman, Reikki & Hood (2011) and Svedholm & Lindeman (2013). The "congruent" condition was replaced with a "color-naming" condition in which a string of Xs appears in red, blue, green, or yellow color and participants are required to respond to the color in which the Xs appear, and the "neutral" condition was replaced with a "word-naming" condition in which the words

[8] Unlike in Experiment 1 where most of the participants who identified as "Christian" reported being affiliated with the local Charismatic church none of the participants who identified as "Christian" in this experiment did.

RED, BLUE, GREEN, and YELLOW appear in black, and participants are required to respond to the word that appears. The "incongruent" condition (the words RED, BLUE, GREEN, and YELLOW appearing in a color different than the one they spell) remained the same. A Stroop response time interference score is then calculated by subtracting response times on the color-naming condition from the incongruent condition. The word-naming condition is not used in calculations; it is meant to discourage a response strategy whereby participants might only attend to the color in which a word appears by either directing their gaze to the periphery of the display or by blurring the written word by squinting their eyes (both strategies participants reported using in informal debriefings).

3.1.4 Procedure

The procedure was the same as in Experiment 1, except that in the speeded condition in addition to instructions emphasizing both response accuracy and speed in the sentence verification task, participants were told that each statement will appear for a short duration, and that the durations will be of variable times (this variability had the added benefit of deterring response strategies). The actual times (M = 3298ms, SD = 962ms; range 1605ms to 6749ms) were determined through pre-testing (n = 15) as the average reading time plus two standard deviations of each statement. In the unspeeded condition the instructions emphasized response accuracy only and statements appeared until participants responded. In both conditions responses were collected via key presses (presented via E-Prime software).

3.2 Results

3.2.1 Sentence response accuracy[9]

As in Experiment 1, the primary hypothesis under test was that participants should be more accurate responding to items in which core intuitions are consistent with acquired beliefs in the domains of religion and science. The accuracy data were subjected to a 2 (Domain; religion versus science) X 2 (Consistency; consistent versus inconsistent) X 2 (Condition; time limit versus no time limit) mixed analysis of variance (ANOVA) with repeated measures on the first two factors, revealing main effects of Domain (science versus religion; $F_{1, 69} = 480.2$, $p < .001$, partial $\eta^2 = .87$), Consistency (consistent versus inconsistent; $F_{1, 69} = 576.4$, $p < .001$, partial $\eta^2 = .89$), and Condition ($F_{1, 69} = 60.5$, $p < .001$, partial $\eta^2 = .47$), qualified by interactions between Domain and Consistency ($F_{1, 69} = 60.0$, $p < .001$, partial $\eta^2 = .46$) and Consistency and Condition ($F_{1, 69} = 4.8$, $p < .05$, partial $\eta^2 = .06$). The interaction between Domain and Condition was not statistically significant ($F_{1, 69} = 2.4$, $p = .123$, partial $\eta^2 = .03$), and the factors did not enter into a three-way interaction ($F_{1, 69} = 2.7$, $p = .105$, partial $\eta^2 = .04$). The interaction between Domain and Consistency is shown for both the unspeeded (no time limit) and speeded (time limit) conditions in Figure 3.

[9] In the speeded (time limit) condition responses that were timed-out were considered incorrect.

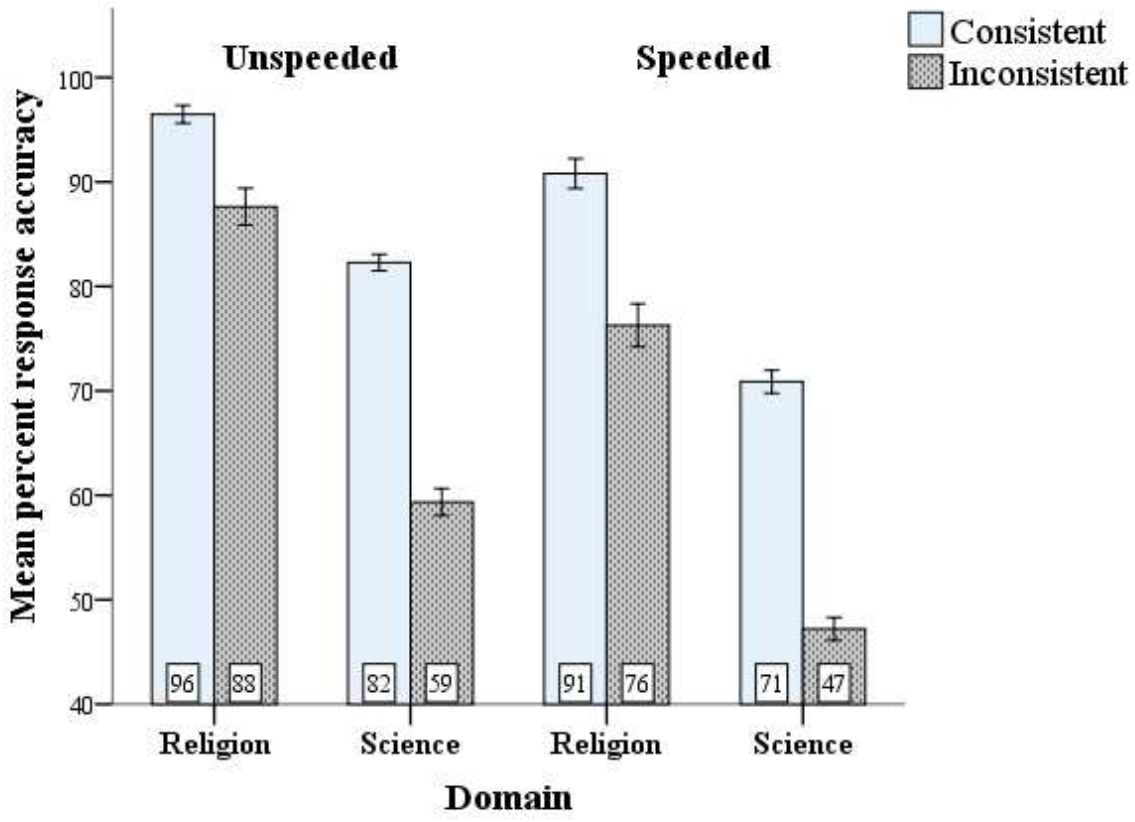

Figure 3. Mean percent response accuracy on consistent and inconsistent items in the domains of religion and science, for both the unspeeded (no time limit) and speeded (time limit) conditions. A mean response accuracy of 50% represents chance responding. Error Bars: +/- 1 SE.

Simple main effect analyses replicated the findings reported in Experiment 1 by confirming that participants performed better on the religion items than on the science items for both consistent ($M_{religion}$ = 93.37%, $SD_{religion}$ = 7.83% versus $M_{science}$ = 76.01%, $SD_{science}$ = 8.19%; $t\ (70)$ = 14.92, $p < .001$, d = 2.17) and inconsistent items ($M_{religion}$ = 81.40%, $SD_{religion}$ = 12.87% versus $M_{science}$ = 52.68%, $SD_{science}$ = 9.25%; $t\ (70)$ = 20.32, $p < .001$, d = 2.56), and that the interaction between Domain and Consistency resulted from the size of the effect for consistency being more than twice as large for the science items ($M_{consistent}$ = 76.01%, $SD_{consistent}$ = 8.19 %; versus $M_{inconsistent}$ = 52.68%, $SD_{inconsistent}$ = 9.25%; $t\ (70)$ =

30.01, p < .001, d = 2.67) than for the religion items ($M_{consistent}$ = 93.37%, $SD_{consistent}$ = 7.83%; versus $M_{inconsistent}$ = 81.40%, $SD_{inconsistent}$ = 12.87%; $t\ (70)$ = 9.30, p < .001, d = 1.12).[10]

An examination of the mean differences, separated by condition, between consistent and inconsistent religion ($M_{unspeeded}$ = 5.67%, versus $M_{speeded}$ = 11.35%) and science ($M_{unspeeded}$ = 11.41%, versus $M_{speeded}$ = 12.13%) items revealed that the two-way interaction between Consistency and Condition was primarily carried by the religion items.[11] The response accuracy data suggest that this was caused by a floor effect in responses to inconsistent science items in the time limit conditions (response accuracy on these items was approximately 50%, which is chance responding).

[10] Shapiro-Wilk tests revealed that several of the response accuracy distributions were not normal (consistent and inconsistent religion items in the no time limit condition: SW = .731, df = 32, p < .001, SW = .872, df = 32, p = .001, respectively; consistent religion items and inconsistent science items in the time limit condition: SW = .874, df = 39, p < .001, SW = .941, df = 39, p < .05, respectively). The simple main effects were therefore analyzed with non-parametric tests to supplement the parametric tests reported here. A series of planned comparisons using the Wilcoxon Signed-Rank Test confirmed all findings reported here (all Zs < -6.67, all ps < .001).

[11] A separate examination of the religion and science items each via a 2 (Consistency; consistent versus inconsistent) X 2 (Condition; time limit versus no time limit) repeated measures analysis of variance (ANOVA) further confirmed that despite the absence of a three-way interaction the two-way interaction between Consistency and Condition was carried by the religion items (religion: $F_{1, 69}$ = 5.1, p < .05, partial η^2 = .07; science: $F_{1, 69}$ = .22 p = n.s., partial η^2 = .00). The results from this mixed ANOVA were additionally replicated with an Extended Linear-Mixed Effects Model, which accommodated different error-variances between the two Consistency factors (the error variance was greater for consistent than for inconsistent items in both Condition factors). The interaction for Consistency and Condition for the religion domain remained significant at p < 0.05.

An analysis of the timed-out responses in the time limit condition further supports this interpretation. A 2 (Domain; religion versus science) X 2 (Consistency; consistent versus inconsistent) repeated-measures ANOVA with proportion of incorrect responses that were due to time-outs as the DV revealed main effects of Domain ($F_{1, 38} = 5.75$, $p < .05$, partial $\eta^2 = .13$) and Consistency ($F_{1, 38} = 15.48$, $p < .001$, partial $\eta^2 = .29$) and no interaction. A larger number of incorrect responses on science items than on religion items were due to time-outs ($M_{science} = 25.60\%$ versus $M_{religion} = 18.70\%$), and a larger number of incorrect responses on consistent items than on inconsistent items were due to time-outs ($M_{consistent} = 27.90\%$ versus $M_{inconsistent} = 16.40\%$). A plausible interpretation of this pattern is that overall the science items were more difficult than the religion items, but since they were allowed approximately similar (or often shorter) responding durations, participants who tried to think about them for too long before responding were timed-out. In contrast, inconsistent items (particularly inconsistent science items) were significantly more difficult than consistent items, so much so that participants often chose an answer at random, causing faster response times and fewer time-outs.

3.2.2 Sentence response time[12]

Response time data for correct responses in the unspeeded condition only[13] were entered into a 2 (Domain; religion versus science) X 2 (Consistency; consistent versus

[12] A very small number of response time data points (<2%) were removed for being more than 3SD beyond the mean response time of any given participant.

[13] Response times in the speeded condition were truncated by the time limit manipulation and were therefore excluded from this analysis.

inconsistent) repeated measures analysis of variance (ANOVA), revealing a main effect of Consistency ($F_{1, 31} = 37.86$, $p < .001$, partial $\eta^2 = .55$) and an interaction between Domain and Consistency ($F_{1, 31} = 21.98$, $p < .001$, partial $\eta^2 = .42$). Simple main effect analyses revealed that the main effect of Consistency was carried by the science items (science: $M_{consistent} = 3176$ms, $SD_{consistent} = 552$ms versus $M_{inconsistent} = 3664$ms, $SD_{inconsistent} = 613$ms, $t(31) = 9.22$, $p < .001$, $d = .84$; religion: $M_{consistent} = 3277$ms, $SD_{consistent} = 693$ms versus $M_{inconsistent} = 3364$ms, $SD_{inconsistent} = 563$ms, $t(31) = 1.21$, p = n.s., $d = .14$), and that despite the lack of a main effect of Domain in the preceding analysis ($F_{1, 31} = 1.53$, p = n.s., partial $\eta^2 = .05$) the response time for inconsistent religion items was significantly faster than for inconsistent science items ($M_{religion} = 3364$ms, $SD_{religion} = 563$ms versus $M_{science} = 3664$ms, $SD_{science} = 613$ms, $t(31) = 3.36$, $p = .002$, $d = .51$). The mean response times are shown in Figure 4.

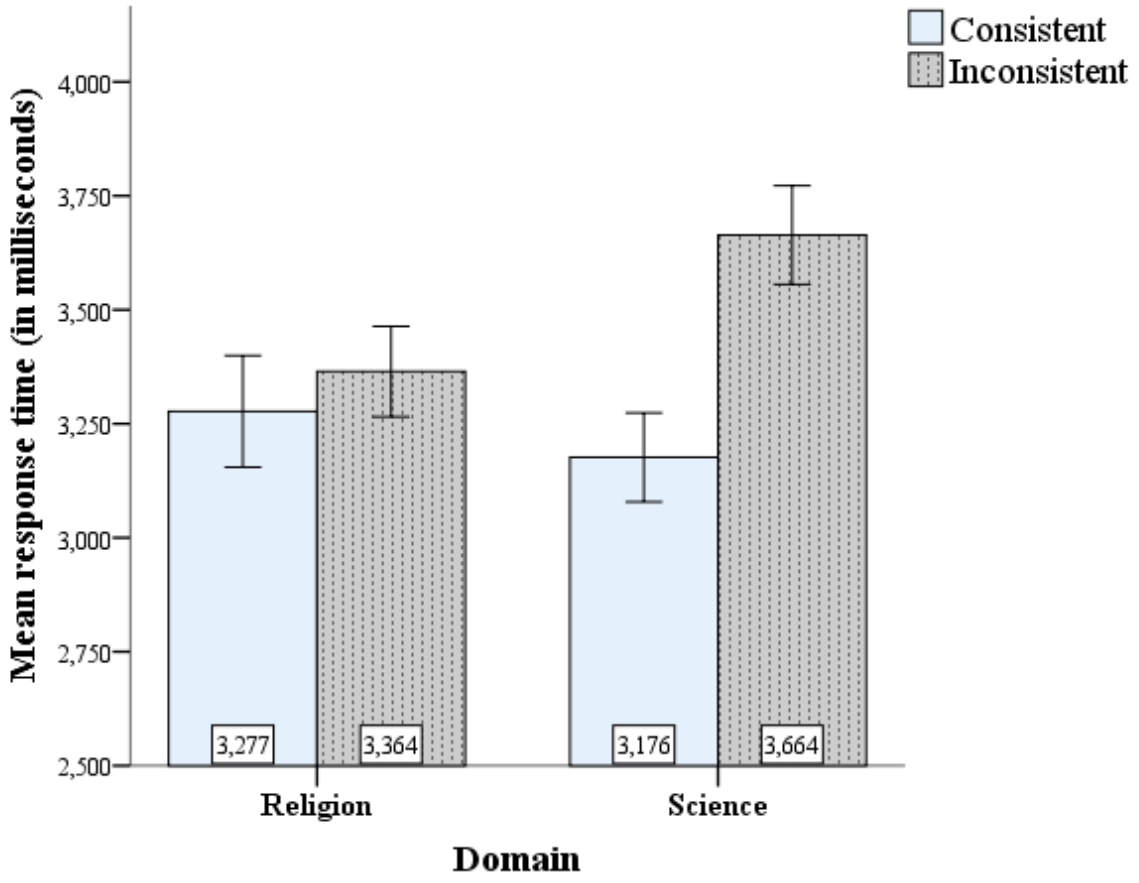

Figure 4. Mean response time (in milliseconds) on consistent and inconsistent items in the domains of religion and science, for the unspeeded (no time limit) condition only. Error Bars: +/- 1 SE.

3.2.3 Associations with measures of executive functions

As in Experiment 1 the Stroop response time interference scores (M = -209ms, SD = 119ms, for correct responses only) along with the working memory measure (M = 22.75, SD = 10.46) were entered into a correlational analysis with interference scores on accuracy (both conditions) and response time (no time limit condition only) on the sentence verification task for both the religion and science domains. Neither the inhibition nor the working memory measures were correlated with either the accuracy or response time interference scores for either the religion or science domains (all ps > n.s.).

3.2.4 Associations with measures of religious education[14]

Two dichotomous measures of religious education were considered: regular monthly attendance at Church (66% of participants reported attending church at least once a month), and theology study (42% of participant reported having studied theology). A one-way analysis of variance (ANOVA) revealed that regular monthly attendance at church did not predict either the accuracy or response time interference scores for religion items (all Fs < 1.0, all ps = n.s.). However, as in Experiment 1, participants who regularly attended church were overall more accurate on both consistent ($F_{1, 69} = 12.82$, $p = .001$, partial $\eta^2 = .16$) and inconsistent ($F_{1, 69} = 5.44$, $p < .05$, partial $\eta^2 = .07$) religion items ($M_{consistent} = 89.06\%$, $SD_{consistent} = 9.81\%$ versus $M_{inconsistent} = 95.57\%$, $SD_{inconsistent} = 5.52\%$).

A one-way analysis of variance (ANOVA) additionally revealed that having studied theology predicted a smaller response time interference score for religion items ($M_{religion} = -90$ms, $SD_{religion} = 468$ms versus $M_{religion} = 209$ms, $SD_{religion} = 321$ms; $F_{1, 30} = 4.62$, $p < .05$, partial $\eta^2 = .13$); it did not predict the accuracy interference score for religion items (F < 1.0, p = n.s.).

4.0 Experiment 3

The primary prediction of the current study whereby inconsistent statements should be responded to less accurately and more slowly than consistent statements was supported in

[14] Associations with measures of math and science education weren't examined in this experiment because there was not sufficient variability in this sample in the number of courses taken within the different math and science content areas.

Experiments 1-2. The goal of Experiment 3 was to control for the possibility that this pattern was caused by low-level biases in the statements used, rather than by co-existence and interference of core intuitions and acquired beliefs. For example, cognitive processing of the inconsistent statements within each group might have been systematically higher than that of the consistent statement as a result of their construction. To test this, in a subset of the religion statements in Experiment 3 the supernatural entity "God" was replaced with a non-supernatural religious agent ("a priest"); it was predicted that since participants hold no acquired beliefs inconsistent with core intuitions about this non-supernatural agent there should be no differences in accuracy or in response time on the modified "inconsistent" and "consistent" (these terms were retained for ease of comparison) statements.

4.1 Method

4.1.1 Participants

Participants were 37 university students (81% female) ranging in age from 18 to 21 (M = 19) drawn from the psychology participant pool at a Southern California university. All participants received class credit for their time. Thirty eight percent of participants identified as White, 24% identified as Hispanic or Latino, 8% identified as black, 19% identified as Asian, and 11% identified as "Other".

As in experiments 1-2 all participants were pre-selected according to criteria meant to ensure primary exposure to and belief in only one religious tradition (see Experiment 1 for details). Four participants did not match these criteria, despite the initial pre-selection, and were therefore not included in the final sample (N = 33).

Of the final sample thirty percent of participants identified as Roman Catholic, 37% identified as non-denominational Christian, and 33% identified as one of a number of Protestant Christian denominations (e.g. Methodist, Pentecostal). On a 4-point Likert scale (range 0 to 3) participants on average reported being moderately religious ($M = 1.85$, $SD = .51$) and moderately spiritual ($M = 1.61$, $SD = .70$), and the two were highly correlated ($r = .44$, $p = .01$).

4.1.2 Design

As in Experiments 1-2 the primary dependent variable was response accuracy and a secondary dependent variable was response time. The design was a 2 (Domain: Religion versus Science) x 2 (Consistency: Consistent versus Inconsistent) factorial design with within-subjects repeated measures.

4.1.3 Materials

A subset of the religion statements (sets T3, T5, T6, T8, and T9 found in the Appendix) was modified by replacing the word "God" with "my priest". The religion statements modified in this manner were not selected at random, but were selected because they could be modified and still remain coherent. For example, a number of statements could not be modified because they had to do with God listening to prayers, and prayers are directed toward God but not toward priests. Example statements appear in Table 2.

The modified religion statements were presented in random order along with the unmodified religion statements (same as in Experiments 1-2) and the entire set of science

statements. No measures of executive functions or of education were administered in this experiment.

Table 2

Sample Modified Statements from the Domain of Religion.

Intuition	Reflection	Religion Statements	Modified Religion Statements
Consistent			
T	T	God has beliefs that are true.	My priest has beliefs that are true.
F	F	All beliefs God has are false.	All beliefs my priest has are false.
Inconsistent			
T	F	God has beliefs that are false.	My priest has beliefs that are false.
F	T	All beliefs God has are true.	All beliefs my priest has are true.

Note. The terms "consistent" and "inconsistent" do not apply to the modified religion statements but are used for illustrative purposes only. The modified religion statements are true or false according to intuition only.

4.1.4 Procedure

The procedure was the same as in Experiments 1-2, except that all participants were assigned to the same unspeeded condition (identical to the unspeeded condition in Experiment 2).

4.2 Results

4.2.1 Sentence response accuracy

A paired-samples t-test on the 20 modified religion items revealed no differences in response accuracy between the modified "consistent" and "inconsistent" items ($M_{consistent}$ = 77.30 %, $SD_{consistent}$ = 15.47%; versus $M_{inconsistent}$ = 73.90%, $SD_{inconsistent}$ = 18.53%; t (32) = .69, p > .5, d = .20). In contrast, a paired-samples t-test on these same unmodified items

(using the "God" rather than "my priest") from the no time limit condition from Experiment 2 revealed a difference between consistent and inconsistent items in response accuracy ($M_{consistent}$ = 96.60 %, $SD_{consistent}$ = 5.45%; versus $M_{inconsistent}$ = 88.70%, $SD_{inconsistent}$ = 13.62%; t (31) = 3.14, p < .01, d = .76).

Next, to examine if the main findings from Experiments 1-2 replicated in the 28 unmodified religion items the response accuracy data were subjected to a 2 (Domain; religion versus science) X 2 (Consistency; consistent versus inconsistent) repeated measures analysis of variance (ANOVA),. The analysis revealed main effects of Domain (science versus religion; $F_{1, 32}$ = 195.41, p < .001, partial η^2 = .86) and Consistency (consistent versus inconsistent; $F_{1, 32}$ = 307.20, p < .001, partial η^2 = .91) qualified by an interaction between Domain and Consistency ($F_{1, 32}$ = 53.23, p < .001, partial η^2 = .62).

Simple main effect analyses confirmed that as in Experiments 1-2 participants performed better on the religion items than on the science items in both the consistent ($M_{religion}$ = 97.19%, $SD_{religion}$ = 6.67% versus $M_{science}$ = 81.61%, $SD_{science}$ = 5.66%; t (32) = 10.78, p < .001, d = 2.52) and inconsistent conditions ($M_{religion}$ = 85.06%, $SD_{religion}$ = 9.68% versus $M_{science}$ = 57.61%, $SD_{science}$ = 7.82%; t (32) = 13.77, p < .001, d = 3.12). The interaction entailed the size of the effect for consistency being more than twice as large for the science items ($M_{consistent}$ = 81.61%, $SD_{consistent}$ = 5.66%; versus $M_{inconsistent}$ = 57.61%, $SD_{inconsistent}$ = 7.82%; t (32) = 19.18, p < .001, d = 3.52) than for the religion items ($M_{consistent}$ = 97.19%, $SD_{consistent}$ = 6.67%; versus $M_{inconsistent}$ = 85.06%, $SD_{inconsistent}$ = 9.68%; t (32) = 8.83, p < .001, d = 1.46).

4.2.2 Sentence response time[15]

A paired-samples t-test on correct responses on the 20 modified religion items revealed no differences between the modified "consistent" and "inconsistent" items in response time ($M_{consistent}$ = 3770ms, $SD_{consistent}$ = 1178ms; versus $M_{inconsistent}$ = 3541ms, $SD_{inconsistent}$ = 992ms; t (32) = .96, p = n.s., d = .21). A paired-samples t-test on correct responses on the unmodified versions of these 20 items from the no time limit condition from Experiment 2 similarly revealed no difference between consistent and inconsistent items ($M_{consistent}$ = 3277ms, $SD_{consistent}$ = 693ms; versus $M_{inconsistent}$ = 3364ms, $SD_{inconsistent}$ = 563ms; t (31) = 1.21, p = n.s., d = .14). However, response time means on consistent versus inconsistent items were in the predicted direction for the unmodified items and in the opposite direction for the modified items. This pattern of findings, along with the response accuracy findings mentioned in section 4.2.1, suggests that the reference to a supernatural agent ("God") in the items used in this series of experiments is critical in causing the effect of consistency observed; the effect is not caused by any idiosyncratic features of inconsistent items, or other differences between consistent and inconsistent items.

Next, response time data on correct responses on the 28 unmodified religion items were subjected to a 2 (Domain; religion versus science) X 2 (Consistency; consistent versus inconsistent) repeated measures analysis of variance (ANOVA). The analysis revealed main effects of Domain (science versus religion; $F_{1, 32}$ = 7.44, p = .01, partial η^2 = .19) and Consistency (consistent versus inconsistent; $F_{1, 32}$ = 68.12, p < .001, partial η^2 = .68)

[15] A very small number of response time data points (<2%) were removed for being more than 3SD beyond the mean response time of any given participant.

qualified by an interaction between Domain and Consistency ($F_{1, 32} = 8.36$, $p < .01$, partial $\eta^2 = .21$).

Simple main effect analyses revealed that as in Experiments 1-2 participants were faster on the science items than on the religion items for consistent items ($M_{science} = 2991$ms, $SD_{science} = 632$ms versus $M_{religion} = 3284$ms, $SD_{religion} = 715$ms; $t(32) = 3.99$, $p < .001$, $d = .43$), and that there was no difference in response time on the science and religion items for inconsistent items ($M_{religion} = 3576$ms, $SD_{religion} = 791$ms versus $M_{science} = 3538$ms, $SD_{science} = 779$ms $t(32) = .49$, p = n.s., $d = .05$). The interaction entailed the size of the effect for consistency being nearly twice as large for the science items ($M_{consistent} = 2991$ms, $SD_{consistent} = 632$ms; versus $M_{inconsistent} = 3538$ms, $SD_{inconsistent} = 779$ms; $t(32) = 9.19$, $p < .001$, $d = .77$) than for the religion items ($M_{consistent} = 3284$ms, $SD_{consistent} = 715$ms; versus $M_{inconsistent} = 3576$ms, $SD_{inconsistent} = 791$ms; $t(32) = 3.91$, $p < .001$, $d = .39$).

5.0 General discussion

5.1 Support for the co-existence of inconsistent acquired beliefs and core intuitions

The goal of the current study was to investigate the hypothesis according to which acquired religious beliefs inconsistent with evolved core intuitions co-exist with, rather than replace, those intuitions in the minds of religious believers, using the case of Christian beliefs about God. The experiments reported here utilized a sentence verification task where participants were required to evaluate as true or false statements that were either consistent or inconsistent on core intuitions (or early-acquired) and (later) acquired religious (or scientific) beliefs. In Experiments 1-2 participants were less accurate and slower in evaluating inconsistent versus consistent statements, supporting our first prediction of an interference of

core intuitions on acquired religious beliefs (or early-acquired on later-acquired scientific beliefs), and thereby both the co-existence and co-option hypotheses.

By replacing the term "God" with "a priest" in a subset of the statements in Experiment 3 we controlled for the possibility that the findings in Experiments 1-2 were caused by low-level biases in the statements used, rather than by co-existence and interference of core intuitions and acquired religious beliefs.

Additionally, the experiments reported here utilized a novel methodology which controls for alternative interpretations whereby attributing person-like characteristics to God is caused by individuals not having acquired the relevant theology[16] or by intentional deviations from theology.

As discussed in the introduction, one alternative interpretation of previous findings is that it is not clear that all participants fully know which characteristics God is assumed to have. In this study all theological beliefs pertaining to the physical and psychological characteristics of God chosen to construct the statements used were ones that would be known to most if not all religious and non-religious Western adults. The very high

[16] This methodology also controlled for an alternative interpretation whereby some participants adhered to a theological doctrine which invalidates the use of this task altogether. For example, the doctrine of negative theology postulates that due to the transcendent essence of God it is only appropriate to describe God by negation: that is, it is only appropriate to describe what God is not, rather than what God is. According to this view any statement about, for instance, the beliefs God has (whether true or false) is false. In this task, however, consistent statements were a baseline against which inconsistent statements were evaluated; the conclusions made are in the difference between the two statement sets, not in their appropriateness with respect to one theological tradition or another. More so, the near perfect response accuracy for our consistent statements confirmed our assumption that the participants in our study adhered to standard theological doctrines.

performance on inconsistent religion statements and near perfect performance on consistent religion statement suggests that the theological beliefs chosen were indeed known to the participants tested.[17]

Additionally, an examination of response accuracies on individual religion statements (see the Appendix) shows that for the most part errors are relatively evenly distributed among the different statements; that is, there was a small but reliable probability that participants would make an error on any given statement regardless of the characteristic the statement pertained to. The alternative whereby participants did not know certain parts of the relevant theology predicts clustering of errors around only few statements, for example statements pertaining to characteristics of God talked about rarely, and/or ambiguously.

Another alternative interpretation of previous findings is that while representations of God co-opt the person concept abstract theological characteristics do replace default inferences about persons in those representations, but believers then intentionally deviate from theology by attributing person-like characteristics to God. The response time data in both the speeded and unspeeded instructions conditions of Experiment 1 and the no time limit condition of Experiment 2 control for this alternative interpretation of the response accuracy data: as predicted by the view whereby attributing person, rather than theological, characteristics to God is the intuitive default, consistent religion statements were responded to more quickly than inconsistent religion statements. If abstract theological characteristics

[17] Whether or not participants had sufficient access to the relevant scientific information across the domains tested in Shtulman & Valcarcel (2012) remains a potential concern, however, since both in their study and in our replication accuracy for some items was as low as would be expected from chance (e.g. less than 50% for inconsistent statements in the domains of 'evolution' and 'mechanics').

replace default inferences about persons, and person-like characteristics are only intentionally and consciously applied, then among correct responses – those responses where person-like characteristics would not have been applied – there should be no difference in response times between different statement kinds.

The response accuracy data in the time limit condition of Experiment 2 additionally control for this alternative. As predicted by the co-existence hypothesis, when participants were given the task with a time limit on responses they disproportionately made more errors on inconsistent versus consistent religion statements. The alternative predicts that under time limit conditions there should be an equal decrease in response accuracies in the different statement kinds.

5.2 Support for a mechanism that resolves conflicts between inconsistent beliefs

If core intuitions co-exist and interfere with inconsistent acquired beliefs then certain mechanisms should exist to resolve the interference or conflict created by tasks in which both representations are engaged (e.g. Kelemen & Rossett, 2009; Kelemen, Rottman, & Seston, 2012; Lindeman, Reikki & Hood, 2011; Svedholm & Lindeman, 2013). Therefore, a second prediction derived from the co-existence hypothesis which was tested here was that the effect of consistency on response accuracy will increase when participants are made less able to resolve conflicts between inconsistent beliefs.

In Experiment 2 a subset of participants was assigned to a speeded responding condition where a time limit was set on responses, and another subset received the task with no such time limit (unspeeded responding condition). The findings demonstrated that when participants are made to respond quickly they disproportionately make more errors on

inconsistent versus consistent religion statements. This supports the prediction that certain cognitive mechanisms resolve conflicts between inconsistent beliefs, and that when individuals are put under time pressure these mechanisms are less able to do so.

5.3 No correlations with executive inhibition

A third prediction derived from the co-existence hypothesis which was tested here was that executive inhibition, as indexed by the behavioral Stroop task, is the process that resolves the interference or conflict of core intuitions on inconsistent acquired beliefs. Previous studies are equivocal in their support for this proposal. For example, on the one hand Kelemen & Rossett (2009) found that performance with a subset of teleological statements was related to the behavioral Stroop task, with the task explaining unique variance in endorsement performance (*ibid*, p. 141), and similarly, Svedholm & Lindeman (2013) showed that a measure of ontological confusion (argued by these authors to be based on core intuitions, and to underlie paranormal belief) was strongly correlated with the behavioral Stroop task (albeit using a different configuration of task conditions and coding than those used by Kelemen & Rossett, 2009). On the other hand, Kelemen, Rottman, & Seston (2012) found no relationship between teleological thinking based errors and the behavioral Stroop task in their study of college students, professional scientists and community members (*ibid*, p. 1079).

The results reported here using task conditions and coding similar to those adopted by Kelemen and colleagues in Experiment 1, and Svedholm & Lindeman (2013) in Experiment 2, repeatedly failed to show a relationship between performances on the sentence verification task and inhibitory control, further muddying the pattern of results on this question. We think

it likely that this failure to find a correlation between performance on the sentence verification task and the behavioral Stroop task is due to the fact that on the one hand, executive inhibition itself can be measured in a variety of ways and that, more broadly, a range of different executive functions may jointly contribute to the coordination of different kinds of representations (e.g. Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). In future researchers might consider the use of a more rigorous battery of inhibitory and other executive functions measures to evaluate more precisely the ways in which executive functions (and specifically inhibition) might be involved in coordinating conflicting core intuitions and acquired beliefs.

5.4 No correlations with practice

A final prediction derived from the co-existence hypothesis which was tested here was that practice with acquired religious and scientific beliefs would attenuate interference from inconsistent intuitions. The failure to find such an effect in both the domains of religion and science in the current study parallels that of Shtulman & Valcarcel (2012) for science. In contrast, the investigation by Kelemen & Rosset (2009) of intuitive teleological thinking about natural phenomena in a similar population of university undergraduates did find an effect of science education (as indexed by questionnaires on geoscience and natural selection) on implicit teleological errors. Furthermore, the investigation by Goldberg & Thompson-Schill (2009) of animacy judgments in biology professors versus university undergraduates found such an effect, with the biology professors showing a smaller bias in preferentially ascribing animacy to animals than to plants (also see Kelemen, Rottman, & Seston, 2012).

Is variability in the strength of acquired beliefs, then, associated with variable susceptibility to interference from core intuitions? The studies by Kelemen and colleagues and the study by Goldberg & Thompson-Schill (2009) all used more nuanced indices of variability in acquired beliefs than the ones used in the current study (or in the study by Shtulman & Valcarcel, 2012); however, their findings are equivocal. Goldberg & Thompson-Schill (2009) find greater animacy judgments for animals versus plants in biology professors versus undergraduates, but not for other categories less studied in biology departments (e.g. nonliving natural kinds versus artifacts). This suggests that the strength of acquired beliefs, rather than another difference between the groups examined, can explain the difference in animacy judgements. On the other hand, while Kelemen, Rottman, & Seston (2012) demonstrated that the tendency to make teleological errors was greater in undergraduates than in science professors, they did not find such a difference between science and humanities professors. In future researchers might consider the use of more nuanced indexes and combinations of multiple indexes (both questionnaires as in Kelemen & Rosset, 2009, and populations that more strongly differ in education or practice as in Kelemen, Rottman, & Seston, 2012, and Goldberg & Thompson-Schill, 2009) to further investigate questions pertaining to the effects of practice on acquired beliefs which are inconsistent with intuitions.

5.5 Future directions

5.5.1 What concept is being co-opted?

In this study we assumed that the person concept is co-opted to form a representation of God. However, other concepts could be co-opted for this function. For example, Shtulman & Lindeman (2015) argue that a disembodied mind concept, rather than a person concept, is

being co-opted in forming representations of God and other supernatural entities; Cohen (2007, 2008) proposed that among believers of the Afro-Brazilian syncretic cult Candomblé some beliefs about possessing spirits co-opt mechanisms for reasoning about pathogens. More so, it is possible that there are individual differences in the core concept that is co-opted for forming a representation of a given supernatural entity. Future studies are needed to further investigate which concepts are being co-opted to form representations of various supernatural entities in both mainstream Christianity and other religious groups.

5.5.2 "On-the-ground" phenomena

The co-existence hypothesis explains two "on-the-ground" phenomena which might otherwise seem mysterious to social scientists and humanists (see Sperber, 1985, for an early discussion): a discrepancy between the different beliefs a given individual reports regarding the same religious or scientific phenomenon, and a discrepancy between a given individual's reported beliefs on the one hand, and behavior on the other (also see Slone, 2004). For example, Christian religious believers often simultaneously describe God as both person-like (e.g. loving, fallible) and abstract (not able to be described with human emotion terms and infallible), and although God is theologically believed to be all-knowing Christian religious believers nonetheless tell Him their wishes (the contradiction, of course, is that if God is all-knowing believers do not need to tell Him anything – He already knows everything). Future studies are needed to further investigate the conditions in which different beliefs are verbalized "on-the-ground", and the role of core intuitions and acquired beliefs in regulating behavior.

References

Baillargeon, R., Scott, R. M., & Bian, L. (In Press). Psychological reasoning in infancy. Annual Review of Psychology, Vol. 67.

Barrett, J., & Keil, F. (1996). Conceptualizing a non-natural entity: Anthropomorphism in god concepts. *Cognitive Psychology*, *31*, 219-247.

Barrett, J. (1998). Cognitive constraints on Hindu concepts of the divine. *Journal for the Scientific Study of Religion*, *37*(4), 608–619.

Barrett, J. (1999). Theological correctness: cognitive constraint and the study of religion. *Method & Theory in the Study of Religion*, *11*, 325-339.

Barrett, J. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences*, *4*(1), 29–34.

Baumard, N., & Boyer, P. (2013). Religious beliefs as reflective elaborations on intuitions: A modified dual-process model. *Current Directions in Psychological Science*, *22*(4), 295-300.

Bloom, P., & Skolnick-Weisberg, D. (2007). Childhood origins of adult resistance to science. *Science*, *316*(5827), 996-997.

Boyer, P. (1994a). Cognitive constraints on cultural representations: Natural ontologies and religious ideas. In L. A. Hirschfeld & S. Gelman (Eds.), Mapping the mind: Domain-specificity in culture and cognition (pp. 391–411). New York: Cambridge University Press.

Boyer, P. (1994b). The naturalness of religious ideas: A cog- nitive theory of religion. Berkeley: University of California Press.

Boyer, P. (2001). *Religion explained: The evolutionary origins of religious thought*. New York: Basic Books.

Boyer, P., & Ramble, C. (2001). Cognitive templates for religious concepts : cross-cultural evidence for recall of counter-intuitive representations. *Cognitive Science*, *25*, 535–564.

Broadway, J., & Engle, R. (2010). Validating running memory span: Measurement of working memory capacity and links with fluid intelligence. *Behavior Research Methods*, *42*(2), 563-570.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Cohen, E. (2007). *The mind possessed: The cognition of spirit possession in an Afro-Brazilian religious tradition*. New York: Oxford University Press.

Cohen, E. (2008). What is spirit possession? Defining, comparing, and explaining two possession forms. *Ethnos*, *73*(1), 1–25.

Epley, N., Converse, B. A., Delbosc, A., Monteleone, G. A., & Cacioppo, J. T. (2009). Believers' estimates of God's beliefs are more egocentric than estimates of other people's beliefs. *Proceedings of the National Academy of Sciences*, *16*(51), 21533–21538.

Evans, J. (2003). In two minds: dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*(10), 454–459.

Evans, J. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–78.

Evans, J., & Stanovich, K. (2013). Dual-Process Theories of Higher Cognition: Advancing the Debate. *Perspectives on Psychological Science*, *8*(3), 223–241.

Goldberg, R., & Thompson-Schill, S. (2009). Developmental "roots" in mature biological knowledge. *Psychological Science*, *20*(4), 480-487.

Granqvist, P., & Kirkpatrick, L. (2008). Attachment and religious representations and behavior. In J. Cassidy & P. R. Shaver (Eds.), *Handbook of attachment: Theory, research, and clinical applications* (2nd ed.). New York: Guilford.

Gray, H., Gray, K., & Wegner, D. (2007). Dimensions of mind perception. *Science*, *315*(5812), 619.

Hyde, K. (1990). *Religion in childhood and adolescence: A comprehensive review of the research*. Birmingham, AL: Religious Education Press.

Ioannidis, J. (2005). Why most published research findings are false. *PLOS Medicine*, *2*(8), 696-701.

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

Kelemen, D., Rottman, J., & Seston, R. (2012). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, *142*(4), 1074-1083.

Kirkpatrick, L. (2005). *Attachment, evolution, and the psychology of religion*. New York: Guilford.

Lane, J., Wellmen, H., Evans, E. (2012). Sociocultural input facilitates children's developing understanding of extraordinary minds. *Child Development*, *8*(3), 1007-1021.

Lane, J. D., Wellman, H. M., & Evans, E. M. (2014). Approaching an understanding of omniscience from the preschool years to early adulthood. *Developmental Psychology*, *50*(10), 2380–2392.

Legare, C., & Gelman, S. (2008). Bewitchment, biology, or both: The co-existence of natural and supernatural explanatory frameworks across development. *Cognitive Science*, *32*(4), 607-642.

Lindeman, M., & Aarnio, K. (2007). Superstitious, magical, and paranormal beliefs: An integrative model. *Journal of Research in Personality*, *41*(4), 731-744.

Lindeman, M., Reikki, T., & Hood, B. (2011). Is weaker inhibition associated with supernatural beliefs? *Journal of Cognition and Culture, 11,* 231–239.

Luhrmann, T. (2012). *When God talks back: Understanding the American Evangelical relationship with God*. New York: Random House, Inc.

Mercier, H., & Sperber, D. (2009). Intuitive and reflective inferences. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 147–170). New York: Oxford University Press.

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: a latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Nosek, B., & Lakens, D. (2014). Registered reports: A method to increase the credibility of published results. *Social Psychology*, *43*, 137-141.

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, *28*(3), 164–171.

Shtulman, A. (2006). Qualitative differences between naïve and scientific theories of evolution. *Cognitive Psychology*, *52*(2), 170-194.

Shtulman, A. (2008). Variation in the anthropomorphization of supernatural beings and its

implications for cognitive theories of religion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1123-1138.

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, *124*, 209-215.

Shtulman, A., & Lindeman, M. (2015). Attributes of God: Conceptual Foundations of a Foundational Belief. *Cognitive Science*. Advance online publication.

Slone, J. (2004). *Theological incorrectness: Why religious people believe what they shouldn't*. New York: Oxford University Press.

Sperber, D. (1985). *On anthropological knowledge: Three essays*. Cambridge University Press.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford: Blackwell Publishers Ltd.

Sperber, D. (1997). Intuitive and reflective beliefs. *Mind and Language*, *12*(1), 67–83.

Sperber, D. (Ed.). (2000). *Metarepresentations: A multidisciplinary perspective*. New York: Oxford University Press.

Svedholm, A., & 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.

Appendix

Religion

				M Proportion Correct		M Response Time (ms)	
Category	**Intuition**	**Reflection**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
T1	T	T	Many people around the world pray, and day after day God listens to their prayers toward Him.	0.97	0.99	2523	2862
	F	F	Many people around the world pray, and God doesn't listen to any of their prayers toward Him.	1.00		3202	
	T	F	Many people around the world pray, and in one day God listens to some of their prayers.	0.77	0.77	1824	2226
	F	T	Many people around the world pray, and in one day God listens to all of their prayers.	0.77		2629	
T2	T	T	When I say the Lord's Prayer quietly at my home, God hears it very well.	0.95	0.95	4895	4961
	F	F	When I say the Lord's Prayer quietly at my home, God has difficulty hearing it.	0.95		5027	
	T	F	When I say the Lord's Prayer on a busy street, God has difficulty hearing it.	0.90	0.91	3607	2818
	F	T	When I say the Lord's Prayer on a busy street, God hears it very well.	0.92		2030	
T3	T	T	God can be present at my church and at other churches as well.	1.00	1.00	3300	3528
	F	F	God is never present at my church, nor is He present anywhere else.	1.00		3757	
	T	F	Sometimes God is at my church, and sometimes He is at other churches.	0.82	0.89	2675	3393
	F	T	God is at all times both at my church and at other churches.	0.95		4111	
T4	T	T	God listens to people's prayers all the time.	0.90	0.95	3475	3218
	F	F	God doesn't listen to people's prayers at all.	1.00		2962	
	T	F	God listens to people's prayers one by one.	0.49	0.64	2275	2148
	F	T	God listens to people's prayers all at once.	0.79		2021	
T5	T	T	God can act on the world in various ways.	1.00	0.98	2929	2973
	F	F	God can't act on the world in any way.	0.95		3016	
	T	F	God needs a body to act on the world.	0.98	0.98	3808	3516
	F	T	God can act on the world without a body.	0.98		3225	
T6	T	T	God can see people's actions.	1.00	1.00	1886	2063
	F	F	God can't see people's actions.	1.00		2240	
	T	F	God needs eyes to see.	0.95	0.95	3217	2737
	F	T	God can see without eyes.	0.95		2256	
T7	T	T	God can see what it is I am doing.	0.98	0.98	3125	2876
	F	F	God can never see what it is I'm doing.	0.98		2627	
	T	F	God needs to look to see what I'm doing.	0.90	0.95	2346	2542
	F	T	God always sees what it is I am doing.	1.00		2738	

T8	T	T	God has beliefs that are true.	0.97	0.99	2621	2867
	F	F	All beliefs God has are false.	1.00		3114	
	T	F	God has beliefs that are false.	0.98	0.95	1881	3028
	F	T	All beliefs God has are true.	0.92		4174	
T9	T	T	God knows of various things that happened in the past.	0.97	0.99	2830	2857
	F	F	God doesn't know of things that happened in the past.	1.00		2883	
	T	F	God knows only some things that happened in the past.	0.95	0.98	2985	2806
	F	T	God knows of all things that happened in the past.	1.00		2627	
T10	T	T	God hears my prayers and the prayers of other people well.	1.00	1.00	2992	3205
	F	F	God hears my prayers worse than the prayers of other people.	1.00		3418	
	T	F	God hears my prayers better than the prayers of other people.	0.95	0.82	3042	3484
	F	T	God doesn't hear my prayers better than those of other people.	0.69		3927	
T11	T	T	God can hear what I say out loud.	0.97	0.97	2707	2648
	F	F	God can't hear what I say out loud.	0.98		2588	
	T	F	God can't hear what I say to myself.	0.95	0.96	2579	3072
	F	T	God can hear what I say to myself.	0.97		3564	
T12	T	T	God knows what I want and what I pray for.	1.00	1.00	2731	3354
	F	F	Even if I pray for it, God won't know what I want.	1.00		3976	
	T	F	God won't know what I want unless I pray for it.	0.92	0.91	4136	3708
	F	T	God will know what I want even if I don't pray for it.	0.90		3280	

Astronomy

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Planet	T	T	Planets are more massive than moons.	0.82	0.85	3366	3337
	F	F	Asteroids are more massive than planets.	0.89		3309	
	T	F	Planets are more massive than stars.	0.68	0.67	2323	3119
	F	T	Stars are more massive than planets.	0.66		3915	
Star	T	T	The sun produces light.	1.00	0.84	2236	2215
	F	F	The sun produces sound.	0.68		2195	
	T	F	The moon produces light.	0.68	0.57	2186	2760
	F	T	The sun produces gravity	0.45		3335	
Solar system	T	T	The moon revolves around the earth.	0.89	0.92	3977	3327
	F	F	The sun revolves around the moon.	0.95		2677	
	T	F	The sun revolves around the earth.	0.90	0.92	2881	2893
	F	T	The earth revolves around the sun.	0.93		2905	
Lunar phase	T	T	Phases of the moon are caused by changes in illumination.	0.35	0.64	5012	4893
	F	F	Phases of the moon are caused by clouds.	0.93		4773	
	T	F	Phases of the moon are caused by the earth's shadow.	0.47	0.62	4617	4119
	F	T	Phases of the moon are caused by the moon's orbit.	0.77		3621	
Season	T	T	A change in overall sunlight causes the seasons.	0.38	0.36	5304	4759
	F	F	The earth's rotation causes the seasons.	0.34		4213	
	T	F	The earth's distance from the sun causes the seasons.	0.48	0.67	4857	4380
	F	T	The earth's tilt causes the seasons.	0.86		3903	

Evolution

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Common Ancestry	T	T	Humans are more closely related to apes than monkeys.	0.61	0.66	4027	4309
	F	F	Whales are more closely related to plants than fish.	0.72		4592	
	T	F	Apes are more closely related to monkeys than humans.	0.23	0.46	5490	4658
	F	T	Whales are more closely related to humans than fish.	0.69		3825	
Phylogeny	T	T	Humans are descended from tree-dwelling creatures.	0.22	0.59	3979	3525
	F	F	Humans are descended from plants.	0.96		3071	
	T	F	Humans are descended from chimpanzees.	0.77	0.54	2681	3057
	F	T	Humans are descended from sea-dwelling creatures.	0.31		3433	
Variation	T	T	Evolution requires differential survival.	0.88	0.85	4209	3993
	F	F	Evolution requires climate stability.	0.82		3776	
	T	F	Evolution requires long periods of time.	0.26	0.56	2081	2936
	F	T	Evolution requires within-species variation.	0.86		3791	
Selection	T	T	Most organisms are adapted to their environment.	0.93	0.65	3503	3438
	F	F	Most organisms live in temperate climates.	0.36		3373	
	T	F	Most organisms have plenty to eat.	0.39	0.30	3504	3955
	F	T	Most organisms die before leaving offspring.	0.22		4406	
Adaptation	T	T	Biological species evolve.	0.95	0.93	2748	2614
	F	F	Inanimate objects evolve.	0.90		2479	
	T	F	Individual organisms evolve.	0.27	0.31	3724	2987
	F	T	Computer viruses evolve.	0.35		2251	

Fractions

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Addition	T	T	1/3 plus 1/3 is 2/3.	0.95	0.97	2815	2973
	F	F	1/3 plus 1/3 is 1/3.	0.98		3131	
	T	F	2/3 plus 2/4 is 2/7.	0.91	0.86	7021	5846
	F	T	1/10 plus 1/10 is 1/5.	0.82		4671	
Division	T	T	4 divided by 2/1 is 2.	0.93	0.95	4199	4021
	F	F	8 divided by 2/1 is 3.	0.98		3842	
	T	F	8 divided by 1/2 is 4.	0.83	0.71	5243	4346
	F	T	4 divided by 1/2 is 8.	0.59		3449	
Conversion	T	T	All decimals are expressible as fractions.	0.89	0.66	3293	4151
	F	F	All fractions are expressible as integers.	0.43		5009	
	T	F	All fractions are expressible as decimals.	0.02	0.48	6058	5315
	F	T	All integers are expressible as fractions.	0.93		4571	
Ordering	T	T	12/13 is greater than 1/13.	0.93	0.88	3467	3653
	F	F	1/17 is greater than 16/17.	0.84		3839	
	T	F	1/17 is greater than 1/9.	0.89	0.92	3017	3252
	F	T	1/13 is greater than 1/30.	0.96		3486	
Infinite density	T	T	There are numbers between 1 and 3.	0.89	0.87	2161	2909
	F	F	There are numbers between 10 and 10.	0.86		3657	
	T	F	There are numbers between 1/10 and 10/100.	0.50	0.70	5940	4552
	F	T	There are numbers between 1/3 and 1/2.	0.91		3163	

Genetics

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Heritalibty	T	T	Hair color is heritable.	1.00	0.98	2257	2172
	F	F	Pierced ears are heritable.	0.95		2086	
	T	F	Immunity to chickenpox is heritable.	0.73	0.71	3120	2931
	F	T	Intelligence is heritable.	0.68		2742	
Chromosome	T	T	Humans have more chromosomes than mosquitoes.	0.70	0.79	3382	3198
	F	F	Men have more chromosomes than women.	0.89		3014	
	T	F	Adults have more chromosomes than children.	0.91	0.65	3032	3898
	F	T	Skin cells have more chromosomes than egg cells.	0.39		4765	
Dominance	T	T	Two brown-eyed parents could have a brown-eyed baby.	1.00	0.63	2871	2757
	F	F	Two blue-eyed parents could have a brown-eyed baby.	0.26		2644	
	T	F	A blue-eyed baby must have two blue-eyed parents.	0.84	0.83	3478	3346
	F	T	Two brown-eyed parents could have a blue-eyed baby.	0.81		3214	
Expression	T	T	Genes that code for eye color can be found in the eye.	0.69	0.67	3802	3683
	F	F	Genes that code for eye color can be found in hair.	0.64		3563	
	T	F	Genes that code for eye color can be found in tears.	0.87	0.67	3251	3450
	F	T	Genes that code for eye color can be found in the liver.	0.48		3650	
Mutation	T	T	Radiation can change one's genome.	0.68	0.80	3294	3006
	F	F	Sunscreen can change one's genome.	0.91		2718	
	T	F	Exercise can change one's genome	0.84	0.62	2508	3050
	F	T	Viruses can change one's genome.	0.41		3591	

Germs

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Contagion	T	T	Being sneezed on can make a person sick.	0.89	0.93	2621	2631
	F	F	Being happy can make a person sick.	0.98		2640	
	T	F	Being cold can make a person sick.	0.30	0.60	2474	2873
	F	T	Being depressed can make a person sick.	0.91		3273	
Contamination	T	T	Rotting meat contains germs.	0.86	0.90	2642	2529
	F	F	Sunshine contains germs.	0.93		2415	
	T	F	Urine contains germs.	0.23	0.56	2204	2415
	F	T	Dish sponges contain germs.	0.89		2627	
Infection	T	T	Germs enter the body through cuts.	0.91	0.91	2164	2418
	F	F	Germs enter the body through the hair.	0.91		2671	
	T	F	Germs enter the body through the skin.	0.32	0.47	2633	2662
	F	T	Germs enter the body through the eyes.	0.61		2692	
Sterilization	T	T	Alcohol kills germs.	0.86	0.93	1989	2010
	F	F	Kindness kills germs.	1.00		2031	
	T	F	Water kills germs.	0.91	0.90	1950	2102
	F	T	Heat kills germs.	0.89		2255	
Microbe	T	T	Germs have a shape.	0.82	0.88	2404	1969
	F	F	Germs have feelings.	0.95		1534	
	T	F	Germs have an odor.	0.73	0.74	2311	2619
	F	T	Germs have DNA.	0.75		2928	

Matter

Category	Naive theory	Scientific theory	Statement	M Proportion Correct		M Response Time (ms)	
				Per statement	Per type	Per statement	Per type
Mass	T	T	Rocks are composed of matter.	0.98	0.97	2302	2699
	F	F	Numbers are composed of matter.	0.95		3097	
	T	F	Fire is composed of matter.	0.38	0.65	2889	2745
	F	T	Air is composed of matter.	0.91		2602	
Weight	T	T	A liter of water weighs more than a liter of air.	0.68	0.83	4652	4925
	F	F	A pound of steal weighs more than a ton of steal.	0.98		5199	
	T	F	A pound of steal weighs more than a pound of feathers.	0.70	0.53	3399	4484
	F	T	A liter of water weighs more than a liter of ice.	0.35		5570	
Density	T	T	Steal is denser than foam.	0.93	0.93	3259	2964
	F	F	Foam is denser than brick.	0.93		2670	
	T	F	Ice is denser than water.	0.55	0.49	4313	4579
	F	T	A cold penny is denser than a hot penny.	0.43		4846	
Divisibility	T	T	A log can be cut in half.	0.98	0.94	2433	2543
	F	F	An idea can be cut in half.	0.91		2653	
	T	F	A shadow can be cut in half.	0.95	0.95	2095	2575
	F	T	A grain of sand can be cut in half.	0.95		3054	
Atom	T	T	Atoms are the constituents of matter.	0.96	0.94	3548	3229
	F	F	Atoms are visible to the naked eye.	0.93		2910	
	T	F	Atoms are weightless.	0.93	0.74	2653	2769
	F	T	Atoms are mostly empty space.	0.54		2886	

Mechanics

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Force	T	T	Forces can be strong or weak.	0.98	0.87	2001	2030
	F	F	Forces can be male or female.	0.76		2060	
	T	F	Forces can be straight or curved.	0.43	0.65	5298	3956
	F	T	Forces can be balanced or unbalanced.	0.86		2614	
Velocity	T	T	A moving bullet loses speed.	0.87	0.88	3153	3248
	F	F	A moving bullet loses weight.	0.89		3342	
	T	F	A moving bullet loses force.	0.22	0.44	2500	2856
	F	T	A moving bullet loses height.	0.66		3212	
Acceleration	T	T	Constant acceleration requires constant force.	0.91	0.80	3950	4920
	F	F	Constant speed requires constant acceleration.	0.69		5889	
	T	F	Constant motion requires constant force.	0.49	0.42	4721	4682
	F	T	Constant force can yield constant rest.	0.34		4642	
Momentum	T	T	Momentum can be transferred from one object to another.	0.89	0.78	3636	3856
	F	F	Color can be transferred from one object to another.	0.68		4075	
	T	F	Forces can be transferred from one object to another.	0.09	0.53	4116	3638
	F	T	Energy can be transferred from one object to another.	0.98		3161	
Gravity	T	T	Anvils fall through air faster than feathers.	0.69	0.84	3218	3502
	F	F	Bright objects fall through air faster than dark objects.	1.00		3787	
	T	F	Heavy balls fall through air faster than light balls.	0.52	0.56	9594	7635
	F	T	Pointy objects fall through air faster than flat objects.	0.59		5676	

Physiology

Category	Naive theory	Scientific theory	Statement	M Proportion Correct		M Response Time (ms)	
				Per statement	Per type	Per statement	Per type
Life	T	T	Fish are alive.	1.00	0.93	1357	1644
	F	F	Rocks are alive.	0.87		1931	
	T	F	The sun is alive.	0.84	0.85	2097	1788
	F	T	Coral is alive.	0.87		1479	
Death	T	T	Turtles can die.	0.98	0.99	1255	1520
	F	F	Screwdrivers can die.	1.00		1786	
	T	F	Clouds can die.	0.96	0.91	1573	1660
	F	T	Mushrooms can die.	0.86		1747	
Reproduction	T	T	Tigers can reproduce.	1.00	0.99	1551	1610
	F	F	Chairs can reproduce.	0.98		1669	
	T	F	Caterpillars can reproduce.	0.18	0.51	3216	2580
	F	T	Ferns can reproduce.	0.84		1943	
Metabolism	T	T	People turn food into energy.	0.86	0.92	2356	2454
	F	F	Rocks turn food into energy.	0.98		2552	
	T	F	Plants turn food into energy.	0.27	0.42	3778	4119
	F	T	Bacteria turn food into energy.	0.57		4461	
Kinship	T	T	A baby can be a niece or a nephew.	1.00	0.86	3309	3434
	F	F	A baby can be a mother or a father.	0.73		3559	
	T	F	A baby can be identical to an older sibling.	0.80	0.81	3539	3594
	F	T	A baby can be an uncle or an aunt.	0.82		3650	

Thermodynamics

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Heat	T	T	The sun has heat.	1.00	0.96	1767	1958
	F	F	Gravity has heat.	0.91		2148	
	T	F	Atoms have heat.	0.43	0.56	2787	2952
	F	T	Ice has heat.	0.68		3117	
Heat source	T	T	Ovens produce heat.	0.98	0.93	2686	2585
	F	F	Rain produces heat.	0.89		2484	
	T	F	Coats produce heat.	0.84	0.73	2971	3210
	F	T	Pressure produces heat.	0.61		3448	
Heat transfer	T	T	Water transfers heat to ice.	0.75	0.78	5085	4816
	F	F	Cold objects transfer heat to warm objects	0.82		4546	
	T	F	Water transfers heat to steam.	0.66	0.74	5423	5500
	F	T	Cold objects transfer heat to even colder objects.	0.82		5578	
Temperature	T	T	Ice has a lower temperature than water.	0.78	0.85	3660	3962
	F	F	Steam has a lower temperature than ice.	0.93		4263	
	T	F	Two cups of ice have a lower temperature than one.	0.87	0.61	6152	5791
	F	T	Boiling water has a lower temperature than steam.	0.36		5431	
Thermal expansion	T	T	Heat increases an object's temperature.	0.98	0.89	3302	3335
	F	F	Heat increases an object's color.	0.80		3367	
	T	F	Heat increases an object's weight.	0.82	0.66	3233	3375
	F	T	Heat increases an object's size.	0.50		3518	

Waves

				M Proportion Correct		M Response Time (ms)	
Category	**Naive theory**	**Scientific theory**	**Statement**	**Per statement**	**Per type**	**Per statement**	**Per type**
Light	T	T	Rainbows contain all colors of light.	0.73	0.77	3097	3359
	F	F	Shadows contain all colors of light.	0.82		3622	
	T	F	Black objects reflect all colors of light.	0.71	0.76	3727	3391
	F	T	White light contains all colors of light.	0.82		3056	
Color	T	T	Red objects reflect red light.	0.72	0.71	3644	3684
	F	F	Red objects reflect blue light.	0.70		3724	
	T	F	Red objects absorb red light.	0.70	0.62	2840	3066
	F	T	Red objects absorb blue light.	0.55		3292	
Sound	T	T	Sounds can be loud or quiet.	0.98	0.95	2699	2795
	F	F	Sounds can be dead or alive.	0.93		2891	
	T	F	Sounds can be near or far.	0.07	0.53	6595	5205
	F	T	Sounds can be direct or reflected.	1.00		3815	
Propagation	T	T	Sound travels through air.	0.95	0.70	1823	2193
	F	F	Sound travels through foam.	0.45		2564	
	T	F	Sound travels through a vacuum.	0.66	0.68	3225	3251
	F	T	Sound travels through metal.	0.71		3277	
Reflection	T	T	Mirrors reflect light.	0.98	0.80	1827	2063
	F	F	Foam reflects sound.	0.62		2298	
	T	F	Prisms reflect light.	0.18	0.24	4647	3880
	F	T	Mirrors reflect sound.	0.30		3113	