CONDITIONAL REFLECTION AND THE RELIGION REFLECTION SCALE: HOW FAMILIARITY WITH DIFFERENT DOMAINS INFLUENCES COGNITIVE REFLECTION

By

Jacob L. Strimaitis

Ralph W. Hood Jr.
Professor of Psychology
(Chair)

Michael D. Biderman
Professor of Psychology
(Committee Member)

David F. Ross
Professor of Psychology
(Committee Member)

CONDITIONAL REFLECTION AND THE RELIGION REFLECTION SCALE: HOW FAMILIARITY WITH DIFFERENT DOMAINS INFLUENCES COGNITIVE REFLECTION

By

Jacob L. Strimaitis

A Thesis Submitted to the Faculty of the University of
Tennessee at Chattanooga in Partial
Fulfillment of the Requirements of the Degree
Master of Science: Psychology

The University of Tennessee at Chattanooga
Chattanooga, Tennessee

May 2018

ABSTRACT

A host of research has shown that those who are nonreligious tend to outscore those who are religious on measures of intelligence and cognitive reflection. The prevailing explanation for this phenomenon is that nonreligious people are more adept at overriding their initial intuitions, thus leading to their increased skepticism of intuitive supernatural beliefs. However, the argument is raised in this investigation that the current measures of cognitive reflection in the field may not encapsulate the totality of dual-process thinking. Just as math students outperform upper-level psychology students on measures of cognitive reflection related to their discipline (Study 1), so, too, may religious individuals outperform nonreligious individuals on measures of cognitive reflection relating to their discipline of familiarity: religion (Study 2). While support for this second hypothesis was wanting, special considerations should be given to find measures of cognitive reflection that appeal to the unique experiences of specific groups in question.

Keywords: cognitive reflection, dual-process theory, religion

DEDICATION

This thesis is dedicated to five important individuals in my life. To my mom and dad, thank you for supporting me emotionally (and financially) throughout this two-year academic endeavor. To my brother, Luke, thank you for your counsel and being a constant reminder to me that academia is not all there is to life. To my fiancée, Shelley, thank you for your love, your intelligent insights, and your much-needed movie nights. Finally, and potentially most importantly, to my cat, Mr. Stubbs. You'll never know I dedicated this thesis to you, nor will you care, but I truly could not have made it this far without you.

ACKNOWLEDGEMENTS

I am forever in debt to my thesis committee, Dr. Ralph Hood, Michael Biderman, and David Ross; thank you for your time, your encouragement, and your patience with me. To my mentors, Dr. Ralph Hood (again), David Ross (again), Bryan Poole, and John Hearn: thank you for pointing me in the right direction at every twist and turn. To Dr. Lucas Van der Merwe: thank you for giving me access to your math students. To Dr. Amye Warren: thank you for reading and critiquing my thesis in its preliminary stages. To Dr. Gordon Pennycook: thank you for conversing with me and sharing with me your syllogisms. To my Hoodlum and honorary Hoodlum colleagues, Sally Swanson, Maggie Dempsey, Shane Littrell, Kaila Rogers, Aaron Eldridge, Cameron Mackey, Shannon Hyder, and Zoe Ross: you are all awesome and I appreciate everything you've done for me. To my good friend, Mitchell Torrence: thank you for your companionship and encouragement throughout this process. Finally, I want to thank the countless others who have invested their time and ideas in me. From the mundane to the extraordinary, I truly believe that every interaction with you has, in some way, influenced the way I think about and perceive the world, and for those reality checks, I am grateful.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BITCH, Black Intelligence Test of Cultural Homogeneity

CR, Cognitive Reflection

CRT, Cognitive Reflection Test

IQ, Intelligence Quotient

MI, Multiple Intelligence

PMP, Mathematical Prerequisites for Psychometrics

RRS, Religious Reflection Scale

UTC, University of Tennessee at Chattanooga

LIST OF SYMBOLS

ANOVA, Analysis of Variance

CI, Confidence Interval

F, F-statistic

M, Mean

N, Number

p, Significance Statistic

SD, Standard Deviation

t, T-statistic

CHAPTER I.

INTRODUCTION

> *Everybody is a genius. But if you judge a fish by its ability to climb a tree, it will live its whole life believing that it is stupid.* – Anonymous quote often attributed to Albert Einstein

Take a moment to imagine the life of a talented flutist preparing for a concert. As she absorbs herself in her piece in her practice room, she feels a sense of accomplishment. After hundreds of hours of practice, her fingers can now dance across the keys with ease. Her breath is in sync with the rhythm of the piece, and the music she creates comes to life in a way that engages her focus and enthralls her soul. She knows how to play this piece, and she is ready for her concert. After a long and satisfying practice session, she rewards her efforts with a nightly jog. As she makes her way past the neighborhood basketball court, she remembers the middle school days when she tried her hand at the sport, but failed miserably. She was certainly tall and physically fit enough to excel, but for some reason, even after tireless nights of tribulation, she could not work her way around the court as her other teammates could. It was not from a lack of motivation, because she frequently watched basketball on T.V. and had the internal drive to get better, nor was it from a lack of practice, because she used her allowance to take lessons from experienced coaches and spent many hours shooting hoops on her own. Rather, despite her potential, her motivation, and her effort, basketball was just a sport that she simply could not master. With a frustrated sigh, her focus returns to the jog and the Christmas music blaring in her

earbuds. She finds herself enjoying the music, not because she is capable of realizing the melodic progression or the complexities of the chords, but because the sounds stir her of their own accord. She loves music. This fact is why she pursued music professionally in the first place, but unlike basketball, she became good at playing the flute after practice. She was made to be a flutist. So, with a newfound solace, she passes the basketball court and the memories of her failures, knowing that she is on the right path.

If society were such that one's intellectual worth was determined by how well one played basketball, this flutist would have been considered stupid. Thankfully, society exists in a pluralistic state that recognizes that one's poor performance in one domain does not necessitate that one's performance in another, even related, domain will likewise be unsatisfactory. For example, even though the flutist had poor hand-eye coordination and spatial awareness when playing basketball, she still had superb dexterity when it came to playing the flute. Similarly, she had incredible gifts in musical performance but was incapable of dissecting music into its components by ear. Again, it is simply folly to assume that one's ability or performance in one domain will match their ability or performance within a different domain or within the same domain but with different circumstances or contexts. A note of caution to remember, which is the central theme of this thesis, is that people with diverse abilities, backgrounds, and interests can and do perform better on tasks that are directly relevant to their specific abilities, backgrounds, and interests. In other words, fish are quite mobile and agile, but only if they are directed to swim in the water, not if they are directed to climb a tree.

The specific topic of interest for this thesis is the cognitive science of religion. Namely, the notion that nonreligious individuals are more reflective thinkers than religious individuals is challenged. This potentially paradigm-shifting research is important, because there is an implicit

stigma among psychologists and scientists that religious individuals are intellectually inferior to atheists and agnostics in measures of intelligence, calculation, and analytical thinking (Nyborg, 2009; Pennycook, Ross, Koehler, & Fugelsang, 2016b; Shenhav, Rand, & Greene, 2012). While many studies show this intellectual disparity between the religious and nonreligious using current measures, and certainly the quality and quantity of evidence in favor of this phenomenon is robust, it is not the case that *all* relevant measures have been considered. Perhaps there is only one hypothetical measure of analytical thinking or one way of thinking that religious individuals do better on or are better at than their nonreligious counterparts, but even one is enough. The point is that such a finding would significantly shift the narrative from one of absolutism to one of nuance. The purpose of this thesis is to explore the extent to which this nuance is warranted by creating and evaluating one potential measure of analytical thinking within a domain that favors the abilities, backgrounds, and interests of religious individuals.

This introduction serves as a map to the imagination, guiding the reader through a story that explains the importance of the issue at hand. The tail of end the map, this paragraph, should serve as final directions for how to navigate through the rest of the thesis. Chapter II contains a detailed account of the literature and history of the topics and includes definitions of terms, like religion, disbelief, intelligence, cognitive reflection, and the content effect; examples of the intellectual differences between religious and nonreligious individuals; counterexamples of how context changes performance on cognitive measures; and a description of the gap in the literature that needs to be filled. Chapters III, IV, and V describe the method, results, and conclusions of Study 1 where the effect of context was tested with math-based questionnaires. A rationale for Study 2 is also included in this section. Chapters VI, VII, and VIII, in turn, describe the method, results, and conclusions of Study 2 where a novel scale, the Religious Reflection Scale, was

tested among a sample of religious and nonreligious individuals. Chapter IX is a final discussion section where all the pieces from the literature review and the conclusions of Studies 1 and 2 are synthesized into a cohesive summary. Suggestions for future research are also included in this chapter.

CHAPTER II

LITERATURE REVIEW

Before diving into the rationale for the current research, one must have a detailed understanding of the history and context of the current topic. Unfortunately, as is the case with every important subject, the background information is vast and can be daunting. To ease this inevitable discomfort, this chapter was crafted in a way that funnels this wealth of information into comprehensible and logically-organized segments. First, key terms and concepts are introduced. Next, the current paradigm between religious belief and cognitive ability is described with multiple examples. After that, alternative examples of the effect of context and content on decision making are reviewed. Finally, all the evidence is coalesced into a rationale section where the purpose of the current research studies is stated.

Key Terms and Concepts

Religion and the Disbelievers

Religion is extremely prevalent in today's society. Since the 1940's, roughly 90% of U.S. respondents have reported belief in a God, and since the early 1990s, about 50-60% of U.S. respondents would claim that religion is "very important" in their own lives (Gallup, 2017). Despite its ubiquity, there is considerable debate among philosophers and psychologists on what religion actually is (for an overview, see Hood, Hill, & Spilka, 2009). Is it a belief system? A

way of living? An association with a particular group or creed? An experience with the divine? All or none of the above? There is no settled, absolute definition of religion, but there are some recurring concepts people have in mind when they think of the word "religion" that deserve appreciation.

In a broad sense, religion can be seen as a search for meaning (Hood et al., 2009). Such a search will involve cognitions, motivations, and sociality. The cognitive search for meaning is one's attempts to comprehend what life is all about, including why certain events occur and how to best get along with others. Beliefs in deities or other transcendent entities act as solutions to the often chaotic nature of the world. The motivational search for meaning is about one's need to control or predict what will happen, even if the inevitable outcome is undesirable. Feelings of self-control or control in general are often enhanced through prayers, rituals, or faith. The social embeddedness of meaning is concerned about the need to belong. Religion does not occur in a vacuum, and it is often through connectivity and communality within institutions that individuals find purpose. Religion involves all three modes of meaning making.

Even with this breakdown, the concept of religion is still somewhat vague. For instance, one could claim that being a fan of a particular sports team could be considered a type of religion, as such a fan could find meaning in rooting for her team every week and justify that meaning through her thoughts ("my place is being a fan of this team"), her motivations ("I feel fulfilled by rooting for this team, and my contributions can help them win"), and her social relationships ("I have fellow fans who share my desires to see our team succeed"). Few readers would recognize such a scenario as being prototypical of religion, even though it would meet the minimum criteria. There must be something missing. One central aspect that is prevalent in most religious traditions, though, and that is not represented by the sports fan analogy, is the divine.

What separates religion from any casual meaning-making system, at least in most Western traditions, is the shared belief in the divine or the supernatural (Cragun, Hammer, & Nielsen, 2015; Hood et al., 2009). These supernatural entities are typically called god(s), angels, demons, and spirits, and they often influence circumstances in the world or provide moral guidance through sacred texts or prophets. The inclusion of the divine is important, because it commonly separates religion from nonreligion (Harrison, 2006). A sports fan could not be considered religious by virtue of the fact that she does not incorporate the supernatural in her fanhood. Granted, she could be considered superstitious. Just as a pigeon could be fooled into thinking there was a causal relationship between its erratic behavior and the presentation of food in Skinner's classic experiments on superstitions (Skinner, 1948), so, too, could the sports fan believe wearing her lucky socks to home games increases the chances of her favorite team winning. The forces that supposedly influence such situations could be considered transcendent, other worldly, or beyond the natural world, and by definition, such forces would be considered supernatural. However, the erroneous equivocation between fate or luck and the term "supernatural" muddies the waters. The supernatural, at least in this thesis, will be used in reference to entities outside of space and time that are well-established within common Western religious traditions (e.g., gods, angels, demons, miracles, heaven, hell, etc.), not in reference to transient superstitions or forces that would better fit within the confines of spirituality (which is discussed later).

Another hallmark of religion worthy of acknowledgement is ritual. Many religions involve commonly repeated practices, or rituals, and include acts such as prayer, meditation, incantations, attending religious services, reading religious texts, contributing money to religious causes, and observing religious holidays (Hood et al., 2009). These practices can occur in

isolation or among a collective body among other religious members, and they usually carry symbolic significance. For example, in some Christian traditions, members will partake of bread and wine in remembrance of the Last Supper taken by Jesus Christ and his disciples before his ensuing arrest, torture, and crucifixion. While this ritual may take different forms and have different symbolic meanings, it is well-established in the Christian faith (Woodhead, Fletcher, Kawanami, & Smith, 2002). Taken as a whole, religion involves beliefs (typically about the supernatural) and behaviors that are incorporated into a cognitive, motivational, and social search for meaning. This definition for religion may not be the most robust or the most familiar to most readers, and it is certainly not the only one (Harrison, 2006), but it will suffice for the purposes of this thesis insofar as it demonstrates that religion is foundational to peoples' lives and encompasses both beliefs and behaviors.

Now that religion has been (somewhat) defined, it would be sensible to describe a related, and just as important, demographic: the nonreligious. As a whole, those who are nonreligious do not subscribe to the beliefs and behaviors of any particular religious tradition. While this definition appears simple at the outset, it is really fairly complex. For instance, is someone who attends religious ceremonies of a particular tradition regularly but does not believe in the tenets of that tradition truly a nonreligious individual? Likewise, if an individual believes in the truthfulness of specific religious creeds but does not practice any of its rituals, is he or she also a nonreligious individual? The answers to these questions are unclear and depend largely on how terms are defined. The tricky delineation between belief and behavior is reminiscent of the familiar and decades-old debate between extrinsic and intrinsic religiosity (Allport & Ross, 1967), though for the sake of this thesis, the motivation behind one's religious orientation is not as important as the objective content of their orientations. The analogy, though, that one's

ritualistic actions are not the same as one's personal convictions, is what stands. Those who are nonreligious in terms of belief are not necessarily the same people who do not participate in religious activities, and vice-versa. My focus, however, is specifically on the subset of nonreligious individuals who do not believe religious tenets and who typically do not engage in other religious rituals: the disbelievers.

Those who do not believe the tenets of religion are disbelievers or nonbelievers (Norenzayan & Gervais, 2013), but as will be demonstrated in a moment, even this definition is not as simple as it appears. Because belief in a god is common among many religions, those who disbelieve are often misconstrued as atheists. However, atheism is a response to the singular issue of whether or not one believes that a god exists. To put it more plainly, an atheist is one who does not believe in god(s) or believes there are no gods (Cliteur, 2009), so by this definition, one could believe the tenets of a conventionally nontheistic religious tradition, like Buddhism, and still be an atheist. Similarly, one could claim to not *know* whether or not god(s) exist (i.e., identify as an agnostic) and still also be an atheist by disbelieving in said god(s), as atheism is in reference to belief and not knowledge. Finally, one could disbelieve many central themes of a religion and be a theist, as could be the case with the Christian who believes that a God exists but is not convinced hell exists or that belief in Jesus is a central criterion for salvation. Are all three aforementioned personas nonbelievers? Well, yes, in some sense, given that every single one of them would disbelieve at least *some* critical element of a typical religion, but it would be improper to claim that they all disbelieve religious claims to the same degree. Just as there are nuances to explore with the nonreligious demographic, there is a spectrum of disbelief to explore within the disbelieving demographic (Silver, 2013).

To make matters not only conceptually tangled but practically difficult as well, at least in terms of measurement, there is a stigma among religious groups against those who would identify as atheist (Doane & Elliott, 2014; Edgell, Gerteis, & Hartmann, 2006). The problem extends to disbelievers or other nonreligious individuals as well (Norenzayan & Gervais, 2013). In other words, even if there were conceptual clarity on what was meant by the terms atheist and disbeliever, the fact that people do not like identifying with these nonreligious affiliations makes measuring these populations troublesome. Despite theoretical complications and measurement shortcomings, though, these terms still have utility. For the sake of clarity in this thesis, the term nonreligious will henceforth be a descriptor of those who primarily do not subscribe to or affiliate with any religion in particular, though said individuals may not entirely distance themselves from *all* beliefs or practices. The term disbelievers will likewise be viewed as a subset of nonreligious individuals who do not believe religious tenets, though it typically refers to agnostics and/or atheists who reject belief in (or claim to not have knowledge of) the existence of god(s).

Spirituality is another term within the religion-nonreligion juncture that deserves consideration, but it is only briefly described here for the sake of saying what it is and why it will not be featured further in this thesis. To begin, spirituality has many definitions and is meaningful to believers and nonbelievers alike, but in short, spirituality is often regarded as one's personal quest for "the sacred" (Zinnbauer & Pargament, 2002; Zinnbauer et al., 1997). This journey may or may not explicitly involve the supernatural or the divine but it almost always entails personal feelings of interconnectedness, wholeness, and meaning in life (Cragun et al., 2015). While religion is usually tied to institutions or rituals, spirituality does not require such a framework and is more personal and subjective (Hood et al., 2009). It is certainly

prevalent in society too, as between 50-80% of individuals across different cultures consider themselves as either spiritual and not religious or equally religious and spiritual (Streib, Silver, Csöff, Keller, & Hood, 2011). While an interesting and conceptually-relevant topic, spirituality will not be explored further because of the limited scope of this investigation. I am primarily concerned about the supernatural religious beliefs that people have, and while there is considerable overlap between the two domains (i.e., a religious person and spiritual person may both profess belief in a Higher Power), I am not going to focus on whether or not those beliefs are tied to personal orientations towards the sacred. Such a discourse deserves another thesis in itself.

Intelligence

An exploration of cognitive abilities would be incomplete if it did not first include a short review of intelligence. The notion of intelligence, at least in its modern form as a measurable concept, first appeared between the 19th and 20th centuries with the work of Alfred Binet (Binet, Beaunis, des Bancels, Simon, & Piéron, 1903; Binet & Simon, 1916; Reisberg, 2013). Binet and his colleagues were directed by the French minister of public instruction to identify children who were not performing well in school and could benefit from special attention. In order to do so, they had to devise a measure that would distinguish the intellectual capacities of those who were performing as expected from those who were not. Such a measure had to include a wide range of tasks associated with cognitive functioning (e.g., copying a drawing, repeating back a string of digits, reasoning through math problems, etc.), as someone with strong intellectual prowess would theoretically be able to complete all or most of them (and not just get lucky with a couple). Likewise, they reasoned that someone who scored low on all the tasks would, by

comparison, be in need of remedial education. The raw score for each student's composite performance on all of these tasks was regarded as their "mental age". This score was compared to their "chronological age," and the ratio, or *quotient*, between a student's mental age and chronological age (multiplied by 100) was that child's intelligence quotient, or IQ. Scores around 100 would indicate higher intelligence, whereas scores significantly lower than or higher than 100 would indicate deficits or advancements in intellectual ability, respectively. This system of calculating IQ scores was meant to be intuitive and simple. For example, a 15-year-old child who received a mental age score of 20 (i.e., who's intellectual capacity was that of an average 20-year-old) would have an overall IQ score of approximately 133, and this score would be appraised as significantly above average. Since Binet's time, a number of social scientists have created a host of popular intelligence measures for children and adults alike (see Reisberg, 2013). They all theoretically tap into this concept of intelligence. However, what does intelligence actually mean?

Unfortunately, there is no one set definition for intelligence upon which every scientist would agree (Legg & Hutter, 2007). To one, it could mean an underlying capacity to think and solve problems, and to another, it could mean one's propensity to learn and adapt to new environments. One psychologist (perhaps glibly) wrote that "intelligence is what is measured in intelligence tests" (Boring, 1923, p. 35). Truly, there is a lack of consensus among the scientific community on what intelligence is. One comprehensive definition, though, that will suffice for the purposes of this thesis, was collected from over 50 psychologists. It follows:

> Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not nearly book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings – "catching on," "making sense" of things, or "figuring out" what to do (Gottfredson, 1997, p. 13).

Indeed, intelligence is a description of cognitive ability. In trying to determine what intelligence is, it may be worthwhile to also describe what it is not. For instance, intelligence, by the above definition, is not a physical skill; rather, it is a cognitive propensity to think and reason. Intelligence is also not specific to one domain of knowledge. One could think or reason about car mechanics and be demonstrating just as much intellectual ability as one who is solving complex calculus problems. Taken together, though, intelligence is cognitive fortitude or ability. It is a general term for how well one can think.

Intelligence is often regarded as hierarchical (Reisberg, 2013), with general intellectual ability (often notated as g) covering all mental abilities. In other words, among different specialized abilities, like language and quantitative reasoning, there is a general mental ability, g, that pervades. This underlying mental ability, while important, is actually not the central focus of this thesis. Rather, it is the subcomponents of intelligence. One model, called Sternberg's Triarchic Theory of Intelligence, delineates between three such subcomponents, or aspects, of intelligence: creative, practical, and analytical (Sternberg, Castejón, Prieto, Hautamäki, & Grigorenko, 2001). Creative intelligence involves being able to discover new ideas and solve novel problems. Practical intelligence involves being able to sensibly react and adapt to real-world problems. Analytical intelligence involves analyzing and evaluating ideas and information. The hierarchical nature of intelligence is addressed here, because from this last subcomponent, analytical intelligence, the concept of cognitive reflection emerges.

Cognitive Reflection and Dual-Process Theory

In terms of definitions, Sternberg's analytical intelligence is very nearly the same thing as cognitive reflection. Analytical intelligence refers to the ability to reason critically (Sternberg et

al., 2001), whereas cognitive reflection refers to the propensity to resist reporting the first response that comes to mind (Frederick, 2005). A key distinction to make here between the two concepts is that the former is about cognitive *ability* while the latter is about cognitive thinking *style* (Pennycook & Ross, 2016; Stanovich, 2012). With analytical intelligence (and also general intelligence), the important component is processing capacity. It is about the efficiency with which one can calculate the solution to a problem and is largely algorithmic. It is about ability. For example, a librarian with high analytical intelligence could be given a task of, say, rearranging the books on her bookshelf and would be able to categorize them systematically and efficiently by topic, color, and size. With cognitive reflection, the important component is disposition (Stanovich, 2012). In other words, it is about one's proclivity or willingness to engage in analytical thinking, depending on the situation or the individual. To use the example of the librarian again, she could potentially decide not organize the books by virtue of the fact that the task was too tedious. She would certainly have the right amount of cognitive ability to complete the task, but in this case, her cognitive reflection (evidenced by her dispositional impatience) would be too low to do so. This example illustrates the important fact that one's cognitive ability, or intelligence, is separate from one's cognitive style, or willingness to *use* their intelligence, and cognitive reflection is primarily about the latter.

There are many synonyms for cognitive reflection: analytical thinking, reflective thinking, deliberative processing, etc. These terms interchange, but they all have roots in dual-process theory (Kahneman, 2003). Dual-process theory has at its core the notion that decisions are made via two types of processing: Type 1 processing, which is characterized by fast, autonomous thinking and gut-based intuitions, and Type 2 processing, which is characterized by slower, more reflective thinking (Evans & Stanovich, 2013; Kahneman, 2003). Everyone uses

both types at various times. For instance, one could engage in Type 1 processing by intuitively deciding to order a burger at a fast food restaurant and then engage in Type 2 processing when deliberating over what car to purchase later that afternoon. Neither type is necessarily or wholly better than the other. When it comes to decision making, engaging in Type 2 processing leads to less bias and errors in reasoning (Pennycook, Fugelsang, & Koehler, 2015b), but it is not necessarily the best cognitive route to take in all circumstances. For example, it may be most wise to follow one's intuition to run away at the sound of an angry, growling dog just around the corner. Such a decision would not be fully reasonable, per se, as it would forego further preponderance of evidence (perhaps the dog was an "all bark and no bite" breed and could be easily threatened to submission with a show of intimidation), but it need not be rational to still be useful. The point, though, is that despite its expediency and utility, Type 1 processing pales in comparison to Type 2 processing when it comes to making decisions without bias or intuition. It would be helpful now to understand dual-process theory in greater detail and delve deeper into how these two types of thinking trap one into or allow one to overcome their bias-prone intuitive thinking.

The dual-process model of analytic engagement follows a three-stage pattern (see Figure 1). In the first stage, intuitive responses are generated from an initial problem or cue. Take, for instance, a problem posed as the square root of 4. An observer of this problem could intuitively think of "2" as a response. This response would probably be the quickest, and thus it would be designated IR_1. After generating one or multiple Type 1 intuitive responses, the second stage is induced: conflict monitoring. In this stage, conflict between Type 1 outputs is either detected or not detected. If no conflict is detected, either because there are no other outputs besides IR_1 or because conflict-detection "mindware" is not activated, then one moves onto the third stage,

called final response selection/generation, where IR_1 is selected as the final Type 2 response. In the case of the observer who only thought of 2 as the IR_1 to the square root of 4, 2 would be selected quickly and without much deliberation in this third and final stage. Such a person would also be deemed an intuitive thinker, as little or no reflective thought was used to make this decision. This is perhaps the scenario where the most prototypical biases, or errors in thinking, occur, given that alternative potential sources of error are not considered.

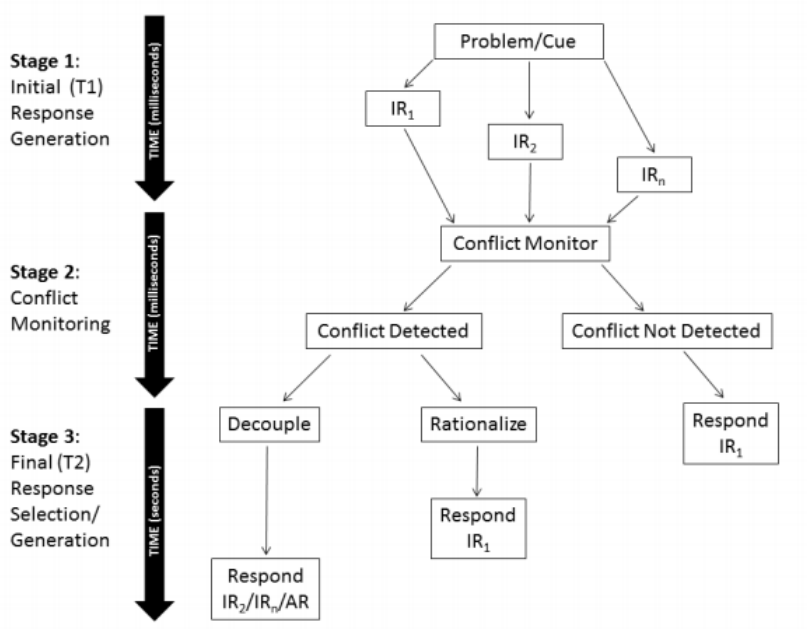

Figure 1 Three-stage dual-process model of analytic engagement (Pennycook, Fugelsang, & Koehler, 2015c)

In Figure 1, AR = alternative response; IR = initial response; T1 = Type 1 "intuitive" processing; and T2 = Type 2 "analytic" processing. IR's are numbered to reflect alternative speeds of generation. IR_1 is the most salient and fluent possible response. IR_n refers to the possibility of multiple, potentially competing, initial responses. IR_n refers to the possibility of an alternative response that is grounded in an initial response.

More Type 2, reflective processing occurs, however, when some conflict between the initial Type 1 outputs is detected. Take, for instance, the observer who initially thinks of 2 but then second-guesses his understanding of the square root symbol and momentarily confuses it with a "squared" symbol, thus leading him to think of 16 as IR_2. Such a detection of conflict from competing IR's in Stage 1 leads to greater Type 2 reflection in Stage 3 where a decision must be selected. It is then that either rationalization or decoupling occurs. Rationalization refers to Type 2 processing where secondary intuitive responses (IR_2 or IR_n) are only briefly considered and then rejected in favor of the initial, but sometimes erroneous, intuitive response (IR_1). Rationalization is often a source of bias, as the strongest intuition (i.e., the first) is favored above all other options. Despite the tendency towards bias, it is regarded as a Type 2 process, because the reasoning involved is still effortful. In the square root of 4 example, the observer uses rationalization when he already assumes 2 is the best answer, but spends some time contemplating 16 just to quickly reject it in light of the preferred answer. In this instance, the second initial response, IR_2 or 16, is not seriously considered and only exists as a means of bolstering belief in the first initial response, IR_1 or 2. Decoupling, which is most often associated with the concept of cognitive reflection, is a class of Type 2 processing that works to inhibit or override an intuitive response. It is contrasted with rationalization insofar as it is about *falsifying* an initial response, not *verifying* said initial response. The observer would use decoupling if he attempts to disprove 2 as the best answer, perhaps by trying to think of numbers that would also suffice as the square of 4 but would still not be 2. It would be in this moment that he would realize that 2 is, in fact, not the sole answer, and this attempt at falsifying the initial Type 1 response could yield an alternative response (AR): plus or minus 2. Indeed, +2 and -2, and not

just +2, is the correct solution to the square root of 4 problem, but it only came about through deeper cognitive reflection.

While there are complexities and nuances to explore with respect to the concepts associated with decision making, dual-process theory is quite simple. In essence, it proposes that there are two kinds of processes in response to problems: intuitive, Type 1 processes that pop into mind quickly and without deliberation and reflective, Type 2 processes that arise from a recognition of potential conflict between initial responses. Dual-process theory is highlighted in this thesis, because it is the theoretical framework behind decision making and is the heart behind one of the most popular measures of cognitive reflection, the Cognitive Reflection Test.

The Cognitive Reflection Test

The Cognitive Reflection Test (CRT) is a three-item questionnaire that measures one's willingness and ability to override their initial intuitions about a problem and reach a correct, analytically-derived answer (Frederick, 2005). The CRT measures cognitive reflection and predicts performance on measures of executive functioning, like heuristics-and-biases tasks and belief bias syllogistic reasoning tasks, much better than measures of intelligence (Toplak, West, & Stanovich, 2011). It also predicts performance on a host of other cognitive tasks, including, but not limited to, ACT performance (Frederick, 2005), performance on calculation choice tasks (Koehler & James, 2010), objective numeracy (Liberali, Reyna, Furlan, Stein, & Pardo, 2012), metacognitive advantage and awareness (Mata, Ferreira, & Sherman, 2013), choosing risk neutral solutions in risky economic games (Oechssler, Roider, & Schmitz, 2009), avoidance of paranormal beliefs (Pennycook, Cheyne, Seli, Koehler, & Fugelsang, 2012), endorsement of evolution (Gervais, 2015), and rejection of pseudo-profound bullshit (Pennycook, Cheyne, Barr,

Koehler, & Fugelsang, 2015a). It also predicts belief in supernatural and religious claims (Pennycook et al., 2012; Pennycook et al., 2016b), but this last relationship will be explored a little later.

How does the CRT work, though? Why is it so popular? For one, it is easy to administer. After all, it only contains three items. The correct answers are also accessible, in that the calculations involved are fairly simplistic. The CRT is valuable to this thesis, though, because it measures analytic thought within the context of dual-process theory. To illustrate this point, consider the following item from the CRT.

> A bat and a ball cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? _____ cents.

This problem provokes an intuitive answer, 10 cents, which is incorrect. The correct answer is actually 5 cents (i.e., if the ball is 5 cents, then a dollar more than the ball, which is the cost of the bat, is $1.05. Taken together, the ball, which, again, is 5 cents, and the bat, which we've calculated to be $1.05, cost $1.10 in total). The simple algebra of the puzzle becomes clear with the following setup:

x = price of bat y = price of ball	write correct equations from information in problem
x + y = $1.10 x = y + $1.00	substitute x
(y + $1.00) + y = $1.10	subtract $1.00 from both sides
y + y = $0.10	add the y's
2y = $0.10	divide by 2
y = $0.05	plug y back into 2nd equation
x = ($0.05) + $1.00	solve for x
x = $1.05	plug x and y back into 1st equation

($1.05) + ($0.05) = $1.10

The intuitive answer of 10 cents is rather enticing, because it is easy to add 10 to $1.00 and get $1.10. Such an initial response is prototypical of Type 1 processing, because it arrives to the mind quickly and with little effort. An individual unwilling to engage in analytical thinking is likely to answer with the incorrect 10 cents response and not contemplate the question further. If an individual wants to think about it a little more, however, they may engage in rationalization or decoupling. If the individual rationalizes, he or she may quickly add 10 cents to $1.00 to verify that the final answer is $1.10 without seriously considering other options. If the individual decouples, which is the stereotypical analytic class within Type 2 processing, he or she may add 10 cents, the intuitive price of the ball, to $1.10, the calculated price of the bat (which is $1.00 more than the price of the ball), and find that answer, $1.20, is incorrect. This mental contemplation and eventual falsification of the initial IR_1 response leads to a juncture where the individual must figure out a way to leave the intuitive trap. Presumably, he or she could set up a series of equations and solve for the prices of both items (as was demonstrated above). Alternatively, he or she could attempt a trial-and-error method where numbers less than 10 are systematically inputted into the model until 5 cents is found. Either way, such an analytic individual would be engaging in decoupling if they continue searching for the correct answer. Thus, one must engage in the three-stage dual-process model of cognitive reflection in order to get this first question of the CRT correct. The other two items of the CRT likewise have strong intuitive traps which require reflective effort to overcome.

A key note that needs to be made is that the CRT assesses both cognitive style and ability (Pennycook & Ross, 2016). It measures cognitive style because it necessarily requires individuals to initially consider the intuitive, but incorrect, answers. Thus, individuals who

answer correctly must override their initial intuitions about the problems they are confronted with in order to solve them and engage in Type 2 reasoning. One could argue that some people may not fall for these initial intuitive traps, and while theoretically true, the fact that most individuals miss at least one of the questions of the CRT (and said missed questions usually contain the intuitive, incorrect response) demonstrates that the initial, intuitive response (IR_1) is usually at least considered, if not outright accepted at the outset (Frederick, 2005; Pennycook, Cheyne, Koehler, & Fugelsang, 2016a). The CRT is also a measure of cognitive ability, because it requires analytic, intellectual processing to complete. Respondents must have a working familiarity with math or with systematic ways of thinking in order to solve the problems, and these processes fall under the umbrella of analytic processing capability. The calculations, while easily accessible, require intellectual fortitude. Could it be that the CRT is just another measure of mathematical ability, though? This question raises a contention that will be discussed in the next section. This current section, however, will conclude with a final thought on the utility of the CRT.

Despite the fact that the CRT measures both cognitive style and ability, it is mainly used as a measure of analytic cognitive style (ACS; Pennycook et al., 2012; Pennycook et al., 2016b). This claim is evidenced by the fact that one must necessarily engage in Type 2 processing in order to override their initial responses to the CRT items, unlike other measures of cognitive ability that do not have this Type 1/Type 2 component. Take, for instance, an item from the measure of mathematical ability by Galli, Chiesi, and Primi (2008):

> The double of ¾ is?
> A) 6/8 B) 3/2 C) 9/16 D) 3/8

The answer to this question requires calculation. One must recall the nature of fractions from previous mathematics classes and realize that adding three quarters of a quantity to another three quarters equals 1.5 or 3/2. Such calculation is algorithmic, but does not quire cognitive reflection. By this, I mean that there is no immediate answer that comes to mind that must be overridden by nearly everyone in order to find the correct solution. Rather, the purpose of this item (and other items like it) is to see how well one can use principles of mathematics, not how well one can override their initial gut feelings. The CRT is also unlike other measures of cognitive style in that it is performance-based. For instance, there is another measure of cognitive style, called the Need for Cognition, that measures one's *purported* propensity to engage in and enjoy thinking (Cacioppo & Petty, 1982; Cacioppo, Petty, Feinstein, & Jarvis, 1996). While related to cognitive reflection, what makes this measure of cognitive style different from the CRT is that it is a self-report measure. Therefore, it only measures how much one believes they engage in analytic thinking, not how much one *actually* engages in analytic thinking. In other words, despite measuring both cognitive style and ability, the CRT is superior to the Need for Cognition, at least for the purposes of this thesis, because it includes an element of how people actually think, not how they *think* they think.

The CRT and Mathematics: Cognitive Ability and Style

To complement the point that was made in the previous section, there are critics who propose that the CRT is nothing more than a glorified test of numerical ability (Sinayev & Peters, 2015; Welsh, Burns, & Delfabbro, 2013). Their argument is that the cognitive reflection components of the CRT have been over exaggerated and that numeric ability, rather than cognitive reflection (as measured by the CRT), more aptly predicts performance on decision-

making outcomes. These claims have been countered by proponents of the CRT who agree that the CRT is at least partially a measure of numeric ability (Pennycook & Ross, 2016). After all, the items require mathematical calculation in order to solve correctly. This fact is not a nail-in-the-coffin for the CRT, though. The CRT predicts moral and religious outcomes that are distinct from numeracy and intimately tied to analytic thinking (Pennycook, Cheyne, Barr, Koehler, & Fugelsang, 2014b), even when controlling for calculation or other measures of numerical ability. This, and other evidence of the CRT's association with non-numeric outcomes, like religion and use of smart phones (Pennycook et al., 2012), suggests that the evidence founded by critics of the CRT are, if anything, the exceptions, not the rule (Pennycook & Ross, 2016). Again, nobody is contending that the CRT only measures cognitive thinking style and does not incorporate some elements of numeric ability; rather, the argument presented is that the CRT contains components of analytic reflection that require numeric ability to solve. This fact is important, because it brings up the implication that the current measures that exist for cognitive reflection are limited. After all, the most popular and robust one, the CRT, only works within the confines of mathematical calculation. Are there perhaps other avenues of cognitive reflection that have not been studied? Put in another way, what would happen if the concepts of cognitive reflection were taken outside the domain of mathematics and incorporated within another domain of reasoning, such as, say, religion? This question is central to this thesis and will be explored in more detail in later sections.

Top-Down and Bottom-Up Factors for Type 2 Processing

There are certain factors that determine whether or not one will engage in analytic thinking. These factors are categorized into two groups: top-down or bottom-up (Pennycook et

al., 2015c). Bottom-up factors are "stimulus driven," whereas top-down factors are not stimulus driven and usually incorporate environmental or context factors. For comparison, and to illustrate the difference between the two factors, consider the bottom-up and top-down delineation in perception (Reisberg, 2013). With perceptual bottom-up processing, individual features of a stimulus are evaluated (Selfridge, 1959). Take, for instance, the English word, CLOCK. Bottom-up processing theory suggests that in order to recognize this word, one must detect (perhaps unconsciously) the features of the individual letters, such as the curvature of the first letter, C, or the vertical and horizontal line elements of the second letter, L, and piece them together to form letters. These letters in turn become the features that form the word as a whole. Bottom-up processing in decision making is similar to its perception counterpart insofar as it is based in the stimulus itself, whether it be in the wording or presentation of the stimulus (which is the problem to be solved). With perceptual top-down processing, the *context* is what influences perception, not the stimulus itself. A classic example of this type processing is shown in Figure 2. The middle "letter" in both words is actually a symbol that could be interpreted as either H or A. In other words, the stimulus is the same. However, the context of the words ensures that only THE and CAT will be read. If one were to misinterpret the ambiguous symbol, one would be left with TAE CHT (or another nonsensical variant), but because these types of incorrect interpretations have little to no meaning, especially in comparison to the most meaningful interpretation, THE CAT, the latter is perceived. Top-down processing in *decision making* is similar to its perception counterpart insofar as it is based on context, though context in the decision-making sense refers to environmental influences and personal dispositions. What follows in this section is a brief overview of top-down and bottom-up factors in decision making and how these factors either impede or encourage greater Type 2 processing.

Figure 2 Depiction of THE CAT perceptual trick that exemplifies top-down processing (after Selfridge, 1955)

Three top-down factors that influence analytic thinking are type of instruction (Daniel & Klaczynski, 2006; Evans, Handley, Neilens, & Over, 2010; Vadeboncoeur & Markovits, 1999), amount of time given to think (Evans & Curtis-Holmes, 2005; Finucane, Alhakami, Slovic, & Johnson, 2000), and individual differences (Frederick, 2005; Stanovich & West, 1998, 2000). Vadeboncoeur and Markovits (1999) investigated how type of instruction influences performance in decision making. They gave participants a conditional reasoning task, or one that required them to consider various "if-then" statements and determine which conclusions logically followed, and evaluated their performance. In one condition, participants were instructed to assume the given if-then statements were true and then make conclusions based off of them. In another condition, participants were given much clearer and thorough instructions on how to consider the if-then statements but were given the same conclusions from which to choose. Those who were in the condition that received the more detailed instructions chose the more logical conclusions than their counterparts who received simple instructions, thus indicating that merely instructing participants to think more thoroughly on conditional reasoning task increases reflective thinking. Having enough time to think about problems also increases reflective thinking. In a study, Evans and Curtis-Holmes (2005) gave half of their participants 10 seconds to complete a reasoning task and the other half an unlimited amount of time to complete

the same task. Those with the time constraint were more prone to responding in an intuitive, belief-influenced way than the other experimental condition who were given more time to reflect on their answers. Therefore, it appears limiting the amount of time one has to think increases the amount of intuitive or biased responses they give. Finally, in a host of evidence suggests that there are individual differences in cognitive reflection capabilities, meaning that some people have certain dispositions or tendencies that make them better at overriding their intuitive pulls than others. For instance, males typically tend to score better on the CRT than females (Frederick, 2005), and performance on typical measures of cognitive ability, like SAT scores, are moderately and positively correlated with measures of cognitive reflection, though the fact that some low-scoring SAT takers perform exceptionally well on these cognitive reflection tasks indicates that performance is not dependent on one's intellectual capability. Instead, cognitive reflection is an entity that is related to and perhaps relies on, but is still distinct from, intelligence. In summary, greater Type 2 processing in problem solving occurs when participants are instructed to think more thoroughly, when participants are given plenty of time to think, and when participants are males or score better on measures of cognitive ability, though this latter observation of individual differences is, indeed, individual-specific.

Introducing a bottom-up change to the stimulus also changes performance on reasoning tasks. Take, for instance, a modified version of the stereotypical ball and a bat problem mentioned earlier:

> A banana and a bagel cost 37 cents. The banana costs 13 cents more than the bagel. How much does the bagel cost?

There is not an intuitive answer that immediately jumps out to one's mind after reading this question, even though it is fundamentally similar to its predecessor. Instead, one must work

through the algebra to reach the appropriate solution. When presented this way, Frederick (2005) reported that more respondents answered it correctly (i.e., responded with "12 cents") than they did with the original ball-and-bat problem, which had the erroneous, intuitive pull. Changing the content of this particular item has likewise had a similar effect across a few studies (De Neys, Rossi, & Houdé, 2013; Mastrogiorgio & Petracca, 2014). This evidence indicates that the CRT problems are unique insofar as they have intuitive traps that oftentimes ensnare respondents, despite evidence that the mathematical computations needed to solve said problems are fairly accessible. Interestingly, this change of performance happens as a result of an alteration of the stimulus itself (a bottom-up factor). The same is true of performance on a belief-bias syllogistic reasoning task, which is another task that measures cognitive reflection. When this task was presented with nonsensical or meaningless words, participants did better on it than others who were given the same task with meaningful words and concepts that invoked bias based on their prior beliefs (Markovits & Nantel, 1989). That is, they were more likely to commit an error in reasoning by evaluating the merits of a syllogistic argument based off of the believability of its conclusion, not its logical structure. Taken together, these studies indicate that when the stimulus, or problem, is changed such that it does not have an intuitive trap or it contains an intuitive trap that is set without the enticement of an intuitive, prior-belief-based option, it will allow analytic processing to proceed.

These top-down and bottom-up factors undoubtedly influence one's propensity to think analytically. The terms themselves, top-down and bottom-up, are important too, because any factors that affect Type 2 processing will fall under either of these two categories. Moving forward, it will be helpful to elucidate to what extent other relatively unexplored factors fit best

within the two distinguishing categories, though such an analysis is not possible without first understanding what they mean.

The Content Effect

There is a tendency for subjects to commit logical errors in the Wason card selection task (described in detail later) when the content they are dealing with is abstract or unfamiliar. Likewise, when the content presented in the problem is familiar and concrete, subjects make logical errors much less frequently. This observation is known as the content effect in reasoning (Cosmides & Tooby, 1992; Cox & Griggs, 1982; Davies, Fetzer, & Foster, 1995).

At first glance, this phenomenon seems to run counter to the belief bias described in the previous section on bottom-up factors in decision making. After all, the tasks associated with the two phenomena in question are quite similar: both the belief bias in syllogistic reasoning tasks and the Wason card selection tasks require deductive reasoning to solve, both have intuitive solutions that most participants fall prey to unless they are overridden with greater analytical thinking, and both can incorporate unfamiliar or familiar terms (Cox & Griggs, 1982; Markovits & Nantel, 1989). Why is it that with the former, changing the stimulus to familiar terms increases error in response (i.e., belief bias), whereas with the latter, changing the stimulus to familiar terms instead *decreases* error in response (i.e., the content effect)? In other words, does familiarity with the content of a cognitive reflection problem increase or decrease Type 2 processing? This question is puzzling. One could answer it by claiming the problems that exemplify belief bias and content effect are fundamentally unique, despite their alleged similarities, and so differences in performance on these tasks are expected. If this is the case, though, then how could one predict whether belief bias or the content effect will prevail in other

tasks of cognitive reflection? Another solution, which supposes the efficacy of both camps and will be explored in the general discussion section of this thesis, is to investigate how both content and context affect Type 2 processing (Pollard & Evans, 1987).

For now, what is important to remember is that familiarity with the concepts of a problem changes how well one will perform on that problem. This observation will be critical moving forward, as the central argument of this thesis is that changing the content of reasoning problems changes performance. To this end, the content effect will be a global term to refer to times when familiarity with the content of a reasoning task increases performance and reduces intuitive responding, whereas belief bias will hereby refer to instances when one's prior belief instead *decreases* performance on reasoning tasks and enhances the power of the intuitive response.

Domain Generality and Specificity

It is worthwhile to take a moment now to distinguish between the content effect and a related term: domain specificity. When cognitive processing is fashioned by specific pieces of information or in specific contexts, then it is considered to be domain-specific; by contrast, when cognitive reasoning incorporates global cognitive processes, then it is considered to be domain-general (Beller & Spada, 2003; Roberts, 2008). What makes the content effect different from domain specificity is that the latter is associated with changes in reasoning *processes* based on context. The content effect says nothing about process. It is simply in reference to the tendency people have to reduce the number of decision-making errors they commit when the content they are thinking about is familiar to them. Whether or not their processes of reasoning are the same, or domain-general, between different tasks does not matter. In fact, the type of reasoning is probably exactly the same, but it gets triggered at different times in different contexts.

Take the example of a baker who is tasked with sorting cakes in a display case. He would perhaps start by labeling them and arranging them by size on the counter. Then, he would take into account color and intrigue of the cakes and think of ways to spatially arrange them such that the most exciting ones were seen first by customers. Finally, he would place them in the display case according to this mental schematic. The completion of such a task would take some time and know-how, but it would presumably be relatively easy for him, given his interest in and familiarity with cakes. Now, think about how this exact same baker would fare on a similar, yet different task: organizing books on a display table. He would probably first mentally organize the books similarly to the cakes, taking size into account first. Next, he would arrange them by color, intrigue, and overall spatial fit. Finally, he would set up the books according to his schematic. In terms of processing, the baker would use the same type of systematic thinking for both tasks. Such consistency in thinking across contents and contexts is domain-general. However, the new book-organizing task introduces the variable of unfamiliarity. That is, he would likely not be as proficient at organizing books on the display table by virtue of the fact that he was not as familiar with how they should be arranged in the first place. Should bestseller's go on top? Should some books be displayed standing up, while others should remain stacked on their backs? Even if he had the foreknowledge to anticipate these ancillary details, his ability to do so would not take into account his motivation, which could also affect his performance. Perhaps the task of piling books is not as personally meaningful or relevant to his own interests, and so while he would have the mental capability to do well, he would falter. The point of this example is to show that reasoning processes need not be domain-specific, in the cognitive sense, in order to be influenced by specific environmental or stimulus-based cues. The content of a problem could just as readily affect motivation or insight without fundamentally

changing the process of thinking as a whole. Thus, a domain-general process of thinking can be applied to different scenarios.

The term domain is still troublesome, though, as it can refer to both a mental process or the referent of a mental process (Pachur & Spaar, 2015). Colloquially, a domain is just a general area of knowledge about which one can reason. In this sense, psychology could be considered a domain. Or electronics. Or baseball. So, even though it has a more robust meaning among cognitive scientists (Roberts, 2008), the term domain will be regarded as a specific context in which thinking occurs. If domain is to refer to a process of thinking, it will be immediately followed by either the term specific or general.

The Current Paradigm

Now that many of the important terms have been discussed and defined, the current paradigm can be introduced. The point of this section is to describe what psychologists have to say about the relationship between religiosity and cognitive capability, which, simply put, is that nonreligious individuals tend to have greater intellectual and cognitive reflective capabilities than religious individuals (Pennycook et al., 2016b). The evidentiary support in favor of this assertion is strong. However, the fact that empirical data point in one direction does not mean they prescribe this direction as the only truth. All research has limitations, and all good scientific hypotheses, if they can be quantified and operationalized, can be challenged. The ending of this literature review will detail one way in which religious individuals could have the upper-hand on nonreligious individuals in matters of cognition. But first, the current paradigm.

Religion and Intelligence

There is significant support for the notion that general intelligence and religiosity are inversely related to each other. Lewis, Ritchie, and Bates (2011) found that IQ was negatively correlated with multiple dimensions of religiosity, including religious identification, private practice, and fundamentalism. Other research found cross-cultural support for this trend, demonstrating that among 137 different countries, disbelief in God is positively correlated to IQ (Lynn, Harvey, & Nyborg, 2009). To put these data in the context of mean-level differences, one large scale study reported that, on average, atheists had IQ scores 5.13 points higher than religious believers (Nyborg, 2009). Indeed, there is a consensus that higher religiosity is associated with lower general intelligence, with overall correlations between religious belief/practice and IQ ranging from -.20 to -.25 (Zuckerman, Silberman, & Hall, 2013). Even though the differences in intelligence between nonreligious and religious individuals are somewhat marginal and the effect sizes of the negative relationship between religion and IQ are modest, the trend is that nonreligious individuals appear to have, on average, somewhat greater intellectual capabilities than religious individuals.

Religion and Cognitive Reflection

Cognitive reflection, which is one's ability to check and detect intuitive errors according to the dual-process theory, is also strongly tied to religious belief (Pennycook et al., 2012). Cognitive reflection deserves attention in the scope of this investigation, because it may actually be a better predictor of religiosity than intelligence (Pennycook et al., 2012; Pennycook et al., 2016b; Shenhav et al., 2012). The evidence for this claim was laid out succinctly by Pennycook (2014), who explained that the CRT, the most popular measure of cognitive reflection, was

shown to positively predict religious beliefs across a number of studies. Importantly, though, this predictive capability remained significant when controlling for cognitive ability, as measured by the Wechsler Adult Intelligence Scale Matrix Reasoning Test and the Shipley Vocabulary Test, age, gender, openness to experience, and extraversion (Shenhav et al., 2012). Another study of a similar nature showed similar results. When controlling for variables such as sex, age, conservatism, education, religious engagement, and, of course, cognitive ability (as measured by the Wordsum verbal intelligence test), measures of ACS, which included the CRT and another purported measure of cognitive reflection, significantly predicted religious beliefs (Pennycook et al., 2012).

Finally, in a recent meta-analysis of over 20 studies containing over 15,000 total participants, Pennycook et al. (2016b) found that performance on a number of measures of ACS, including the CRT, belief-bias syllogisms, and base-rate neglect problems, correlated negatively with various measures of religious belief. In their own series of four studies with over 1,000 participants (that were included in the aforementioned meta-analysis), they found that purported atheists scored the highest on the cognitive reflection measures. As the gradation continued from atheist, to agnostic, and to belief in a personal deity, participants scored increasingly lower, suggesting that, on the spectrum of affiliations relating to belief of God, those who deviate more from the conventional belief in a personal deity tend to be more reflective and more prone to overriding their initial intuitions than their more religious counterparts. Additionally, ACS remained a significant predictor of religious belief when numeracy and verbal intelligence were controlled for in a multiple regression analysis. This final meta-analysis and four-study investigation acts as a proverbial lighthouse for the current paradigm: across multiple studies and

measurement parameters, nonbelieving individuals are more reflective thinkers than religious believers.

Alternatives to Current Paradigm: A Case for Conditional Reasoning

The main argument posed in this thesis is that familiarity with certain tasks gives certain individuals or groups of individuals advantages over others. As was just reviewed, the current paradigm suggests that religious individuals are generally not as analytical or reflective as their nonreligious counterparts. However, this conclusion is made based off of performance on only a few tasks (Pennycook et al., 2015b). The first is the CRT, which was discussed in an earlier section. The second is the base-rate neglect task, which requires participants to suppress intuitive stereotypes about groups of people and select answers in favor of statistical probabilities. The third is the belief-bias in syllogistic reasoning task, which prompts participants to assess the validity of syllogistic arguments, despite the fact that the correct answers to these syllogisms come in conflict with their intuitive beliefs. These are the most popular and often-used measures of cognitive reflection, and overwhelmingly, nonreligious individuals perform better on them than religious individuals (Pennycook et al., 2016b). However, the existence of even one measure of cognitive reflection in which religious individuals consistently outperformed nonreligious individuals would add nuance to the current paradigm, suggesting instead that nonreligious individuals are more reflective than religious individuals in *most* or *some* measures of cognitive reflection.

The purpose of this section is to explore instances in which familiarity with certain concepts or cultures influenced performance on reasoning tasks. I argue that the process of reasoning, while perhaps structured under domain-general "mindware", is multifaceted and

selective. Intuitive and reflective inclinations in decision making are subject to the influence of individual skills/talents, group cultures, individual familiarities, and the contents of the tasks themselves. In short, reasoning is conditional.

Multiple Intelligence Theory

In his seminal work, *Frames of Mind; The Theory of Multiple Intelligences*, Howard Gardner proposed that there are eight different intelligences (Gardner, 1983, 1999). Three of the eight intelligences, visual-spatial, verbal-linguistic, and logical-mathematical, were essentially reiterations of the traditional model of IQ. However, Gardner argued for the existence of more intelligences, such as musical-rhythmic intelligence, bodily-kinesthetic intelligence, interpersonal intelligence, and naturalistic intelligence, that had functional, culturally-valued outcomes and could helped people solve problems. Importantly, the inclusion of such new intelligences partly revolutionized the way teachers and educators evaluated the abilities of their students. For instance, the student athlete who excelled in basketball but still received low scores on math tests in school could still be regarded as having a high degree of intelligence, though just in a different domain (i.e., bodily-kinesthetic). So, too, could the quiet mathematical prodigy find solace in having high logical-mathematical intelligence but low interpersonal intelligence. This framework was meant to be inclusive and show that everyone had unique abilities above-and-beyond the rigid structure of academic intelligence. Indeed, after its inception, multiple intelligence (MI) theory gained positive support among students, teachers, and parents who regarded it as a happier alternative to the idea that there is only one kind of "smart" (Díaz-Lefebvre, 2004; Mettetal, Jordan, & Harper, 1997).

Among academics, however, MI theory did not gain as much support, with many critics challenging its theoretical foundations, its definitions, and its empirical backing (Visser, Ashton, & Vernon, 2006; Waterhouse, 2006b; White, 2008). For example, in one study, intercorrelations between performances on verbal and nonverbal versions of the eight intelligences were quite high (Visser et al., 2006). In other words, participants who did well in one domain, like musical-rhythmic intelligence, were fairly likely to succeed in other domains, like verbal-linguistic intelligence. This evidence appears to run counter to MI theory, which suggests that participants who do well in one domain would not predictably do as well in other domains. Gardner has responded to this and other waves of criticisms (Gardner & Moran, 2006), which has, in turn, sparked more controversy (Waterhouse, 2006a).

Whether or not MI theory holds significant merit for intelligence as a whole is still up for debate, but the key takeaway is that it was an attempt to undermine the current order at that time. And it succeeded, insofar as it raised the question of individual differences in intelligence theory. Even though the norm seems to be that general intelligence pervades among many skillsets and abilities (Visser et al., 2006), this fact is the trend, not the absolute standard. Part of Gardner's theory was based on the savant, or the genius who excelled significantly above average in one domain, despite deficiencies in others (Gardner, 1983). Such individuals are extraordinary and uncommon. They do not necessarily signify a trend. However, they do indicate the *possibility* of individual differences, or the possibility that the talents of some individuals are varied (even in the cognitive domain). Such diversity is meaningful, as it gives rise to the idea that measures of cognitive ability are limited if they do not take into account the unique life experiences of the individual. Such a distinction becomes much more apparent in the next example.

The BITCH-100 Test

In order to demonstrate that mainstream intelligence tests were culturally biased, Robert Williams developed the Black Intelligence Test for Cultural Homogeneity (BITCH-100; Williams, 1972). This test was meant to be a measure of intelligence for African-American individuals pertaining to their unique language, attitudes, and lifestyles. As such, its items were culture-specific, not culture-fair. A culture-fair test would be designed to not give any particular groups or individuals advantages over others; by contrast, a culture-specific test would be designed to give unfair advantages to certain groups. As a culture-specific test, the BITCH-100 was designed to give Black individuals an advantage over White individuals, and it did. In the original sample comparing 100 White individuals to 100 Black individuals on the BITCH-100, the Black individuals outscored the White individuals by a mean of 36.00 points. This effect has been repeated multiples times with the same results: when given a normal IQ test, White individuals outscore Black individuals, but when given the culture-specific BITCH-100, Black individuals outscore White individuals (Matarazzo & Wiens, 1977; McNiel, 1975).

Clearly, the cultures and experiences of particular groups of individuals play crucial roles in how well they perform on reasoning tests, especially if those tests do not take into account these unique narratives. The BITCH-100 serves as a sober reminder that participants must be familiar with the content they are presented with in order to have a chance of demonstrating their cognitive prowess. It also demonstrates the case that cognitively disparaged groups of individuals (e.g., Black individuals) are not as incapable at thinking and reasoning as their relatively poor performance on test items would indicate. A solution to this problem, which Williams showed in 1972, is to create measures with items that both require mental fortitude and are unique to the experiences of these minority groups.

Wason Card Selection Task

A reasoning task need not pander to one group solely in order to show marked improvement among participants. Simply changing the nature of the problem to something familiar is enough. Take, for instance, the Wason card selection task (Cosmides & Tooby, 1992; Wason, 1966). In this task, participants are presented with four cards (two letters and two numbers) and told that every card has on it a letter on one side and a number on the other (Figure 3). They are also given a rule that they must determine the truth of. One such rule could be "if the card shows a D on one side [*p*], then its opposite face must be 3 [*q*]." They are then instructed to turn over as many cards (and only as many cards) as is necessary to determine the truthfulness of this rule. When the problem was presented in this abstract format (see Figure 3, top), only around 10-25% of participants answered correctly by flipping over the card with the D (*p*) and the card with a 7 showing (not *q*) (Cosmides & Tooby, 1992). However, when the context of the problem was changed and participants were instead asked to determine whether or not actors were violating social contracts (i.e., by drinking beer while underage) (Cox & Griggs, 1982; Pollard & Evans, 1987), they answered correctly (*p*, not *q*) around 75% of the time (Cosmides & Tooby, 1992).

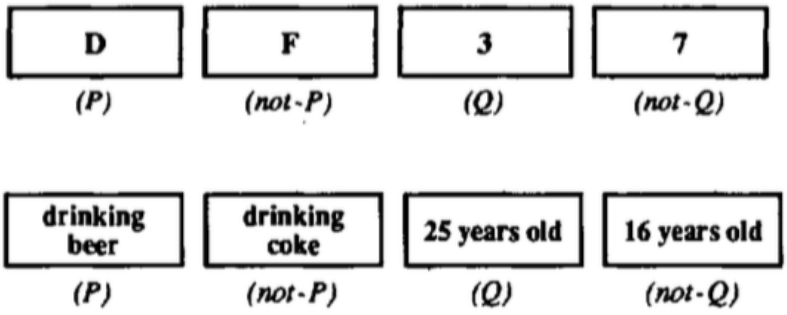

Figure 3 Depiction of Wason card selection tasks (images from Cosmides & Tooby, 1992)

The Wason card selection task demonstrates that there is a content effect to thinking critically (Davies et al., 1995). That is, when the content of a problem changes such that it incorporates terms or ideas which are more familiar, performance on the problem increases. This logic could also be extended to performance on measures of cognitive reflection, because after all, the Wason card selection task could be considered a type of cognitive reflection task. When given in its abstract form, most people (~41%) erroneously select *p* and *q* (Cox & Griggs, 1982). In a sense, participants would have to overcome the intuitive pull to select the cards that first come to mind with the statement of the problem (i.e., determine the truth of the rule "if *p*, then *q*"). Thus, they would have to activate Type 2 processing in order to test their initial response of *p* and *q*, realize it does not work, and then try further alternatives. Such algorithmic processing is indicative of reflective thinking, and thus, the Wason card selection task is a task of cognitive reflection. So, altogether, there is evidence that suggests that changing the content of one measure of cognitive reflection to something familiar increases performance.

Domain-Specificity Preferences in Decision Making

Not only do content effects apply to a very narrow scope of alterations for a specific measure of reasoning, but there appears to be a spectrum of topics to which intuitive-deliberative preferences in decision making are applied. Take, for example, the tasks of selecting a mate or purchasing a new electronic item. Which would influence one to rely on their intuitions? Which would invoke a greater preference for deliberation? The answer may be a little surprising. Selecting a mate is associated with more intuitive modes of thinking, whereas buying a new electronic device is associated with more deliberative modes of thinking. In a study evaluating how much participants reportedly relied on their intuitions or their deliberations for making

decision in a number of domains, Pachur and Spaar (2015) found that, in addition to selecting a mate, people were more likely to select their clothing and choose restaurants based on their intuitions. Likewise, they were more inclined to use deliberative processing when choosing a doctor or buying electronic devices. Interestingly, they were equally likely to prefer intuitive and deliberative processing when making vacation plans, probably because such a judgment required them to balance their intuitions (e.g.., do they feel they want to go somewhere warm, like a beach, or somewhere exciting, like an amusement park) with their deliberations (e.g., deciding how to budget the vacation).

This study illustrates the important point that one's purported willingness to engage in intuitive or deliberative thinking is dependent on what one is deciding, even if one is more inclined to think intuitively or deliberatively overall. For instance, one could hypothetically claim to be a deliberative thinker in general and apply critical thought to math problems but then intuitively choose one particular color scheme for an unpainted room because it "looks better". The task of painting a room would not invoke the same type of reflective thinking in this individual. Truly, deliberative processing is selective and contingent on the issue at hand.

The Familiarity Factor

In another example of the effect of content on the ability to reason, college students from a variety of academic backgrounds were posed three types of formal reasoning problems, which required use of abstract thought, hypothesis testing, deductive reasoning, and problem-solving to complete at the formal operations level (De Lisi & Staudt, 1980; Piaget, 1972). The first type was a physics-based problem that required the participants to systematically isolate the variables that would change the frequency oscillations of a pendulum. The second type was a political

socializations concepts task that required participants to make political decisions, justify them, and then consider opposing viewpoints. The third type was a literary styles analysis task that required participants to match the literary styles of various works. As expected, the researchers found that a high percentage of those who were well-versed in physics and mathematics completed the physics-based task at the formal operations level (~90%) (De Lisi & Staudt, 1980). The political science and English students who were given the same physics-based task also scored at the formal operations level, but because they were not as familiar with physics, a lesser percentage of them scored at this highest level (only ~40%). Interestingly, the math and physics students did not outperform the other students on the other two tasks. Instead, when the styles of the tasks were switched, those who had a task related to their field of interest did better than their academic counterparts. That is, a higher percentage of those who were political science and English majors answered formal operations questions on political socialization and literary styles, respectively, at the formal operations level compared to everyone else. For a graphical visualization of this phenomenon, see Figure 4.

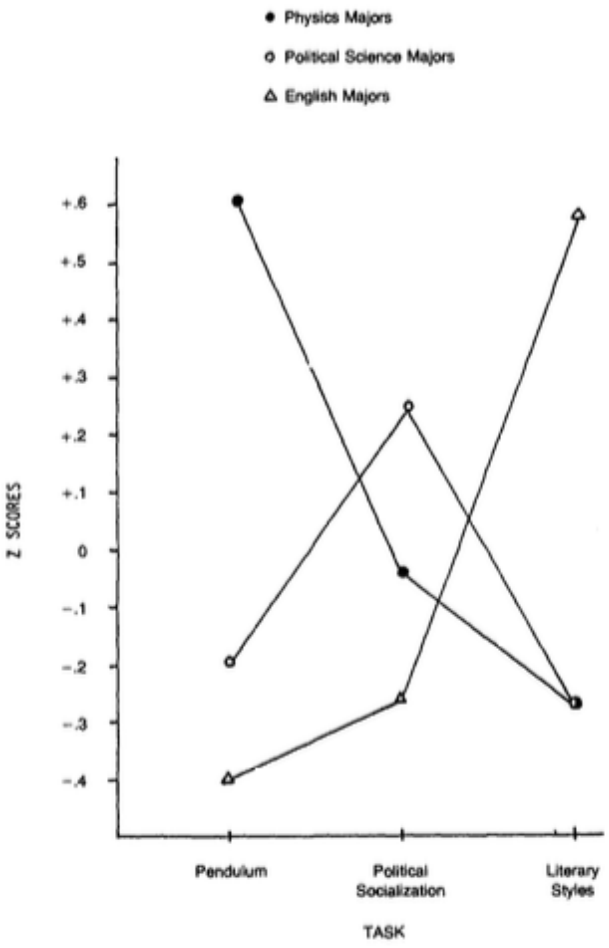

Figure 4 Mean scores on three formal operations tasks by college major groups (taken from De Lisi & Staudt, 1980)

This pivotal study lays the foundation for the argument posed in this thesis, which is that familiarity with certain domains of reasoning affects performance on reasoning tasks. This argument, which is essentially an updated version of the content effect, will be called the familiarity factor. It is distinct from the content effect in one major way: instead of focusing purely on bottom-up (i.e., stimulus-based) factors, as the content effect does, the familiarity factor focuses on both bottom-up *and* top-down factors. An explanation follows.

The content effect is based on the fact that changing the stimulus, or the content, of a reasoning problem changes performance. Namely, in the case of the Wason card selection task, making the content more familiar to participants increases performance (Cox & Griggs, 1982). This effect is solely based on bottom-up factors, though. Sure, participants must have some working knowledge of policing social norms in order to (unconsciously) appreciate the ease of the age drinking problem over its more difficult abstract counterpart, but such acknowledgement is readily accessible to all who understand social norms in the first place. Therefore, changing the content to align with the familiarity of the participant happens nearly ubiquitously, and such near universality is, for all intents and purposes, a bottom-up phenomenon. By contrast, the familiarity factor incorporates both bottom-up and top-down processing. To begin, the familiarity factor depends on the stimulus-in-question changing. In the aforementioned study with the three types of college students, the stimulus changed from physics-based, to political science-based, to English-based. However, and most importantly, this change in the stimulus only had an effect on the participants *who belonged to different groups*. That is, changing the stimulus to appeal to the familiarities of each of the three represented groups either changed or did not change their performance, depending on which group they belonged to. Take, for instance, the change of the task from physics-based to political science-based. Such a change in stimulus had a drastic effect on both the physics and math students and the political science students (i.e., the change hurt the former and aided the latter). The English students, however, experienced very little change in performance at all. Such a recognition that the content of a stimulus can alter performance for different groups, based specifically on a top-down factor delineation, is the familiarity factor.

Rationale for Studies

I have attempted to demonstrate in this literature review that reasoning is complicated. Even if reflective thinking follows the three-stage dual-process model (Pennycook et al., 2015c), the issue remains of when such reflective thinking is invoked. According to one major camp, analytical thinking, skepticism, and religious disbelief go hand-in-hand (Norenzayan & Gervais, 2013; Pennycook, Cheyne, Koehler, & Fugelsang, 2013; Pennycook et al., 2016b), alluding to the fact that some individuals are more predisposed to thinking reflectively than others. According to another camp, success in reasoning is based instead on culture, content, and familiarity (Cox & Griggs, 1982; De Lisi & Staudt, 1980; Matarazzo & Wiens, 1977). The purpose of these studies is to explore the middle ground between these two camps and challenge the paradigm set forth by the first camp that religious disbelievers are more reflective thinkers than religious believers. The purpose of the first study is to replicate the major findings of Pennycook et al. (2016b) and take the first step to show that the CRT measures cognitive reflection within the domain of mathematics. The purpose of the second study is to replicate the major findings of Pennycook et al. (2016b) again and also take the first step to show that a novel measure of cognitive reflection within the domain of religion, the Religion Reflection Scale, can invoke reflective thinking in religious individuals on par with their nonreligious counterparts.

Figure 5 illustrates the theoretical basis behind this exploration. The left graph depicts predicted scores between math majors and psychology majors on the original, math-based CRT and on a hypothetical "Psych CRT," which would have questions within the domain of psychology that measure cognitive reflection. If indeed cognitive reflection occurred independent of content or familiarity, then upper-level math and psychology students, who presumably have the same access to higher-level cognitive functioning, would score equally well

on a general measure of cognitive reflection, as the math-based CRT is purported to be. If, however, cognitive reflection dependent on content or familiarity, then upper-level math students would outperform upper-level psychology students if the cognitive reflection task were based in mathematics, whereas the reverse would be true if the task were based in psychology. Both versions of the CRT would still measure cognitive reflection, but would do so *in the domain of the discipline in question*, which would give an unfair advantage to those individuals who were most familiar with that domain.

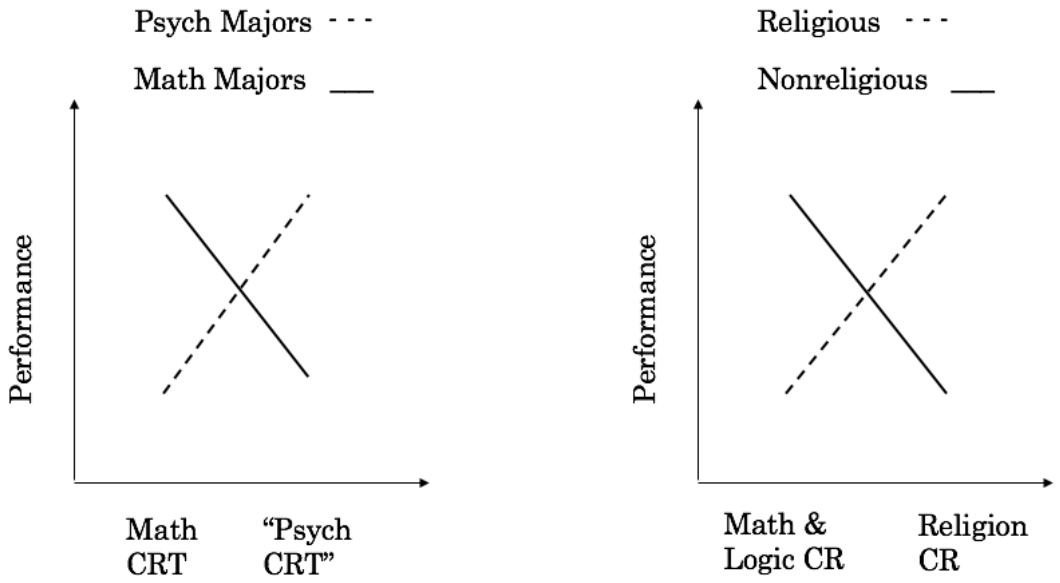

Figure 5 Side-by-side comparison of theoretical basis for Study 1 (left) and Study 2 (right); CRT = Cognitive Reflection Test; CR = Cognitive Reflection

The right graph in Figure 5 depicts the logical extension of the aforementioned argument. Nonreligious individuals tend to do better than religious individuals on cognitive reflection tasks that involve math-based reasoning and logic with respect to general facts (Pennycook et al., 2013; Pennycook et al., 2016b). This pattern, while somewhat marginal, appears to be consistent.

However, I argue that these tasks may not assess the *totality* of all domains of cognitive reflection. Perhaps those who are religious perform poorly on the CRT and other cognitive reflection tasks because they are less familiar with problem-solving or analytical thinking in the domain of mathematics or general content syllogistic logic, not because they are less reflective thinkers overall. It stands to reason that a cognitive reflection task that was designed to measure cognitive reflection *within the domain of religion* would yield results consistent with the notion that religious individuals may think analytically within their familiar domain. Thus, religious individuals would perform better than nonreligious individuals on a measure of religious cognitive reflection.

Main Research Hypotheses

Hypothesis 1: The major findings of Pennycook et al. (2016b) will replicate. That is, measures of cognitive reflection will be negatively related to religious belief.

Hypothesis 2: Students in upper-level mathematics courses will score higher on the CRT than students in upper-level psychology courses.

Hypothesis 3: Religiosity will be positively associated with religious reflection.

Study 1 Research Hypotheses

Hypothesis 1a: The CRT and a measure of analytic cognitive style (ACS; i.e., a composite score of the original CRT, the additional CRT, and the belief bias in syllogistic reasoning task) will negatively correlate with religious belief.

Hypothesis 1b: Categorically, atheists and agnostics will outperform theists on all measures of ACS.

Hypothesis 1c: ACS will predict religious belief, even when controlling for numeracy, gender, age, and reported GPA.

Hypothesis 2: Students in upper-level mathematics courses will score higher on the CRT than students in upper-level psychology courses.

CHAPTER III

STUDY 1 METHOD

Participants

Participants were selected via convenience sampling (class-to-class or student lounge recruitment) between March 2017 and April 2017 from the University of Tennessee at Chattanooga (UTC), a public university located in southern United States. There were 110 participants: 82 from upper-level psychology courses (80% female) and 28 (36% female) from upper-level mathematics courses. All participants answered all questions. The average age was 24.20 years (*SD* = 7.637) and the average reported GPA was 3.34 (*SD* = 0.439) out of a possible 4.0. Most participants were seniors (81.8%) and most identified as having a belief in a personal God or higher power (74.5%). Twenty-eight students were math majors (e.g., mathematics, actuary science, applied statistics, etc.) with experience in upper-level math concepts, whereas the remaining 82 were non-math majors (e.g., psychology, child and family studies, communications, etc.) recruited from upper-level psychology courses. The breakdown of religious identification in decreasing percentage was as follows: Christian (41.3%), Protestant Christian (18.3%), Agnostic (15.6%), Catholic Christian (10.1%), No Religion (6.4%), Atheist (3.7%), Other (3.7%), and Jewish (0.9%). Participation was voluntary and some students (psychology majors) received course credit for participating. This research was approved by the Institutional Review Board of the University of Tennessee at Chattanooga (Appendix A).

Materials

Demographics

Data were collected on age, gender, current major, academic classification, and current purported GPA of each student (Appendix B).

Religious Beliefs

The religious beliefs of the participants were measured through multiple means (Appendix C). The first series of questions was an 8-item scale that assessed participants' conventional beliefs in religious/supernatural entities, including heaven, hell, miracles, afterlife, angels, demons, soul, and the devil/Satan (Pennycook, Cheyne, Barr, Koehler, & Fugelsang, 2014a; Pennycook et al., 2016b). Participants indicated their level of belief in each religious/supernatural entity by responding to a 5-point Likert scale (from strongly disagree to strongly agree). Higher scores indicated higher religious belief. The estimate of internal reliability for this scale from this sample was excellent ($\alpha = .942$). The next question asked the participant to identify what *type* of God, if any, they believed in. Those who indicated they believed in 1) A personal God, 2) God as an impersonal force, or 3) A God who created everything but does not intervene in human affairs were coded as theists. Those who indicated they 4) Don't know whether or not any Gods exist, or 5) Don't know whether or not any Gods exist and no one else does either were coded as agnostics. Those who selected the options 6) I don't believe in Gods of any sort, or 7) I believe that God does not exist were coded as atheists. Finally, participants were given a question where they indicated what religion, if any, they subscribed to.

Numeracy

To measure numeracy, an abridged version of the Mathematical Prerequisites for Psychometrics (PMP) was crafted and administered. The PMP was originally created to measure student mathematical ability across six domains: Operations, Fractions, Set theory, first order Equations, Relations, and Probability (Galli et al., 2008). It shows good construct validity with positive attitudes towards statistics, $r = .41$, $p < .01$, and good predictive validity with failure and non-failure of exams, $t(305) = 7.45$, $p < .001$. The estimate of internal reliability of this scale from this sample was acceptable ($\alpha = .788$).

This questionnaire was chosen as a measure of numeracy, because it requires participants to have prerequisite knowledge of statistical operations at a level that most non-math majors can understand (Galli, Chiesi, & Primi, 2011). Therefore, success on this measure was accessible to all students in the sample. In order to craft the abridged version of the PMP, for the purpose of reducing participant fatigue, the 30 original PMP items were sorted into two 15 item pools of equal difficulty. One of the pools was selected at random to be used in this pilot study (Appendix D).

Belief Bias in Syllogistic Reasoning

This task consists of eight syllogistic reasoning problems, derived from the work of Markovits and Nantel (1989), that have been used to evaluate non-numerical, logical reasoning (Toplak, West, & Stanovich, 2014). Participants were asked to assume the premises of each syllogism were true and assess whether or not the conclusions followed logically from the premises (Appendix E). In order to answer them correctly, however, they had to set aside their prior knowledge and make conclusions from new, counterintuitive premises. For example, one

item reads, “Premise 1: All things that are smoked are good for the health; Premise 2: Cigarettes are smoked; Conclusion: Cigarettes are good for the health.” The correct answer is that this syllogism is valid, but because most people hold the belief that smoking is bad for the health, their belief bias may lead them to (incorrectly) conclude that this syllogism is invalid. The other seven items also have conclusions *inconsistent* with their validity structures. As such, this task measures cognitive processing akin to dual-process theory, though unlike the CRT, it does not require mathematical skill to solve correctly. Responses to this task are dichotomous, and scores are calculated as sum totals of correct answers. Possible scores range from 0-8 with higher scores indicating greater cognitive reflection.

The task has received poor estimates of internal reliability in other studies (Chronbach’s $\alpha = 0.64$; Toplak et al., 2014), but in the current sample, the estimate was higher ($\alpha = .786$). It has also been used as a predictive measure of religious belief, just like the CRT (Pennycook et al., 2013).

CRT (Original)

The CRT is a measure of cognitive reflection, given that it measures respondents’ propensity to override their intuitions on a series of mathematics questions and reach an analytical answer (Frederick, 2005). It consists of three questions, each of which has an intuitive, incorrect answer and an analytical, correct answer (Appendix F). Estimates of the internal reliability of the CRT have been low-to-adequate (α’s between 0.60 and 0.74) (Liberali et al., 2012; Stieger & Reips, 2016), though for this sample, the estimate was stronger ($\alpha = .797$). Participants were not allowed to use calculators, but they could spend as much time as they wanted on these questions. Scores on this measure range from 0-3, with each number

representing how many questions each participant got correct. High scores indicate strong analytical cognitive style.

CRT (Additional)

This measure contains 4 items that are similar in design to the original CRT questions (Appendix G), because they each have an intuitive and an analytical answer, but they are useful in this research in that they are not as well-known as the original three and thus are not prey to carryover effects (Toplak et al., 2014). Scores on this measure range from 0-4, with each number representing how many questions each participant got correct. As before, high scores indicate stronger analytical cognitive style. The CRT (additional) correlates fairly well with the CRT (original) (r = .55+; Pennycook et al., 2016b). The correlation between the two scales reached this range in the current sample (r = .645, p < .001). Estimates of internal consistency for the two CRT's combined is slightly higher than the original CRT alone as well (α = 0.72; Toplak et al., 2014); the estimate from the current sample was similar in magnitude (α = .793). For the sake of time and participant fatigue, the last item of the CRT (additional) was removed from all surveys in Study 1, but it returned in Study 2.

Procedures

Participants were given hard copies of the battery of questionnaires in mixed order. Most surveys were administered in large classroom settings, but for the sake of recruiting math majors, some were administered in more casual, "student lounge" settings. Participants were instructed to go through the packet and answer the questions accurately, honestly, and with no calculators. The average time that students needed to complete the survey was 15-25 minutes.

CHAPTER IV

STUDY 1 RESULTS

Descriptive statistics for the variables and a correlation matrix of the cognitive tasks can be found in Table 1 and Table 2 below. Participants in this sample responded quite highly to the religious beliefs questionnaire (M = 3.68, SD = 1.060), reflecting the nearly 75% belief in a God or higher power reported earlier. Participants also excelled at the numeracy task. The reverse was true of the CRT items, all of which showed high positive skewness. Therefore, ceiling and flooring effects could have partially reduced the efficacy of the results of the statistic tests described in the next sections. The responses to the belief bias in syllogistic reasoning task and the measurement of analytic cognitive style (ACS), which was the sum total of the original CRT items, the additional CRT items, and the syllogism task, were much more normally-distributed. Finally, all the cognitive tasks were significantly and positively related to teach other, as expected.

Table 1 Descriptive Statistics of Variables in Study 1

	N	**Min**	**Max**	**Mean**	**SD**	**Skewness**	**Kurtosis**
Religious Beliefs	110	1	5	3.68	1.060	-0.696	-0.289
Numeracy	110	3	15	11.45	2.970	-0.680	-0.172
CRT (original)	110	0	3	0.98	1.125	0.745	-0.894
CRT (additional)	110	0	3	0.97	1.036	0.762	-0.604
CRT Total (w/o recognition)	67	0	6	1.48	1.761	1.256	0.765
CRT Total	110	0	6	1.96	1.958	0.866	-0.393
Syllogisms	110	0	8	4.70	2.337	-0.293	-0.874
ACS	110	0	14	6.66	3.801	0.295	-0.761

Table 2 Correlation Matrix of Cognitive Tasks in Study 1 (all *p*'s < .001)

	CRT (original)	CRT (additional)	Numeracy	Syllogisms
CRT (original)	1			
CRT (additional)	.645	1		
Numeracy	.565	.550	1	
Syllogisms	.525	.489	.558	1

Hypothesis 1

Overall performance on the original CRT negatively correlated with religious belief (r = -.210, p = .028; see Table 3). Another measure of cognitive reflection, the belief bias syllogisms, negatively correlated with religious belief (r = -.265, p = .005). ACS showed a similar negative relationship with religious belief (r = -.232, p = .015). Numeracy did not significantly correlate to religious belief (r = -.039, p = .683). Unexpectedly, these were the only relationships that coincided with the Pennycook et al. (2016b) study. The additional CRT questions did not correlate significantly with religious belief (r = -.021, p = .824), nor did performance on all the CRT items, whether calculated with those who were familiar with at least one of the items (r = -.134, p = .162) or without them (r = -.116, p = .352, N = 67). The correlations of the non-math majors from the upper-level psychology courses fell in line with the expected trend, but the correlations of the math majors showed atypical variations. These anomalies will be discussed in the next chapter.

Table 3 Correlations (*r*) Between Religious Belief and Performance on Cognitive Tasks in Study 1

	CRT[1]	**CRT[2]**	**CRT[3]**	**CRT Total**	**Syllogisms**	**ACS**	**Numeracy**
Non-Math Majors$_{(82)}$	-.309**	-.049	-.182$_{(54)}$	-.226*	-.273*	-.298**	-.090
Math Majors$_{(28)}$	-.078	.051	-.180$_{(13)}$	-.017	-.299	-.173	.332
Combined$_{(110)}$	-.210*	-.021	-.116$_{(67)}$	-.134	-.265**	-.232*	-.039

In Table 3, CRT[1] = Accuracy on original 3-item CRT; CRT[2] = Accuracy on additional CRT problems; CRT[3] = Accuracy on all 6 CRT problems excluding participants who indicated seeing at least one of the problems; CRT Total = Accuracy on all 6 CRT items; and ACS = Analytic Cognitive Style (addition of CRT Total and Syllogisms). *N* listed in parentheses. * indicates $p < .05$ and ** indicates $p < .01$

A series of one-way ANOVA's were conducted to determine which of the three theistic groups, theist, agnostic, or atheist, performed best on the cognitive tasks. When ACS was inputted as the dependent variable, there was a statistical difference between the three groups, $F(2, 107) = 4.260$, $p = .017$, $\eta^2 = .074$ (Figure 6). A post-hoc Tukey HSD test revealed that agnostics ($M = 8.57$, $SD = 3.488$) scored significantly higher than theists ($M = 6.07$, $SD = 3.800$, $p = .014$), though atheists ($M = 7.60$, $SD = 2.302$) did not score significantly higher than theists ($p = .643$) or agnostics ($p = .857$). There were only two other cognitive tasks for which this trend continued: the original CRT ($F = 5.694$, $p = .004$) and the syllogisms ($F = 4.941$, $p = .009$). Analyses of all other models of performance on the remaining cognitive tasks, including those that excluded participants who had seen the CRT items before, revealed no statistically significant differences between the three groups (F's < 2.423, p's $> .097$).

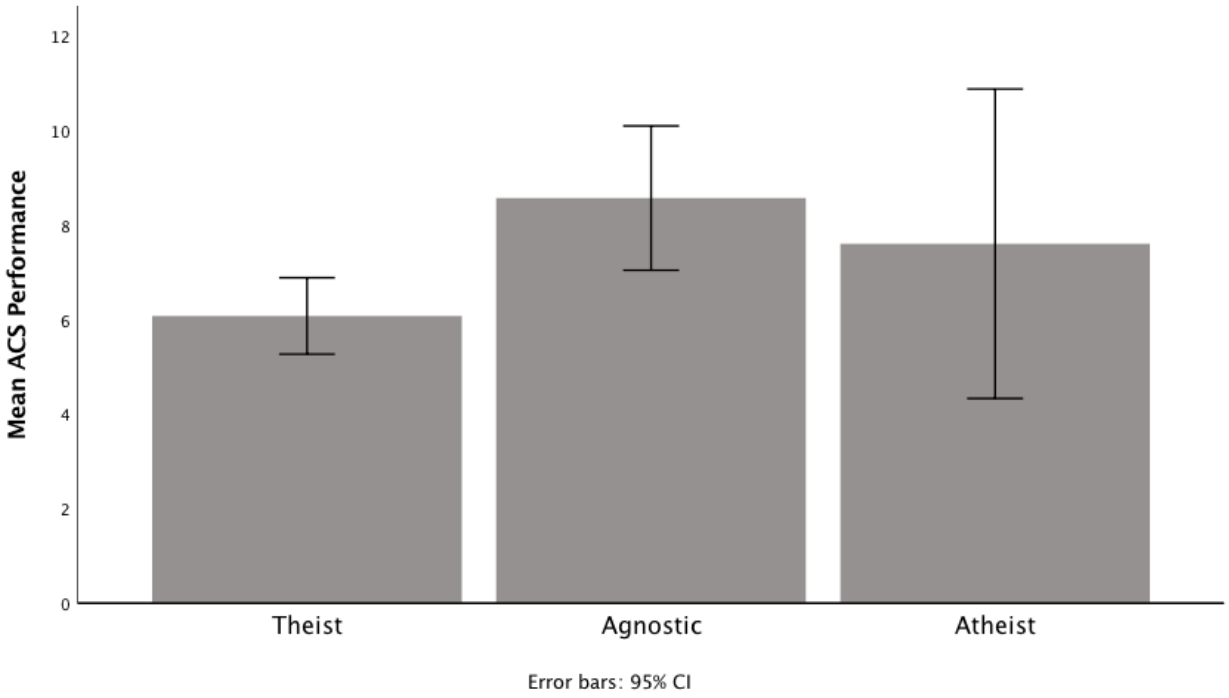

Figure 6 Mean accuracy on analytic cognitive style (ACS) as a function of theism in Study 1 (*N*'s = 82, 23, and 5, respectively)

Religious belief was regressed onto ACS, numeracy, self-reported GPA, age, and gender. This multiple regression model was statistically significant, $F(5, 101) = 2.662$, $p = .027$, and ACS emerged as a significant predictor of religious belief ($\beta = -0.116$, $p = .002$, 95% CI [-.188, -.044]), whereas numeracy ($\beta = 0.073$, $p = .133$, 95% CI [-.022, .168]), gender ($\beta = -0.266$, $p = .267$, 95% CI [-.739, .207]), age ($\beta = -0.002$, $p = .906$, 95% CI [-.035, .031]), and self-reported GPA ($\beta = 0.281$, $p = .272$, 95% CI [-.224, .786]) did not. This evidence indicates that analytic cognitive style is a powerful and unique predictor of religious beliefs.

Hypothesis 2

According to an independent-samples *t*-test, math majors ($M = 1.93$, $SD = 1.215$) significantly outperformed non-math majors ($M = 0.66$, $SD = 0.892$) on the CRT (original),

$t(108) = 5.904$, $p < .001$ (Figure 7). In a statistically significant multiple regression model, $F(4, 102) = 19.369$, $p < .001$), math majors still performed better than non-math majors on the CRT (original) ($\beta = 0.492$, $p = .038$, 95% CI [.029, .956]), even when controlling for numerical ability ($\beta = 0.118$, $p = .003$, 95% CI [.040, .196]), GPA ($\beta = 0.145$, $p = .485$, 95% CI [-.266, .557]), and performance on the syllogisms task ($\beta = 0.135$, $p = .003$, 95% CI [.048, .222]).

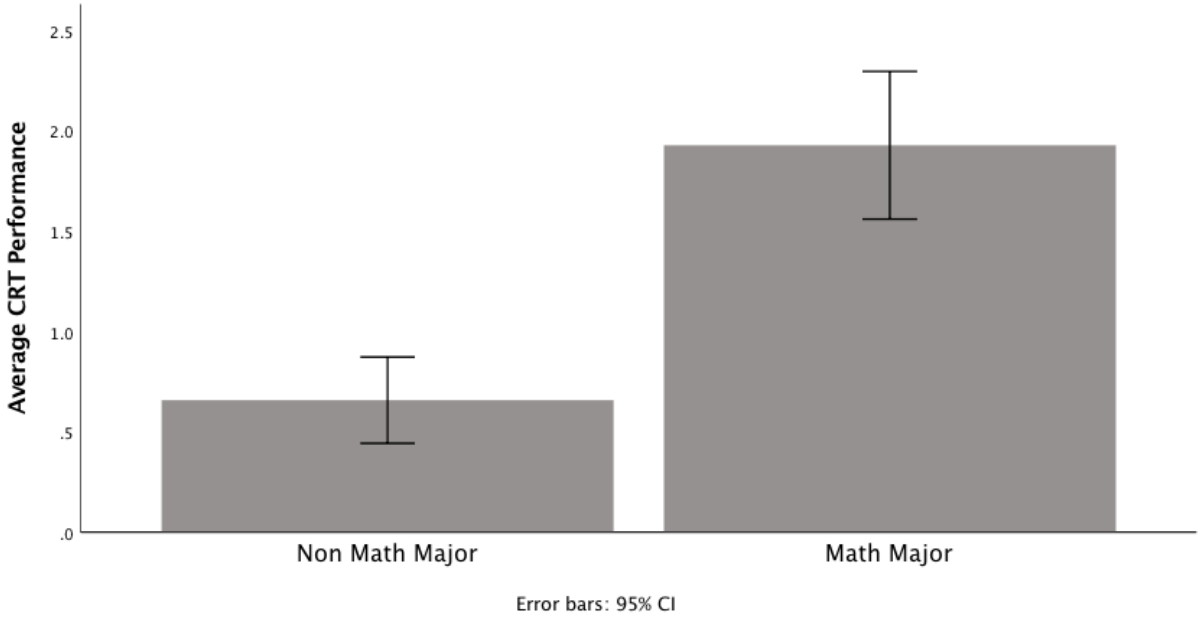

Figure 7 Mean accuracy on CRT (original) as a function of major (*N*'s = 82 and 28, respectively)

CHAPTER V

STUDY 1 DISCUSSION AND RATIONALE FOR STUDY 2

The purpose of Study 1 was to replicate the results of prior research on the relationship between analytical thinking and religiosity and to demonstrate that upper-level math majors would outperform upper-level non-math majors on a reasoning task that required math skills to solve successful. To this end, this initial study was a success. In support of Hypothesis 1a, measures of cognitive reflection, such as the original CRT and the belief bias in syllogistic reasoning task, negatively correlated with religious belief. Overall ACS, which was a sum total of three measures of cognitive reflection, also negatively correlated with religious belief (r = -.232). This result coincided neatly with the expected magnitudes of this relationship (r's between -.18 and -.33; Pennycook et al., 2016b). There were a couple of oddities that deserve explanation, however. For one, not all measures of cognitive reflection (e.g., the additional CRT) were significantly negatively related to religious belief, despite research suggesting that they should have been. This inconsistency could be explained by low sample size and power. Second, there could have been something about the math majors that enabled them able to accept supernatural religious claims and still think reflectively about math, thereby making it so that there was no relationship between their analytic cognitive styles and their religious beliefs. The explanation for such a puzzle lies beyond the scope of this current investigation, but it remains plausible that math students are fundamentally different than psychology students in how they think about religion and math.

In partial support of Hypothesis 1b, agnostics outperformed theists on some of the key measures of cognitive reflection (i.e., the original CRT, the syllogisms, and ACS). Atheists, however, did not do better than theists and agnostics on any of the measures. Previous evidence has shown that agnostics do better than theists and atheists do better than both groups (Pennycook et al., 2012; Pennycook et al., 2016b), but the overarching trend, which has shown up consistently in this study, is that disbelievers do better than believers. In other words, even though those in the "most extreme category" of disbelief (i.e., atheists) did not do substantially better than theists or agnostics, a larger representation of disbelievers (i.e., agnostics) still did. This outcome raises an interesting question: what if the relationship between religiosity and cognitive reflection were truly curvilinear? That is, maybe those who too readily accept or outright reject religious beliefs are the most intuitive thinkers, whereas those who exercise more caution with respect to both extreme positions (i.e., agnostics) think the most reflectively.

Hypothesis 1c was supported with a multiple regression model, showing that ACS is a significant predictor of religious belief, even when controlling for numerical ability, gender, age, and GPA. This finding also coincides with past research, where various other controls (such as intelligence, education, and socioeconomic status) have been included in other regression models to ensure that ACS is the best predictor of religious beliefs (Pennycook et al., 2013; Pennycook et al., 2016b; Shenhav et al., 2012).

Finally, Hypothesis 2 was supported with an independent samples *t*-test, which showed that upper-level math students considerably outperformed upper-level psychology students on the original CRT. This finding is critical, because it demonstrates that one's familiarity with a specific domain of expertise or experience (i.e., math) significantly increases one's ability to solve reasoning tasks related to that domain. The CRT is a math-based test of cognitive reflection

(Frederick, 2005). Despite the fact that the calculations required to solve the CRT questions, especially the ball-and-bat problem, are rudimentary (De Neys et al., 2013; Mastrogiorgio & Petracca, 2014), the format of the questions appeals mainly to those who normally use math to solve real world problems (Mastrogiorgio, 2015).

The upper-level psychology and math students should have been equally fooled by the CRT, at least according to prior evidence suggesting that most students from top-level universities, like Princeton University and Carnegie Mellon University, make at least one mistake on the CRT (Frederick, 2005). However, based off of the evidence in this current study, only certain types of students get fooled by the CRT. Namely, they are students who do not come from robust math-based backgrounds (as is the case for most psychology students in their final years of undergraduate education). Such a finding makes sense. Psychology students spend time learning the intricacies of personality theory, how to write in proper APA format, and how traumas to different lobes of the brain affect behavior. Their education is primary qualitative and experience-based. Math students, by contrast, are trained to solve complex calculus problems and reason through differential equations. Their education is quantitative and based in abstract logic. If, instead of solving math- and logic-based problems, the math students were given cognitive reflection problems relating to psychology, the paradigm could theoretically be reversed. The upper-level psychology students, who would have spent their education reasoning through theories of psychology, would not be so fooled by these hypothetical "Psych CRT" problems; instead, they would outperform the math students.

If the evidence of Study 1 shows anything at all, it shows that the traps set by the current measures of cognitive reflection are only ensnaring insofar as they are placed before inexperienced thinkers. An experienced mathematician does not fall prey to his or her intuitions

when thinking through an intentionally tricky math problem. An experienced psychologist who is inexperienced in math, however, does. The same could be true for any number of domains of reasoning. An experienced car mechanic does not hear the rattling sound of a lawn mower motor and mistake it for the roar of a car engine; however, an experienced businessman who is inexperienced in car mechanics might. A mother familiar with the wailings of her child might be able to consistently distinguish between a cry for food and a cry for attention; however, an experienced chess player unfamiliar in child rearing might not. Nobody would make the argument that these individuals who would fall for the "intuitive traps" in the previous examples would be cognitively handicapped, yet if these "measures" of cognitive reflection were used to assess analytic thinking in general, surely everyone who was not a mathematician, a car mechanic, or a babysitter would be labeled as relatively poor thinkers.

How well people appear to reason depends entirely on what they are asked to reason about. This argument is the prevailing motivation for Study 2. In Study 1, religious individuals were shown to be relatively poor thinkers compared to their nonreligious counterparts. However, this claim is based solely on their performance on just a few, very specific tasks of cognitive reflection. If the logic of the above paragraph is even remotely true, then surely religious individuals are not completely incapable of thinking reflectively! Instead, they just appear to not think reflectively within the domains of math-based and general-logic-based reasoning. If given the opportunity, however, they may reason quite well within a domain in which they have intimate familiarity and experience: religion. If there were a measure of cognitive reflection that tapped into the unique experiences of religious individuals, then on that measure, religious individuals would outperform nonreligious individuals.

Religion Reflection Scale

In an attempt to arouse analytical thinking among religious individuals, the Religion Reflection Scale (RSS) was created (Appendix J). The RRS is unique in that it incorporates the same dual-process thinking of other measures of cognitive reflection (i.e., its items have intuitive traps that must be overridden to solve correctly), but it does so with content that is unique to the experiences and beliefs of religious individuals. Take, for instance, the following syllogism:

> Premise 1: All nonphysical beings do not exist
> Premise 2: God is a nonphysical being
> Conclusion: God does not exist

Presumably, a religious individual who believed that God existed would find immediate conflict between this belief and the conclusion presented here. Such an individual would feel an intuitive pull to assume that the internal logic of the syllogism, or its validity, was flawed, since its conclusion would intuitively be "untrue". However, an individual with greater cognitive reflection capabilities would be able to mentally process the syllogism with greater reflection and (correctly) realize that the conclusion follows logically from the premises (provided one was, of course, asked to assume the premises were true for the sake of the exercise). As such, the RRS measures cognitive reflection within the context of religion.

Belief in God was not the only realm of religious belief or experience that was represented in the RRS. Other religious-based domains in the current, predominantly Christian, zeitgeist were incorporated, including beliefs in the creation of the world, the inerrancy of the Bible, the divinity of Jesus, and the common practice of prayer among religious believers. These themes were selected based off of their relevance among believers in Christian traditions (Hood et al., 2009), informal conversations with religious believers, the input of an expert in religious

studies (Dr. Ralph Hood), and the input of an expert in dual-process theory and religion (Dr. Gordon Pennycook).

As one will immediately notice, the RRS is reminiscent of the belief bias in syllogistic reasoning task (Study 1 and Markovits & Nantel, 1989). This use was intentional, as the syllogistic reasoning task has already been used to assess reasoning across different domains. For example, Markovits and Nantel (1989) discovered that changing the content of the syllogisms from abstract to meaningful changed performance. That is, when the syllogisms were presented with nonsensical terms, participants were more likely to separate themselves from their intuitions and correctly identify whether or not the syllogisms were sound; alternatively, when the exact same structured syllogisms were presented with meaningful terms, participants made more intuitive errors. An example follows.

> Premise 1: All things that are smoked are good for the health
> Premise 2: Cigarettes are smoked
> Conclusion: Cigarettes are good for the health

Markovits and Nantel (1989) found that, when the meaningful term "Cigarettes" was replaced by the nonsensical term "Ramadion", participants made less errors. This phenomenon, to repeat again, is known as belief bias. It is the tendency for one to erroneously judge the strength of an argument based on the believability of its conclusion rather than its merits (Evans, Newstead, & Byrne, 1993). The issue of belief bias will be addressed in the following paragraph, but the takeaway here is that the syllogistic reasoning task offers a unique opportunity for researchers to change the stimulus, or the content, of the items to suit their needs. In the past, the items have been changed from abstract to meaningful; presumably, they can also change to reflect different

areas of knowledge, like religion. The RRS is based off of the syllogistic reasoning task for this reason.

As was argued in regards to the content effect, there is a tendency for participant's familiarity with the content of a reasoning task to instead positively influence their performance (Cox & Griggs, 1982). So, too, can the content of syllogistic reasoning tasks affect performance and even attenuate belief bias. For instance, Goel and Vartanian (2011) found that when participants were given meaningful syllogisms with neutral conclusions, they fell prey to belief bias, but when they were given the same types of syllogisms with inflammatory conclusions, they made less errors. This evidence suggests that the more provocative a syllogism is, the less likely participants go with their guts when judging its conclusion. Given the prevalence and meaningfulness of religions in society (Gallup, 2017), surely one of the most provocative things one could do is present ideas that run contrary to deeply held religious beliefs. This is the point of the RRS: to evoke Type 2 reasoning among religious believers by provoking or challenging their religious-based intuitions.

Rationale for Study 2

The purpose of Study 2 was to extend the logic established in Study 1. In Study 1, math students far outscored psychology students in the math-based CRT, even when controlling for academic performance, numerical ability, and performance on the logic-based syllogistic reasoning task. This evidence suggests that there is something unique about the experiences and education of the math students that enables them to do better on the CRT than psychology students. The goal of Study 2 is to likewise show that religious individuals are reflective thinkers, though they have not been given the unique opportunity to show their prowess in a

domain that is familiar to them. Just as math would be familiar to math students, religion would be familiar to religious individuals. When given a chance to reason about religious concepts, religious individuals should do better than nonreligious individuals. The RRS acts as a means of tapping into the reflective powers of religious individuals, specifically in regards to their domain of familiarity: religion.

Study 2 Research Hypotheses

Hypothesis 1a: The CRT will negatively correlate with religious belief.

Hypothesis 1b: Measures of ACS will negatively correlate with religious belief.

Hypothesis 1c: ACS will predict religious belief, even when controlling for numeracy, gender, and reported GPA

Hypothesis 3a: Religious belief will positively correlate with religious reflection.

Hypothesis 3b: Religious familiarity will positively and uniquely predict religious reflection.

CHAPTER VI

STUDY 2 METHOD

Participants

Data were collected from 185 participants (80% female) between November 2017 and February 2018 in classroom and lab settings at UTC. Of the 21 participants who had any missing data (~11% of total), only five failed to answer more than two questions. To accommodate for this missing data, cases were deleted listwise in individual analyses (but not excluded from the dataset entirely). The average age was 21.14 years (*SD* = 4.139) and the average reported GPA was 3.29 (*SD* = 0.483) out of a possible 4.0. There were 32 freshmen (17%), 48 sophomores (26%), 59 juniors (32%), 39 seniors (21%), and six graduate students (3%). Most students identified as having a belief in a personal God or higher power (i.e., theist; 79%). The breakdown of religious identification in decreasing percentage was as follows: Christian (48.6%), Agnostic (13.5%), Protestant Christian (9.7%), Catholic Christian (9.7%), No religion (8.6%), Other (4.3%), Atheist (1.6%), Muslim (1.1%), Buddhist (0.5%), Hindu (0.5%), and Muslim (0.5%). Participation was voluntary and all students received course credit for participating. This research was approved by the Institutional Review Board of UTC (Appendix A).

Procedures

The procedures for Study 2 largely mimicked those of Study 1. Participants gave informed consent to participate and then took the battery of questionnaires in a mixed order so

that there was counterbalancing of the measures (Finley, Tang, & Schmeichel, 2015). They completed the battery on paper in either a laboratory room or a classroom and were instructed to write in or circle only one answer per item and not use calculators. They were given as much time as they needed to complete the questionnaires, though most took between 15-30 minutes on average.

Materials

Demographics

The same demographic information that was collected for Study 1 was again collected for Study 2 (Appendix B).

Religious Beliefs

The same measure from Study 1 was also used for Study 2 (Appendix C). The estimate of internal reliability for the 8-item scale assessing beliefs in supernatural religious concepts was excellent for this study as well ($\alpha = .901$).

Numeracy

The same measure from Study 1, the half-PMP (Galli et al., 2011), was also used for Study 2 (Appendix D). The estimate of internal reliability for this scale in this sample was acceptable ($\alpha = .758$).

Religious Familiarity

At the end of the religious beliefs questionnaire, there was a brief section where the participants reported their personal familiarity with religious concepts (Appendix H). These three questions came from the intellect dimension of the Centrality of Religiosity Scale (CRS; Huber & Huber, 2012). Participants indicated their level of cognitive interest or involvement in religious topics by selecting one of the five Likert-style response options. Scores for this measure can range from 1 to 5 with high scores representing higher familiarity or interest in religious topics (regardless of religious belief itself). The CRS has high internal consistency across its five dimensions (α's = .80 - .93) and high construct validity with respect to religious identity in a students' sample (r = .83) and in the international Religious Monitor (r = .73; for overview, see Huber & Huber, 2012). The three items of the religious familiarity subscale of the CRS showed good internal reliability in this sample (α = .805).

Belief Bias in Syllogistic Reasoning (Updated Version)

This task was used as a measure of non-numerical, logic-based cognitive reflection (Appendix I). It differs from the original Syllogistic Reasoning measure (Appendix E) in two marked ways: one, the items coincide with modern terms (e.g., Rockefeller was replaced by Bill Gates), and two, half of the items were replaced with alternative items whose internal validity structures were *consistent* with the believability of their conclusions (i.e., participants now could not get all eight items correct by simply guessing the "non-intuitive" responses). These changes were implemented to reduce participant confusion with outdated terms and disrupted their pattern recognition. This technique of altering the content and structure of the belief bias syllogisms has theoretical and practical precedence in the literature (Morley, Evans, & Handley,

2004; Sá, West, & Stanovich, 1999; Toplak, West, & Stanovich, 2017). Surprisingly, the estimate of internal reliability for this scale in this study was incredibly low (α = .378), despite indications that it measures similar constructs to the other cognitive measures (Table 5).

CRT (Original)

The CRT was used again as a measure of math-based cognitive reflection (Appendix F). The estimate of reliability for measure in this sample was lower than for Study 1 (α = .631).

CRT (Additional)

The same additional CRT items were included in this study (Appendix G), though this time, all four items were included instead of just the first three. The estimate of reliability of this measure for this study was incredibly poor (α = .401).

Religion Reflection Scale

This eight-item measure consists of religious-based syllogisms, and as such, it was intended to measure cognitive reflection within the domain of religion (Appendix J). The structure of the RRS is similar to that of the updated belief bias in syllogistic reasoning task. Half of the syllogisms are valid, whereas the other half are invalid. Of the valid syllogisms, half have conclusions that are believable to a religious individual (consistent items) and half have conclusions that are not believable (inconsistent items). Of the invalid syllogisms, half have conclusions that are believable to a religious individual (inconsistent items) and half have conclusions that are not believable (consistent items). Answers are coded dichotomously as correct or incorrect. Scores range from 0-8 with higher scores indicating higher religious

cognitive reflection. The measure of internal reliability for this scale, in line with the estimate of reliability for the updated general syllogisms, was poor (α = .398).

CHAPTER VII

STUDY 2 RESULTS

Descriptive statistics for the variables and a correlation matrix of the cognitive tasks can be found in Table 4 and Table 5 below. There were no marked differences between participants in the two studies in terms of descriptive statistics. Participants in this sample responded just as highly to the supernatural religious beliefs questionnaire as did those in Study 1 (M = 3.74, SD = 0.874), adding to the potential for ceiling effects. Responses to the CRT questionnaires were also highly positively skewed, adding to the potential for floor effects. The measure with the highest degree of normality, in terms of shape, was the new religious familiarity questionnaire (M = 3.22, SD = 0.891, skewness = 0.134). This finding indicates that religious familiarity and religious beliefs are two different constructs, despite their statistical relatedness (r = .271, p < .001). Interestingly, the distributions of the responses to the general syllogisms, the RRS, and ACS (which, in this case, was the addition of all three original CRT items, all four additional CRT items, and both 8-item syllogistic reasoning tasks) were more positively skewed than those in Study 1. Granted, all three of these measures were not the same as those in Study 1, but this finding is still noteworthy. Finally, all the cognitive tasks were significantly and positively related to each other, as expected (though to a lesser magnitude than Study 1).

Table 4 Descriptive Statistics of Variables in Study 2

	N	Min	Max	Mean	SD	Skewness	Kurtosis
Religious Beliefs	184	1	5	3.74	0.874	-0.812	0.556
Religious Familiarity	183	1.33	5	3.22	0.891	0.134	-0.720
Numeracy	185	2	15	9.77	3.171	-0.284	-0.839
CRT (original)	183	0	3	0.50	0.851	1.588	1.463
CRT (additional)	183	0	4	0.75	0.921	1.204	1.027
CRT Total (w/o recog.)	136	0	6	0.93	1.162	1.538	2.579
CRT Total	183	0	7	1.25	1.501	1.436	1.786
General Syllogisms	185	2	8	4.75	1.416	0.923	0.102
RRS	185	1	8	4.01	1.546	0.445	-0.085
ACS	185	5	22	9.99	3.549	1.178	1.030

Table 5 Correlation Matrix of Cognitive Tasks in Study 2 (all p's < .001)

	CRT (original)	CRT (additional)	Numeracy	General Syllogisms	RRS
CRT (original)	1				
CRT (additional)	.443	1			
Numeracy	.486	.358	1		
General Syllogisms	.320	.323	.418	1	
RRS	.336	.329	.419	.589	1

For the sake of consistency between Study 1 and Study 2, all eight items of both the general syllogisms and the RRS were included in the analysis of the following hypotheses. However, it is worth noting that there were mean differences in subtotal scores between the consistent and inconsistent syllogisms. Collectively, participants scored higher on the consistent general (M = 3.39, SD = 0.72) and religious (M = 2.64, SD = 0.87) syllogisms than the inconsistent general (M = 1.36, SD = 1.25) and religious (M = 1.37, SD = 1.19) syllogisms, indicating that when the syllogisms were constructed with logic that was consistent with the believability of their conclusions, people answered them more correctly. Likewise, when they were inconsistent with the believability of their conclusions, as was the case with the eight

original syllogisms in Study 1, people tended to answer them incorrectly. This trend remained when the scores of only believers (theists) were analyzed; however, nonbelievers (agnostics and atheists) deviated from this trend a little bit in that they scored about as high on the consistent religious syllogisms (M = 2.46, SD = 0.96) as they did on the inconsistent syllogisms (M = 2.32, SD = 1.18). Statistically, though, nonbelievers did not do any better than believers with respect to the consistent religious syllogisms, $t(181) = 1.37$, $p = .174$. It was only with respect to the inconsistent religious syllogisms that nonbelievers did better than their believing counterparts (M = 1.14, SD = 1.07), $t(181) = 5.87$, $p < .001$. These moderating differences in performance on the RRS between believers and nonbelievers suggest that the two groups are equally adept at answering the consistent religious syllogisms, but nonbelievers get an edge when answering the reflection-requiring inconsistent religious syllogisms. This finding is inconsistent with Hypothesis 3 (as will be described later), but is potentially consistent with the notion that religious familiarity plays a stronger role than religious affiliation in predicting performance on the RRS (as will also be described later).

Hypothesis 1

Performance on the original CRT negatively correlated with religious belief for Study 2 (r = -.203, p = .006) and both studies combined (r = -.208, $p < .001$; see Table 6). The additional CRT items also negatively correlated with religious belief, even when those who were familiar with at least one of the items were removed (r = -.181, p = .035, N = 136). Both of these relationships became nonsignificant in the combined analysis, though. The general belief bias syllogisms correlated negatively with religious beliefs in the first study when presented with only inconsistent conclusions (r = -.265, p = .005), but in Study 2, the relationship became

nonsignificant (r = -.088, p = .236). In the combined analysis, this relationship was also significant (r = -.182, p = .002). As will be repeated later, the RRS did not positively correlate with religious belief; instead, it significantly negatively correlated to religious belief (r = -.149, p = .043), just like most of the other measures of cognitive reflection. All measures of ACS, whether from just the six CRT items and the inconsistent syllogisms (Study 1; r = -.232, p = .015) or from the seven CRT items, the consistent and inconsistent general syllogisms, and the RRS (Study 2; r = -.189, p = .010), were significant. The combined model showed the same relationship (r = -.212, p < .001). Numeracy showed significance in Study 2 (r = -.205, p = .005), which brought the combined relationship to significance as well (r = -.140, p = .016).

Table 6 Correlations (r) Between Religious Belief and Performance on Cognitive Tasks Across Study 1 and Study 2

	CRT[1]	**CRT[2]**	**CRT[3]**	**CRT Total**	**General Syllogisms**	**RRS**	**ACS**	**Numeracy**
Study 1$_{(110)}$	-.210*	-.021	-.116$_{(67)}$	-.134	-.265**	-	-.232*	-.039
Study 2$_{(184)}$	-.203**$_{(182)}$	-.154*$_{(182)}$	-.181*$_{(136)}$	-.209**$_{(182)}$	-.088	-.149*	-.189**	-.205**
Combined$_{(294)}$	-.208**	-.090	-.131$_{(203)}$	-.172**	-.182**	-.149*$_{(184)}$	-.212**	-.140*

In Table 6, CRT[1] = Accuracy on original 3-item CRT; CRT[2] = Proportionate accuracy on additional CRT problems; CRT[3] = Proportionate accuracy on all 6 or 7 CRT problems excluding participants who indicated seeing at least one of the problems; CRT Total = Proportionate accuracy on all 6 or 7 CRT problems; General Syllogisms = Accuracy on either original or updated belief bias in syllogistic reasoning task; RRS = Religion Reflection Scale; ACS = Analytic Cognitive Style (proportionate addition of CRT Total, General Syllogisms, and RRS, where appropriate). * indicates p < .05 and ** indicates p < .01

A series of one-way ANOVA's were conducted to determine which of the three theistic groups, theist, agnostic, or atheist, performed best on the cognitive tasks. When ACS was inputted as the dependent variable, there was a statistical difference between the three groups in Study 2, $F(2, 180) = 7.395$, $p = .001$, $\eta^2 = .076$ (Figure 8). A post-hoc Tukey HSD test revealed that agnostics ($M = 11.97$, $SD = 4.151$) significantly outscored theists ($M = 9.51$, $SD = 3.158$, $p = .001$), whereas atheists ($M = 11.83$, $SD = 5.913$) did not outscore neither agnostics nor theists (p's = .996 and .240, respectively). With the following cognitive tasks as dependent variables, only agnostics outscored theists: numeracy ($F = 4.324$, $p = .015$), the total CRT excluding those who had seen at least one of the items ($F = 3.873$, $p = .023$), the general syllogisms ($F = 4.755$, $p = .010$), and the RRS ($F = 6.311$, $p = .002$). With two measures, the additional CRT items ($F = 4.211$, $p = .016$) and total CRT ($F = 4.677$, $p = .010$), only atheists outscored theists. The only model where there were no significant differences between the three groups was with the original CRT ($F = 2.518$, $p = .083$). All the others indicated that either atheists or agnostics did better than theists.

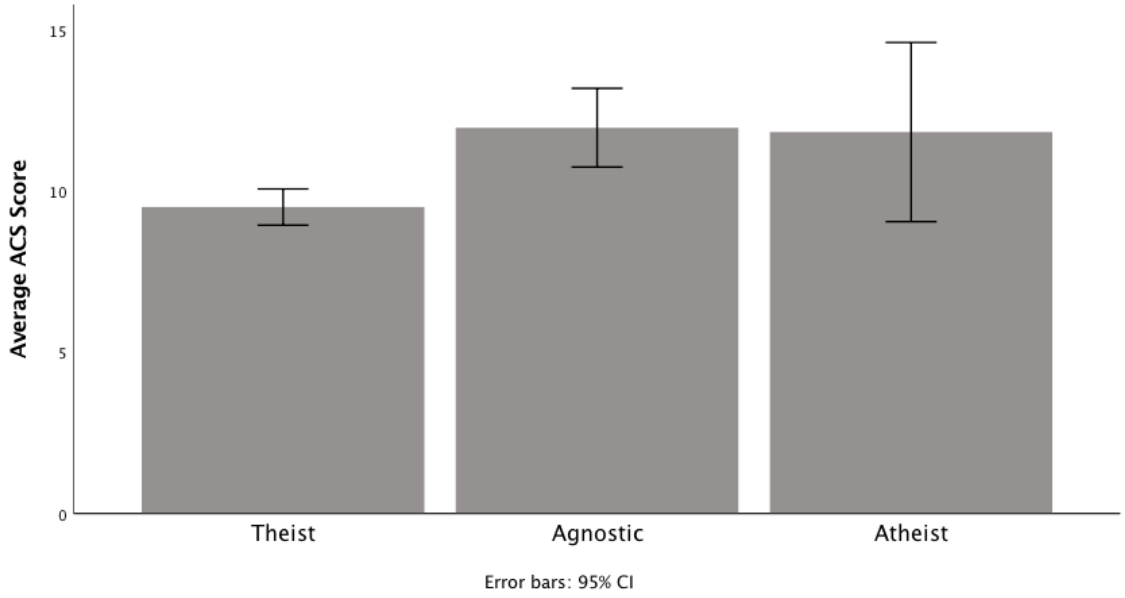

Figure 8 Mean accuracy on analytic cognitive style (ACS) as a function of theism in Study 2 (*N*'s = 146, 31, and 6, respectively)

For both studies combined, there was also a significant difference between the three theistic groups when proportionate ACS performance was inputted as the dependent variable, $F(2, 290) = 11.316$, $p < .001$, $\eta^2 = .072$ (Figure 9). A post-hoc Tukey HSD analysis revealed a consistent finding: agnostics ($M = 0.560$, $SD = 0.215$) outscored theists only ($M = 0.421$, $SD = 0.013$, $p < .001$), whereas atheists ($M = 0.527$, $SD = 0.210$) neither outscored agnostics ($p = .880$) nor theists ($p = .197$). This trend continued for numeracy ($F = 5.175$, $p = .006$), the original CRT ($F = 7.896$, $p < .001$), proportionate accuracies on all CRT items ($F = 5.486$, $p = .005$), proportionate accuracies on all CRT items when people familiar with them were removed ($F = 3.228$, $p = .042$, though no significant differences were found in post-hoc analysis) and the general belief bias syllogisms tasks ($F = 9.541$, $p < .001$). The only model that was nonsignificant was the one with only the additional CRT items ($F = 1.603$, $p = .203$).

Collectively, this evidence indicates that nonbelievers performed better than theists on multiple measures of cognitive reflection.

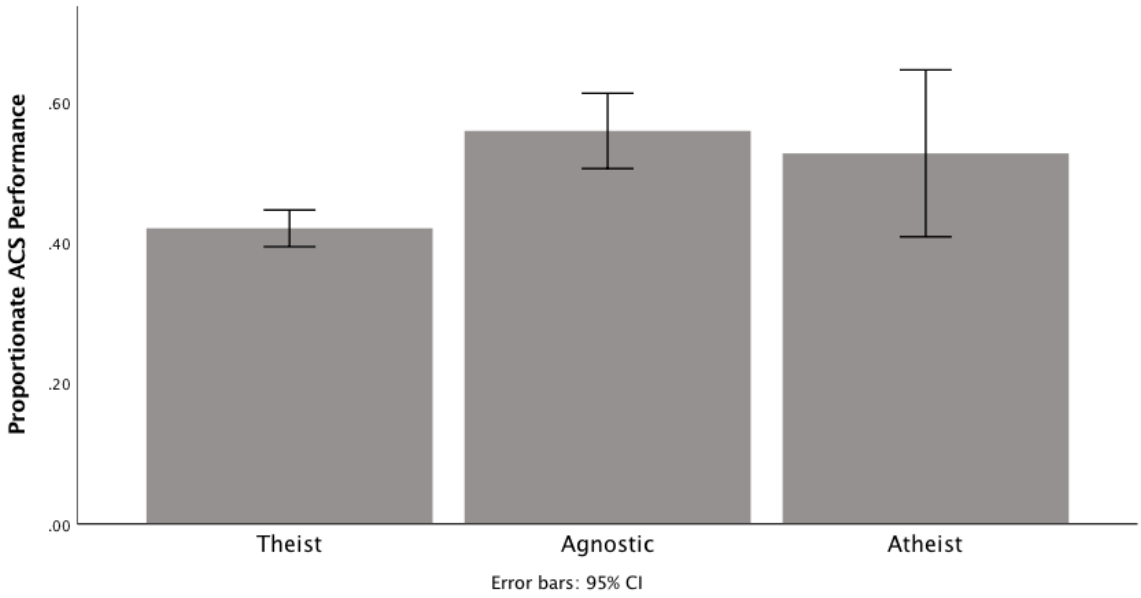

Figure 9 Mean proportional accuracy on analytic cognitive style (ACS) as a function of theism for both studies combined (*N*'s = 228, 54, and 11, respectively)

Religious belief was regressed onto ACS, numeracy, self-reported GPA, age, and gender. This model was non-significant, $F(5, 167) = 1.830$, $p = .110$. When the studies were combined, though, a stronger pattern emerged. In a multiple regression model with the same controls, $F(5, 274) = 3.849$, $p = .002$, ACS emerged as the only significant predictor of religious belief ($\beta = -1.069$, $p = .002$, 95% CI [-1.750, -.389]; all other p's > .2). Collectively, this evidence indicates that analytic cognitive style is a powerful and unique predictor of religious belief.

Hypothesis 3

The RRS was moderately and significantly related to the other measures of cognitive reflection (Table 5), providing some construct validation that it was, at least to some extent, measuring cognitive reflection. The lack of multicollinearity with these scales indicates, however, that the RRS was also measuring some other construct. This other construct was not religious belief, as indicated by the negative relationship between the RRS and religious belief (r = -.149, p = .043; see Table 6).

Religious familiarity positively and significantly correlated with RRS performance (r = .151, p = .041, Table 7), suggesting that the RRS was drawing out more of a cognitive reflection component of religiosity than simply a belief-based one. The religious familiarity questionnaire was also not simply a measure of a need for cognition (Cacioppo & Petty, 1982), as such a measure of self-reported willingness to think should have correlated positively with *all* the measures of cognitive reflection, not just the religious-based one. However, when controlling for religious beliefs (β = -0.145, p = .189), numeracy (β = 0.064, p = .067), performance on all seven CRT items (β = 0.143, p = .047), and performance on the general syllogisms (β = 0.502, p < .001), religious familiarity did not come out as a unique predictor of RRS performance (β = 0.189, p = .077), $F(5, 175) = 24.599$, $p < .001$.

Table 7 Correlations (r) Between Religious Familiarity and Other Measures of Cognition

Religious Beliefs	**CRT[1]**	**CRT[2]**	**CRT[3]**	**CRT Total**	**General Syllogisms**	**RRS**	**Numeracy**
.271**	.002	.032	.132	.023	.132	.151*	.038

In Table 7, CRT^1 = Accuracy on original 3-item CRT; CRT^2 = Accuracy on additional CRT problems; CRT^3 = Accuracy on all 7 CRT problems excluding participants who indicated seeing at least one of the problems; CRT Total = Accuracy on all 7 CRT problems; General Syllogisms = Accuracy on updated belief bias in syllogistic reasoning task; RRS = Religion Reflection Scale. * indicates $p < .05$ and ** indicates $p < .01$

CHAPTER VIII

STUDY 2 DISCUSSION

The purpose of Study 2 was to replicate the results of prior research on the relationship between analytical thinking and religiosity and to demonstrate that a novel measure of cognitive reflection, the Religion Reflection Scale, would instead positively correlate with religious belief. To this end, this study was only partially successful. In support of Hypothesis 1, all of the classic measures of cognitive reflection, including the original CRT items, the additional CRT items, and the "new" general syllogisms task, negatively correlated with religious belief. Overall ACS, which was a sum total of the CRT items, the general syllogisms, and the RRS, was negatively correlated with religious belief. This relationship manifested across both studies (r = -.212), which coincides with expectations (r's between -.18 and -.33; Pennycook et al., 2016b). Mean comparisons indicated that nonbelievers (i.e., agnostics and atheists) outperformed believers (i.e., theists) on nearly all the measures of cognitive reflection. Oddly, in Study 2, ACS did not significantly predict religious belief when other standard demographics were controlled for in a multiple regression model, but across both studies, the expected relationship was found again. Taken together, this evidence suggests that nonreligious individuals are collectively better analytical thinkers than their religious counterparts.

The main purpose of Study 2, though, was to assess the efficacy of the RRS and how well it could measure the reflective prowess of religious individuals in their domain of expertise, religion. The RRS moderately positively correlated with the other scales of cognitive reflection,

indicating that it was measuring cognitive reflection *and* at least some other component(s). To this extent, the RRS showed some promise as a measure of cognitive reflection within the domain of religion. In contrast to the expectation of Hypothesis 3a, however, the RRS and religious belief were negatively related to each other, indicating that the additional component that the RRS was measuring was not religious belief. Indeed, in a one-way ANOVA comparing the performance of theists, agnostics, and atheists on the RRS, agnostics outscored theists, as would be expected if the RRS were just another measure of cognitive reflection.

In anticipation of this result, the three-item religious familiarity scale was included in Study 2. It was reasoned that not all religious believers contemplate their religious beliefs to the same extent. For instance, a religious believer could potentially believe in their religious beliefs by fiat without much contemplation, whereas a nonbeliever could show great interest in studying religious concepts and culture without believing any of its precepts. The low correlation between religious belief and familiarity confirmed that the two concepts are quite different. This distinction is important, because scores on the RRS were not meant to show who was religious and who was not; rather, the RRS was meant to positively correlate with *familiarity* with religion. Therefore, the nonbeliever familiar with religion could hypothetically outperform the disinterested believer by virtue of simply engaging in the topics more.

In support of Hypothesis 3b, religious familiarity was positively related to RRS performance. This relationship was also unique insofar as religious familiarity did not significantly correlate with any of the other measures of cognitive reflection (although this finding is somewhat curtailed by the lack of retained significance of religious familiarity in a multiple regression model with other controls). This evidence indicates that the attempt to make a measure of cognitive reflection within the domain of religion was at least partially successful.

No, religious individuals did not suddenly outperform nonbelievers on this measure, as was anticipated, but those who expressed interest in investigating religious concepts were more likely to outscore their less interested counterparts, meaning that being familiar with religion gave some individuals a unique advantage on the RRS (above and beyond even their beliefs).

CHAPTER IX

GENERAL DISCUSSION

Across two studies, a central question was tested: are nonbelievers truly more reflective thinkers than religious individuals? The results from the replication portions of these studies indicated that, yes, agnostics and atheists consistently outperform theists on classic measures of cognitive reflection, such as the CRT and the belief bias in syllogistic reasoning task. The relationship between religious belief and analytic cognitive style (ACS) was always significantly negative, with the magnitude hovering just below the middle point of the expected -.18 to -.33 range (Pennycook et al., 2016b). This relationship retained significance, even when controlling for other important variables like numeracy, self-reported GPA, age, and gender. In terms of group differences, atheists did not outscore agnostics and theists in all measures (and, in fact, they only outscored theists in some), but in most measures of cognitive reflection, agnostics consistently did better than theists. Overall, this replication was successful, with nonbelievers (mainly agnostics) outperforming believers on nearly all counts of cognitive reflection.

However, the main goal of this research was to determine if there is an element of familiarity that influences performance on analytic thinking measures. This hypothesis was explored across the two studies. The purpose of Study 1 was to see if those who were more familiar with math (i.e., upper-level math majors) would outperform those who were presumably just as intellectually capable, but relatively not as familiar with math (i.e., upper-level psychology majors) on the math-based CRT items. Indeed, math students were not fooled by the

intuitive solutions to the math problems. They significantly outperformed psychology students on this measure (even after controlling for GPA, numerical ability, and performance on the general logic syllogisms). This evidence indicates that there is something special about being a math major that makes them predisposed to scoring well on the math-based CRT, a measure which does not require sophisticated math knowledge to answer correctly. It is suggested that this extra factor is familiarity. Presumably, if these same students were given a similar battery of questions in which dual-process thinking was tested with questions relating to psychology, the upper-level psychology students would do better than their math student counterparts by virtue of being more familiar with *that* domain of knowledge. In short, Study 1 demonstrated, at least initially, that familiarity is an influential variable that predicts performance on measures of cognitive reflection. Therein lies the main purpose of Study 2.

The goal of Study 2 was to tease out the unique reflection capabilities of religious individuals by getting them to reason with content related to *their* domain of familiarity: religion. This objective was tested by crafting and analyzing a novel measure of cognitive reflection in the domain of religion, the Religion Reflection Scale (RRS). Despite showing low internal reliability, the RRS correlated neatly with the other measures of cognitive reflection in the study. These correlations were significant enough to suggest that the RRS was tapping into cognitive reflection, but they were low enough to suggest that it was measuring cognitive reflection through some other means. This other means was hypothesized to be religious familiarity. Unfortunately, the relationship between religious beliefs and performance on the RRS was still negative. A groups comparison between theists, agnostics, and atheists confirmed that agnostics (and to a lesser extent, atheists) still did better on this measure of cognitive reflection than theists.

As was stated before, however, the RRS was meant to relate to familiarity with religion, not just having religious beliefs. One could presumably be a religious believer and attend religious services but not mentally engage with the concepts themselves; likewise, one could be a nonbeliever but show great interest in learning more about the histories, beliefs, and cultures of other religions. It was hypothesized that, beyond merely religious beliefs, religious *familiarity* would positively predict performance on the new RRS. Indeed, self-reported interest in religion positively related to RRS performance. Interestingly, this relationship was the only one that was significant; religious familiarity did not significantly relate to performance on any of the other secular measures of cognitive reflection, including the parallel general syllogistic task. This evidence suggests that, at least on a small level, being interested in and familiar with religion can enable one to succeed on a measure of cognitive reflection within the domain of religion.

Limitations

There was a lack of representative sample sizes among the different religious groups. By virtue of sampling from a predominantly Christian university in the southern United States, this demographic pitfall was unavoidable. However, if more atheists and agnostics could have been included (with numbers at least on par with the theists), there could have been less ceiling effects with respect to the supernatural religious beliefs questionnaire and a more striking effect of religious belief category on cognitive reflection. These tweaks in sampling may not have substantially altered the overall results of the studies, but they could have enhanced their generalizability.

Another limitation was the lack of systematic alterations between Study 1 and Study 2. As was stated before, the changes to the syllogistic reasoning task should not have significantly

altered the constructs they were measuring, but they did, in fact, alter some critical elements of the scale. Primarily, the estimate of reliability of the updated syllogistic reasoning task (and by extension, the RRS) was extremely poor in Study 2. This downside was somewhat alleviated by the significant correlations between the updated syllogism tasks and all of the other measures of cognitive reflection, but the fact still remains that the syllogisms in Study 2 may not have been reliable. Another non-systematic alteration (and potential limitation) between the studies was the inclusion of the fourth additional CRT item in Study 2. It was originally removed in Study 1 in order to lower the number of questions in the study (reducing participant fatigue) and limit the types of questions that were presented so that only short, easy-to-read questions were included. For the sake of consistency with past research and for the sake of variability in the CRT measures as a whole, though, the lengthy and convoluted item made a reappearance in Study 2. Without repeating these studies over again with these exact alterations, it is difficult to tell whether any differences in results between Study 1 and Study 2 were due to the differences in the samples or the novelties of the measures themselves. Fortunately, the primary findings still replicated, at least to a significant enough degree to suggest that these alterations in the measures did not substantially diminish their quality in measuring cognitive reflection.

Another limitation, one that has more to do with the utility or generalizability of the studies and less to do with the design of the research itself, was the volatile nature of the results. This limitation is specifically in regards to the religious familiarity questionnaire in Study 2. While the only measure of cognitive reflection that was significantly related to religious familiarity was the RRS, the magnitude of this relationship was quite small and just hovering on the line of significant (r = .151, p = .041). With a slightly different sample, this relationship could have been pushed into the non-significant category.

Beyond volatility, the magnitude of this correlation is also troubling, because it is so small. What does a correlation of .152 with a sample of less than 200 people mean in the real world? That among small groups of people, being familiar with religion *only slightly* gives one an edge in solving religious-based cognitive reflection problems? How useful is that information? Moreover, how useful is the negative correlation between religious beliefs and ACS when it only exists between -.18 and -.33 (-.2 in this study)? Yes, measures of cognitive reflection may be better predictors of religious belief than intelligence, but only to a statistically significant degree (Pennycook, 2014), not to a degree that appears to have any significant impact on the real world. There is a deeper limitation of meaningfulness that exists in this field of inquiry. Take, for example, the results of Study 2. Out of 23 items, atheists and agnostics together answered around 12, whereas theists answered around 10 correctly (Figure 8). Let us say, though, that instead of 23 tricky questions on a psychology study, this sample was instead asked to make 23 real world life decisions. These results imply that, on average, nonbelievers would think reflectively on about half of those decisions, whereas religious individuals would think reflectively on slightly less than half. Already, the outcomes are suggesting that these differences in performance between religious groups are marginal, even if they are statistically significant. Also, who is to say that the decisions on which nonbelievers thought reflectively were the best decisions? A religious individual could think reflectively over which doctor to choose but use intuition when selecting a reliable meal at a restaurant, whereas a nonreligious individual could haphazardly decide to drive under the influence of marijuana or alcohol but deliberate obsessively over what type of outfit to wear at a dinner party. Just by thinking reflectively on *slightly* more issues than religious individuals, nonreligious individuals do not suddenly get a prize for living the safest life, the healthiest life, or the best life. If anything, the

best decision one could make could be to don the religious beliefs of their culture, as being alone in a minority group can be ostracizing and damaging to one's physical and emotional wellbeing (Galen, 2015; McBrayer, 2014). The statistically significant, but small, correlations between peoples' purported religious beliefs and their performance on cognitive tasks does not indicate much or account for much in the real world. Granted, this research still has value. The fact that nonbelievers outscore believers on cognitive reflection tasks in many studies is interesting, but it bears repeating that these differences are marginal and perhaps not indicative of much. It is certainly not the case that religious people are substantially cognitively inferior to nonbelievers.

Future Directions

The easiest route to extending this current research would be to find or create the measure of cognitive reflection on which religious individuals would do better than nonbelievers. The key may lie in the Wason Card Selection Task (Wason, 1966). As was stated before, when participants were no longer asked to assess an abstract rule (i.e., if a card has a vowel on one face, it must have an even number on the opposite face), but were instead asked to police a familiar law (i.e., if a person is drinking a beer, they must be over 21 years old), most of them answered the new problem correctly (Cox & Griggs, 1982; Davies et al., 1995). What accounted for this change? For one, the content changed. Instead of vowels and numbers, a familiar law involving alcohol and age was introduced. When creating the RRS, it was assumed that content was all that would matter. Just by reasoning about religious-based items, religious individuals should have done better than those who were presumably not as familiar with those concepts (just as math students did better than psychology students on the math-based problems). However, a second, critical component to this paradigm is missing, and that component is

context. In the updated Wason card selection task, participants were not just asked to think about alcohol and ages; instead, they were prompted to *police* the rule about alcohol and ages and find the cheaters. In other words, they were asked to think about familiar items in a more familiar context as well. Pollard and Evans (1987) tested this hypothesis by placing participants in one of four conditions. The first condition, the Abstract condition, had participants solving the Wason card selection task with the normal abstract rules (vowel/number). The second condition, the Policeman condition, was the same updated task where participants were asked to pretend to be policemen on duty and check whether or not people in a bar were over the legal age (alcohol/age). As expected, participants in the Policeman condition significantly outperformed participants in the Abstract condition. In the remaining two conditions, something surprising happened. In the third condition, the Beers condition, participants were asked to check a drinking rule without legal enforcement (alcohol/age). In other words, in this condition, the content was familiar to participants, and they were still asked to assess the rule, but they were asked to do so outside of the familiar context of policing the rule as an on-duty cop (presumably to bust cheaters). In the fourth condition, the Security condition, participants were asked to pretend to be security guards who were policing people with ID cards that had strings of numbers and letters on both sides (vowel/number). In other words, in this condition, the context of policing was familiar to the participants but the content of the ID cards they were policing was abstract and unfamiliar. Amazingly, participants in these last two conditions, where only content or context was familiarized, *did not improve their performance*. They did about as well as those in the Abstract control condition, whereas only those in the Policeman condition had significantly higher success. This evidence indicates that, in order for a reasoning task to effectively increase

the scores of any one group, it must incorporate familiar content *and* familiar context related to that group.

The RRS was only different from the other measures of belief bias in syllogistic reasoning in terms of content, not context. Outcomes could change if instead the RRS was crafted to uniquely pander to the experiences and challenges of religious individuals, both in terms of content and context. Perhaps instead of asking participants to assume the premises of the religious syllogisms as true and work out the logic internally, they were instead asked to evaluate the arguments (i.e., the presented religious syllogisms) of other religious individuals with different beliefs than theirs. With these new directions, the context would involve policing the arguments of other religious individuals, not just assessing the merits of unrelatable syllogisms. With relatable content and context, religious individuals could potentially outscore nonreligious individuals on this measure. Or, at the very least, the current difference between them could be attenuated. In a related study, participants of high cognitive ability outscored participants with low cognitive ability when they were both asked to solve syllogistic reasoning problems according to the boilerplate "assume the following is true" method of presenting syllogisms; however, when the instructions were changed to be more situational and involve non-dichotomous responding, the differences in performances between high and low cognitive groups disappeared (Evans et al., 2010). This evidence indicates that merely changing the instructions, or the context in which reasoning problems are presented, can reduce the effects of belief bias. If the instructions were changed to pander to one group over another, the effects could be even more profound.

Over the past century (Larson & Witham, 1997; Leuba, 1916; Stirrat & Cornwell, 2013) scholars have found that most eminent scientists (80-90%) do not believe in a God or higher

power. This finding is incredibly interesting and leads to a whole new question that has yet to be resolved: if disbelief is so intimately tied to critical thought, what accounts for the outliers? In other words, what separates the ~10% believing faction of scientists from the vast majority of their disbelieving peers? Surely both groups can reason critically! In order to become a prominent scientist, one must analyze data, think of solutions to problems, and produce noteworthy results, all of which require one to deeply process information. Cognitive ability cannot be a reliable predictor of religiosity at this top level. Herein lies a case where domain specificity has merits. Perhaps some upper-level scientists intuitively believe in God and devote their reflective powers to their science. Put differently, scientists may be able to distinguish which domains they choose to reason about. For most, they reason about their science and about religion and find the latter wanting, whereas for the others, they reason about their science and intuitively accept their religions without much contemplation. Their reasoning is, thus, contingent on which domain they are engaged in.

Such an explanation is still wanting, as it fails to account for the deeply religious theologians who clearly have strong reflective capabilities in the realms of metaphysics, mathematics, and philosophy. If reflective thought and religiosity were truly so incompatible, then how come great thinkers like Blaise Pascal, Thomas Aquinas, and St. Augustine slipped through the cracks? What makes it possible for someone to reflect deeply about religion and still maintain religious beliefs?

Conclusion

This research deserves consideration in the field of analytical thinking and religiosity, because it represents an initial attempt to challenge the current paradigm, a paradigm which may

be partially incorrect. Are nonbelievers truly more reflective thinkers than religious believers? The initial answer is yes. When given classic measures of cognitive reflection, like the CRT and belief bias in syllogistic reasoning syllogisms, agnostics and atheists outperform theists (Pennycook et al., 2016b; though for a cross-cultural challenge of this claim, see Gervais et al., 2017). However, this finding is limited in scope in terms of the instruments used. Yes, religious believers fall prey to intuitive pitfalls in math-based and general-knowledge, logic-based questions, but does this evidence mean they are not reflective thinkers overall? What would happen if they were asked to reason about questions related to the most fundamental aspects of their identities? Admittedly, the answer is still disappointing: the attempt in this thesis to show that religious individuals could outperform nonreligious individuals on a measure of religious cognitive reflection failed (though religious familiarity could have influenced this relationship).

Regardless of this failure, there is a valuable takeaway from this research that deserves consideration, and this takeaway is that not all conceivable measures of cognitive reflection have been exhausted. Finding or creating such a measure that draws out the reflective capabilities of religious individuals may take years and may seem as futile as finding a needle in a haystack, but it remains possible that this measure could exist. Such a finding would bring incredible nuance to the table, further establishing the continuing truth that reality is not black-and-white. Instead, reality is fluid, and truth is contingent on context. No longer could it be said that nonbelievers are more reflective thinkers than their believing counterparts. Instead, the prevailing nuanced claim would be that nonbelievers are more reflective than believers with respect to some, *but not all,* measures of cognitive reflection. Again, it remains to be seen what kind of measure this could be, but the finding of such a measure would be invaluable to those seeking to understand the refined truth.

REFERENCES

Allport, G. W., & Ross, J. M. (1967). Personal religious orientation and prejudice. *Journal of Personality and Social Psychology, 5*(4), 432-443.

Beller, S., & Spada, H. (2003). The logic of content effects in propositional reasoning: The case of conditional reasoning with a point of view. *Thinking and Reasoning, 9*(4), 335-378.

Binet, A., Beaunis, H. É., des Bancels, J. L., Simon, T., & Piéron, H. (1903). *L'année Psychologique* (Vol. 9). Paris, France: Presses Universitaires de France.

Binet, A., & Simon, T. (1916). *The development of intelligence in children: The Binet-Simon Scale*. Baltimore, MD: Williams & Wilkins Company.

Boring, E. G. (1923). Intelligence as the tests test it. *New Republic, 35*, 35-37.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*, 116-131.

Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119*(2), 197-253.

Cliteur, P. (2009). The definition of atheism. *Journal of Religion and Society, 11*, 1-23.

Cosmides, L., & Tooby, J. (1992). *The adapted mind: Evolutionary psychology and the generation of culture*. New York, NY: Oxford University Press.

Cox, J. R., & Griggs, R. A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition, 10*(5), 496-502.

Cragun, R. T., Hammer, J. H., & Nielsen, M. (2015). The Nonreligious–Nonspiritual Scale (NRNSS): Measuring everything from Atheists to Zionists. *Secularism and Nonreligion, 2*(3), 36-53.

Daniel, D. B., & Klaczynski, P. A. (2006). Developmental and individual differences in conditional reasoning: Effects of logic instructions and alternative antecedents. *Child Development, 77*(2), 339-354.

Davies, P. S., Fetzer, J. H., & Foster, T. R. (1995). Logical reasoning and domain specificity: A critique of the social exchange theory of reasoning. *Biology and Philosophy, 10*, 1-37.

De Lisi, R., & Staudt, J. (1980). Individual differences in college students' performance on formal operations tasks. *Journal of Applied Developmental Psychology, 1*(3), 201-208.

De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin & Review, 20*(2), 269-273.

Díaz-Lefebvre, R. (2004). Multiple intelligences, learning for understanding, and creative assessment: Some pieces to the puzzle of learning. *Teachers College Record, 106*(1), 49-57.

Doane, M. J., & Elliott, M. (2014). Perceptions of discrimination among atheists: Consequences for atheist identification, psychological and physical well-being. *Psychology of Religion and Spirituality, 7*(2), 1-12.

Edgell, P., Gerteis, J., & Hartmann, D. (2006). Atheists as "other": Moral boundaries and cultural membership in American society. *American Sociological Review, 71*(2), 211-234.

Evans, J. S. B., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning, 11*(4), 382-389.

Evans, J. S. B., Handley, S. J., Neilens, H., & Over, D. (2010). The influence of cognitive ability and instructional set on causal conditional inference. *Quarterly Journal of Experimental Psychology, 63*(5), 892-909.

Evans, J. S. B., Newstead, S. E., & Byrne, R. M. (1993). *Human reasoning: The psychology of deduction*. East Sussex, UK: Psychology Press.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223-241.

Finley, A. J., Tang, D., & Schmeichel, B. J. (2015). Revisiting the relationship between individual differences in analytic thinking and religious belief: Evidence that measurement order moderates their inverse correlation. *PloS One, 10*(9), e0138922.

Finucane, M. L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000). The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making, 13*(1), 1-17.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25-42.

Galen, L. (2015). Atheism, wellbeing, and the wager: Why not believing in God (with others) is good for you. *Science, Religion and Culture, 2*(3), 54-69. doi:10.17582/journal.src/2015/2.3.54.69

Galli, S., Chiesi, F., & Primi, C. (2008). The construction of a scale to measure mathematical ability in psychology students: An application of the Rasch Model. *TPM-Testing, Psychometrics, Methodology in Applied Psychology., 15*(1), 1-16.

Galli, S., Chiesi, F., & Primi, C. (2011). Measuring mathematical ability needed for "non-mathematical" majors: The construction of a scale applying IRT and differential item functioning across educational contexts. *Learning and Individual Differences, 4*(21), 392-402.

Gallup. (2017). Religion. Retrieved from http://www.gallup.com/poll/1690/religion.aspx

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences* New York, NY: Basic Books.

Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. New York, NY: Basic Books.

Gardner, H., & Moran, S. (2006). The science of multiple intelligences theory: A response to Lynn Waterhouse. *Educational Psychologist, 41*(4), 227-232.

Gervais, W. M. (2015). Override the controversy: Analytic thinking predicts endorsement of evolution. *Cognition, 142*, 312-321.

Gervais, W. M., van Elk, M., Xygalatas, D., McKay, R., Aveyard, M., Buchtel, E. E. K., . . . Riekki, T. (2017). *Analytic atheism: A cross-culturally weak and fickle phenomenon?* Retrieved from psyarxiv.com/92r8x

Goel, V., & Vartanian, O. (2011). Negative emotions can attenuate the influence of beliefs on logical reasoning. *Cognition and Emotion, 25*(1), 121-131.

Gottfredson, L. S. (1997). Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography. *Intelligence, 24*(1), 13-23.

Harrison, V. S. (2006). The pragmatics of defining religion in a multi-cultural world. *International Journal for Philosophy of Religion, 59*(3), 133-152.

Hood, R. W., Hill, P. C., & Spilka, B. (2009). *The psychology of religion: An empirical approach* (4th ed.): Guilford Press.

Huber, S., & Huber, O. W. (2012). The centrality of religiosity scale (CRS). *Religions, 3*(3), 710-724.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*(9), 697-720.

Koehler, D. J., & James, G. (2010). Probability matching and strategy availability. *Memory & Cognition, 38*(6), 667-676.

Larson, E. J., & Witham, L. (1997). Scientists are still keeping the faith. *Nature, 386*(6624), 435-436.

Legg, S., & Hutter, M. (2007). A collection of definitions of intelligence. *Frontiers in Artificial Intelligence and Applications, 157*, 1-12.

Leuba, J. (1916). *The belief in God and immortality. A psychological, anthropological and statistical study*. Boston, MA: Sherman, French & Company.

Lewis, G. J., Ritchie, S. J., & Bates, T. C. (2011). The relationship between intelligence and multiple domains of religious belief: Evidence from a large adult US sample. *Intelligence, 39*(6), 468-472.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25*(4), 361-381.

Lynn, R., Harvey, J., & Nyborg, H. (2009). Average intelligence predicts atheism rates across 137 nations. *Intelligence, 37*(1), 11-15.

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition, 17*(1), 11-17.

Mastrogiorgio, A. (2015). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6*, 1-2.

Mastrogiorgio, A., & Petracca, E. (2014). Numerals as triggers of System 1 and System 2 in the 'bat and ball' problem. *Mind & Society, 13*(1), 135-148.

Mata, A., Ferreira, M. B., & Sherman, S. J. (2013). The metacognitive advantage of deliberative thinkers: A dual-process perspective on overconfidence. *Journal of Personality and Social Psychology, 105*(3), 353.

Matarazzo, J. D., & Wiens, A. N. (1977). Black Intelligence Test of Cultural Homogeneity and Wechsler Adult Intelligence Scale scores of black and white police applicants. *Journal of Applied Psychology, 62*(1), 57-63.

McBrayer, J. P. (2014). The wager renewed: Believing in God is good for you. *Science, Religion and Culture, 1*(3), 130-140.

McNiel, N. D. (1975). IQ tests and the black culture. *Phi Delta Kappan, 57*(3), 209-210.

Mettetal, G., Jordan, C., & Harper, S. (1997). Attitudes toward a multiple intelligences curriculum. *Journal of Educational Research, 91*(2), 115-122.

Morley, N. J., Evans, J. S. B., & Handley, S. J. (2004). Belief bias and figural bias in syllogistic reasoning. *Quarterly Journal of Experimental Psychology Section A, 57*(4), 666-692.

Norenzayan, A., & Gervais, W. M. (2013). The origins of religious disbelief. *Trends in Cognitive Sciences, 17*(1), 20-25.

Nyborg, H. (2009). The intelligence–religiosity nexus: A representative study of white adolescent Americans. *Intelligence, 37*(1), 81-93.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization, 72*(1), 147-152.

Pachur, T., & Spaar, M. (2015). Domain-specific preferences for intuition and deliberation in decision making. *Journal of Applied Research in Memory and Cognition, 4*(3), 303-311.

Pennycook, G. (2014). Evidence that analytic cognitive style influences religious belief: Comment on Razmyar and Reeve (2014). *Intelligence, 43*, 21-26.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2014a). Cognitive style and religiosity: The role of conflict detection. *Memory & Cognition, 42*(1), 1-10.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2014b). The role of analytic thinking in moral judgements and values. *Thinking & Reasoning, 20*(2), 188-214.

Pennycook, G., Cheyne, J. A., Barr, N., Koehler, D. J., & Fugelsang, J. A. (2015a). On the reception and detection of pseudo-profound bullshit. *Judgment and Decision Making, 10*(6), 549-563.

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2013). Belief bias during reasoning among religious believers and skeptics. *Psychonomic Bulletin & Review, 20*(4), 806-811.

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2016a). Is the Cognitive Reflection Test a measure of both reflection and intuition? *Behavior Research Methods, 48*(1), 341-348.

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition, 123*(3), 335-346.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015b). Everyday consequences of analytic thinking. *Current Directions in Psychological Science, 24*(6), 425-432.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015c). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34-72.

Pennycook, G., & Ross, R. M. (2016). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 7*, 1-4.

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

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development, 15*(1), 1-12.

Pollard, P., & Evans, J. S. B. (1987). Content and context effects in reasoning. *American Journal of Psychology, 100*(1), 41-60.

Reisberg, D. (2013). *Cognition: Exploring the science of the mind* (5th ed.). New York, NY: W. W. Norton & Company.

Roberts, M. J. (2008). *Integrating the mind: Domain general versus domain specific processes in higher cognition*. Hove, England: Psychology Press.

Sá, W. C., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91*(3), 497-510.

Selfridge, O. G. (1955). *Pattern recognition and modern computers.* Paper presented at the Western Joint Computer Conference.

Selfridge, O. G. (1959). *Pandemonium: A paradigm for learning.* Paper presented at the Mechanisation of Thought Processes, London.

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

Silver, C. F. (2013). *Atheism, agnosticism, and nonbelief: A qualitative and quantitative study of type and narrative.* (Doctoral dissertation). Retrieved from https://scholar.utc.edu/cgi/viewcontent.cgi?article=1536&context=theses

Sinayev, A., & Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6*, 1-16.

Skinner, B. F. (1948). "Superstition" in the pigeon. *Journal of Experimental Psychology, 38*, 168-172.

Stanovich, K. E. (2012). On the distinction between rationality and intelligence: Implications for understanding individual differences in reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *Oxford Handbook of Thinking and Reasoning* (pp. 343-365). Oxford, UK: Oxford University Press.

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127*(2), 161-188.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23*(5), 645-665.

Sternberg, R. J., Castejón, J. L., Prieto, M. D., Hautamäki, J., & Grigorenko, E. L. (2001). Confirmatory factor analysis of the Sternberg Triarchic Abilities Test in three international samples: An empirical test of the triarchic theory of intelligence. *European Journal of Psychological Assessment, 17*(1), 1-16. doi:10.1027//1015-5759.17.1.1

Stieger, S., & Reips, U.-D. (2016). A limitation of the Cognitive Reflection Test: Familiarity. *PeerJ, 4*, e2395. doi:10.7717/peerj.2395

Stirrat, M., & Cornwell, R. E. (2013). Eminent scientists reject the supernatural: a survey of the Fellows of the Royal Society. *Evolution: Education and Outreach, 6*(33), 1-5.

Streib, H., Silver, C. F., Csöff, R.-M., Keller, B., & Hood, R. W. (2011). *Deconversion: Qualitative and quantitative results from cross-cultural research in Germany and the United States of America* (Vol. 5). Göttingen, Germany: Vandenhoeck & Ruprecht.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*(7), 1275-1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*(2), 147-168.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2017). Real-world correlates of performance on heuristics and biases tasks in a community sample. *Journal of Behavioral Decision Making, 30*(2), 541-554.

Vadeboncoeur, I., & Markovits, H. (1999). The effect of instructions and information retrieval on accepting the premises in a conditional reasoning task. *Thinking & Reasoning, 5*(2), 97-113.

Visser, B. A., Ashton, M. C., & Vernon, P. A. (2006). Beyond g: Putting multiple intelligences theory to the test. *Intelligence, 34*, 487-502.

Wason, P. C. (1966). *Reasoning* (B. M. Foss Ed.). Harmondsworth, Middlesex, England: Penguin Books.

Waterhouse, L. (2006a). Inadequate evidence for multiple intelligences, Mozart effect, and emotional intelligence theories. *Educational Psychologist, 41*(4), 247-255.

Waterhouse, L. (2006b). Multiple intelligences, the Mozart effect, and emotional intelligence: A critical review. *Educational Psychologist, 41*(4), 207-225.

Welsh, M., Burns, N., & Delfabbro, P. (2013). *The Cognitive Reflection Test: How much more than numerical ability*. Paper presented at the 35th Annual Conference of the Cognitive Science Society, Austin, TX.

White, J. (2008). Illusory intelligences? *Journal of Philosophy of Education, 42*(3-4), 611-630.

Williams, R. L. (1972). *The BITCH-100: A culture-specific test*. Paper presented at the American Psychological Association Annual Convention, Honolulu, Hawaii.

Woodhead, L., Fletcher, P., Kawanami, H., & Smith, D. (2002). *Religions in the modern world: Traditions and transformations*. London and New York: Routledge.

Zinnbauer, B. J., & Pargament, K. I. (2002). Capturing the meanings of religiousness and spirituality: One way down from a definitional Tower of Babel. *Research in the Social Scientific Study of Religion, 13*, 23-54.

Zinnbauer, B. J., Pargament, K. I., Cole, B., Rye, M. S., Butter, E. M., Belavich, T. G., . . . Kadar, J. L. (1997). Religion and spirituality: Unfuzzying the fuzzy. *Journal for the Scientific Study of Religion*, 549-564.

Zuckerman, M., Silberman, J., & Hall, J. A. (2013). The relation between intelligence and religiosity: A meta-analysis and some proposed explanations. *Personality and Social Psychology Review, 17*(4), 325-354.

APPENDIX A

IRB APPROVAL LETTER

Institutional Review Board
Dept. 4915
615 McCallie Avenue
Chattanooga, TN 37403-2598
Phone: (423) 425-5867
Fax: (423) 425-4052
instrb@utc.edu
http://www.utc.edu/irb

MEMORANDUM

TO: Jacob Strimaitis **IRB # 16207**
Dr. Ralph Hood

FROM: Lindsay Pardue, Director of Research Integrity
Dr. Amy Doolittle, IRB Committee Chair

DATE: 1/25/2017

SUBJECT: IRB #16207: Domain Specificity of the Cognitive Reflection Task: A Look at Numerical Ability and Religion

The IRB Committee Chair has reviewed and approved your application and assigned you the IRB number listed above. You must include the following approval statement on research materials seen by participants and used in research reports:

***The Institutional Review Board of the University of Tennessee at Chattanooga (FWA00004149) has approved this research project #* 16207.**

Annual Renewal. All approved research is subject to UTC IRB review, at least once a year. Please visit our website (http://www.utc.edu/research-integrity/institutional-review-board/forms.php) for the Form B (continuation / change / completion form) that you will need to complete and submit if your project remains active and UTC IRB approval needs to be renewed for another year. Unless your research moves in a new direction or participants have experienced adverse reactions, then renewal is not a major hurdle. You as Principal Investigator are responsible for turning in the Form B on time (2 weeks before one year from now), and for determining whether any changes will affect the current status of the project. When you complete your research, the same change/completion form should be completed indicating project termination. This will allow UTC's Office of Research Integrity to close your project file.

Please remember to contact the IRB immediately and submit a new project proposal for review if significant changes occur in your research design or in any instruments used in conducting the study. You should also contact the IRB immediately if you encounter any adverse effects during your project that pose a risk to your subjects.

For any additional information, please consult our web page http://www.utc.edu/irb or email instrb@utc.edu.

Best wishes for a successful research project.

APPENDIX B

DEMOGRAPHICS

What is your age (in years)?

With what gender do you most identify?

What is your ethnicity?

What is your current major?

What is your academic classification (i.e. sophomore, junior, senior, etc.)?

What is your current GPA?

APPENDIX C

RELIGIOUS BELIEFS

	I strongly disagree (1)	I disagree (2)	I don't know (3)	I agree (4)	I strongly agree (5)
There is a life after death. (1)	❍	❍	❍	❍	❍
There is a heaven where people who have led good lives are eternally rewarded. (2)	❍	❍	❍	❍	❍
There is a hell where people who have led bad lives and die without being sorry are eternally punished. (3)	❍	❍	❍	❍	❍
Religious miracles occur. (4)	❍	❍	❍	❍	❍
Angels are active in the world. (5)	❍	❍	❍	❍	❍
Demons are active in the world. (6)	❍	❍	❍	❍	❍
People have an immaterial soul, a part of themselves that is beyond their merely physiological and physical properties. (7)	❍	❍	❍	❍	❍
There is a devil/Satan. (8)	❍	❍	❍	❍	❍

What sort of God, if any, do you believe in?

- ❍ A personal God
- ❍ God as an impersonal force
- ❍ A God who created everything, but does not intervene in human affairs
- ❍ Don't know whether or not any Gods exist
- ❍ Don't know whether any Gods exist, and no one else does either
- ❍ I don't believe in Gods of any sort
- ❍ I believe that God does not exist

With which of the following do you identify?

- ❍ Agnostic
- ❍ Atheist
- ❍ Baha'i
- ❍ Buddhist
- ❍ Chinese Traditional
- ❍ Christian
- ❍ Christian (specifically Catholic)
- ❍ Christian (specifically Protestant)
- ❍ Hindu
- ❍ Humanist
- ❍ Jewish
- ❍ Muslim
- ❍ No religion
- ❍ Other not listed
- ❍ Sikh
- ❍ Taoist

APPENDIX D

NUMERACY

1. What is the result of $\frac{2}{3} + \frac{3}{4}$?

A) $\frac{5}{7}$ B) $\frac{6}{12}$ C) $\frac{17}{12}$ D) $\frac{5}{12}$

2. What is the result of $(-8) \div (-2)$?

A) 4 B) – 4 C) 16 D) 0.25

3. The double of $\frac{3}{4}$ is?

A) $\frac{6}{8}$ B) $\frac{3}{2}$ C) $\frac{9}{16}$ D) $\frac{3}{8}$

4. Set A consists of all odd numbers between 8 and 20, whereas set B consists of all numbers less than 10. How many elements do the two sets share?

A) None B) 1 C) 2 D) 3

5. Place five red marbles, three green marble and two yellow marbles in a bag. Draw one red marble out of the bag. Without replacing the marble, what is the probability of drawing one green marble in a second drawing?

A) $\frac{3}{10}$ B) $\frac{3}{9}$ C) $\frac{4}{10}$ D) $\frac{4}{9}$

6. What is the result of the following equation: (5 + 3)x = 0 ?

A) x = 5 – 3 B) x = 0 C) x = $\frac{5}{3}$ D) x = 5 + 3

7. Which of the following relations is true?

A) 0.01 x 0.01 < 0.01
B) 0.01 x 0.01 = 0.01
C) 0.01 x 0.01 > 0.01
D) 0.01 x 0.01 = 0.1

8. Knowing that xy = 3, which of the following is true?

A) $y = \frac{3}{x}$ B) $y = 3 - x$ C) $y = 3x$ D) $\frac{xy}{3} = 0$

9. The value of 0.05 is:

A) Lower than 0
B) Within – 1 and 0
C) Higher than 0.1
D) Within 0 and 1

10. By rolling a dice and tossing a coin at the same time, how many combinations are possible?

A) 6 + 2 B) 6 x 2 C) 6 + 6 D) 6 x 6

11. In a school there are 125 students. Out of those students, the number who passed an exam are 116. The percentage of students that failed is:

A) 6.2% B) 9% C) 7.8% D) 7.2%

12. There are four aces in a deck of 40 cards. What is the probability of drawing one ace from this deck?

A) $\frac{1}{40}$ B) $\frac{1}{4}$ C) $\frac{4}{10}$ D) $\frac{4}{40}$

13. When drawing 2 cards (without replacement) from a standard deck of playing cards, the probability of drawing two kings is:

A) Equal to the probability of drawing one king
B) Higher than the probability of drawing one king
C) Lower than the probability of drawing one king
D) None of the above

14. The value of –0.98 is:

A) Within –1 and 0
B) Lower than –1
C) Higher than 0
D) Within –2 and –1

15. Which is the result of the following operations? $[(13 - 10)^2 + (17 - 20)^2 + (10 - 10)^2]$

A) 0 B) 18 C) 6 D) 9

APPENDIX E

BELIEF BIAS IN SYLLOGISTIC REASONING

Directions: You are going to receive a series of eight problems. You must decide whether the stated conclusion *follows logically* from the premises or not

You must suppose that the premises are all true and limit yourself only to the information contained in these premises. This is very important.

For each problem, decide if the given conclusion *follows logically from the premises*. Circle YES if, and only if, you judge that the conclusions can be derived *unequivocally* from the given premises, otherwise circle NO

Premise 1: All things that are smoked are good for the health
Premise 2: Cigarettes are smoked
Conclusion: Cigarettes are good for the health

YES NO

Premise 1: All unemployed people are poor
Premise 2: Rockefeller is not unemployed
Conclusion: Rockefeller is not poor

YES NO

Premise 1: All flowers have petals
Premise 2: Roses have petals
Conclusion: Roses are flowers

YES NO

Premise 1: All animals with four legs are dangerous
Premise 2: Poodles are not dangerous
Conclusion: Poodles do not have four legs

YES NO

Premise 1: All mammals walk
Premise 2: Whales are mammals
Conclusion: Whales walk

YES NO

Premise 1: All eastern countries are communist
Premise 2: Canada is not an eastern country
Conclusion: Canada is not communist

YES NO

Premise 1: All animals love water
Premise 2: Cats do not like water
Conclusion: Cats are not animals

YES NO

Premise 1: All things that have a motor need oil
Premise 2: Automobiles need oil
Conclusion: Automobiles have motors

YES NO

APPENDIX F

CRT (ORIGINAL)

1. A ball and a bat cost $1.10 in total. The bat costs $1.00 more than the ball. How much does the ball cost? _____ cents

2. If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? _____ minutes

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? _____ days

APPENDIX G

CRT (ADDITIONAL)

1. If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? _____ days

2. Jerry received both the 15^{th} highest and 15^{th} lowest mark in the class. How many students are in the class? _____ students

3. A man buys a pig for $60, sells it for $70, buys it back for $80, and sells it finally for $90. How much has he made? _____ dollars

4. Simon decided to invest $8,000 in the stock market one day early in 2008. Six months after he invested, on July 17, the stocks he had purchased were down 50%. Fortunately for Simon, from July 17 to October 17, the stocks he had purchased went up 75%. At this point, Simon has: a. broken even in the stock market, b. is ahead of where he began, c. has lost money

APPENDIX H

RELIGIOUS FAMILIARITY

	Not at all	Not very much	Moderately	Quite a bit	Very much so
How interested are you in learning more about religious topics?	O	O	O	O	O

	Never	Rarely	Occasionally	Often	Very often
How often do you think about religious issues?	O	O	O	O	O
How often do you keep yourself informed about religious questions through radio, television, internet, newspapers, or books?	O	O	O	O	O

APPENDIX I

BELIEF BIAS IN SYLLOGISTIC REASONING: UPDATED VERSION

Directions: You are going to receive a series of eight problems. You must decide whether the stated conclusion *follows logically* from the premises or not

You must suppose that the premises are all true and limit yourself only to the information contained in these premises. This is very important.

For each problem, decide if the given conclusion *follows logically from the premises*. Circle YES if, and only if, you judge that the conclusions can be derived *unequivocally* from the given premises, otherwise circle NO

Premise 1: All flowers have petals
Premise 2: Roses have petals
Conclusion: Roses are flowers

YES NO

Premise 1: All birds have wings
Premise 2: Crows are birds
Conclusion: Crows have wings

YES NO

Premise 1: All meat products can be eaten
Premise 2: Apples can be eaten
Conclusion: Apples are meat products

YES NO

Premise 1: All European countries are cold
Premise 2: Mexico is not cold
Conclusion: Mexico is not a European country

YES NO

Premise 1: All things that are smoked are good for the health
Premise 2: Cigarettes are smoked
Conclusion: Cigarettes are good for the health

YES NO

Premise 1: All humans have brains
Premise 2: Cats are not humans
Conclusion: Cats do not have brains

YES NO

Premise 1: All unemployed people are poor
Premise 2: Bill Gates is not unemployed
Conclusion: Bill Gates is not poor

YES NO

Premise 1: All animals with four legs are dangerous
Premise 2: Poodles are not dangerous
Conclusion: Poodles do not have four legs

YES NO

APPENDIX J

RELIGION REFLECTION SCALE

Directions: You are going to receive a series of eight problems. You must decide whether the stated conclusion *follows logically* from the premises or not

You must suppose that the premises are all true and limit yourself only to the information contained in these premises. This is very important.

For each problem, decide if the given conclusion *follows logically from the premises*. Circle YES if, and only if, you judge that the conclusions can be derived *unequivocally* from the given premises, otherwise circle NO

Premise 1: All things are created
Premise 2: The world is a thing
Conclusion: The world was created

YES NO

Premise 1: All books with errors were written by people
Premise 2: The Bible was written by people
Conclusion: The Bible has errors

YES NO

Premise 1: All religious beliefs are sacred
Premise 2: Secular beliefs are not sacred
Conclusion: Secular beliefs are not religious beliefs

YES NO

Premise 1: All prophets pray
Premise 2: Common religious followers are not prophets
Conclusion: Common religious followers do not pray

YES NO

Premise 1: All nonphysical beings do not exist
Premise 2: God is a nonphysical being
Conclusion: God does not exist

YES NO

Premise 1: All good things are loved by God
Premise 2: Evil-doing is not good
Conclusion: Evil-doing is not loved by God

YES NO

Premise 1: All beings made in God's image have souls
Premise 2: Humans have souls
Conclusion: Humans are made in God's image

YES NO

Premise 1: All divine beings are nonphysical
Premise 2: Jesus was not nonphysical
Conclusion: Jesus was not a divine being

YES NO

VITA

Jacob Strimaitis was born in Santa Cruz, CA, to the parents of Lenny and Lanette Strimaitis. He has one older brother, Luke. He attended Poplar Grove Elementary School in Franklin, TN, and continued to Franklin High School in the same town. After graduation, he attended Lee University in Cleveland, TN, where he developed a fascination for science and human behavior. Also, by being in a religious institution, he was exposed to many different worldviews that significantly impacted his own beliefs and perspectives. He received a Bachelors of Science in May 2016 in Biochemistry with dual minors in Psychology and Religion. Jacob immediately went on to pursue a Masters of Science degree in Research Psychology from the University of Tennessee at Chattanooga and is expected to graduate with this degree in May 2018. He is intending to take a short hiatus from academia in order to use his skillset and love for science in an industrial setting, though he hopes to pursue a PhD after gaining this more applied research experience.