An Experimental Examination of Religious Teaching and Thought-Action Fusion

Johanna Anita Younce
younce94@gmail.com

Follow this and additional works at: https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations

Recommended Citation
https://huskiecommons.lib.niu.edu/allgraduate-thesesdissertations/7800
ABSTRACT

AN EXPERIMENTAL EXAMINATION OF RELIGIOUS TEACHING
AND THOUGHT-ACTION FUSION

Johanna A. Younce, MA
Department of Psychology
Northern Illinois University, 2019
Kevin D. Wu, Director

Scrupulosity is an understudied symptom dimension of obsessive-compulsive disorder (OCD) that involves moral- and religious-based obsessions and compulsions. There are several different underlying cognitive beliefs that have been shown to be related to the development and maintenance of OCD, but one—moral thought-action fusion (TAF-moral)—has been highly associated with scrupulosity. TAF-moral is the belief that thoughts are morally equivalent to actions, including thoughts that are intrusive in nature. It has been found that individuals from different religious traditions display different levels of TAF-moral and display different strengths of association between TAF-moral and religiosity. However, these differences have not been examined experimentally and thus causality has not been determined. Since a highly punishing concept of God is closely associated with scrupulosity, the purpose of this study was to examine causal links between religion and TAF-moral by testing two opposing God concepts (punishing or forgiving) as a cause of heightened TAF-moral in certain religious traditions.

Ninety-nine participants met inclusion criteria for the study. Participants were randomly assigned to be primed with either a forgiving-God (FG) or punishing-God (PG) concept before completing a negative thought induction task paradigm meant to induce TAF-moral. It was found
that there were no significant differences between priming conditions on perceived moral
wrongness scores, and the frequency of neutralizing behavior did not depend on condition. Both
groups were significantly more anxious following the negative thought induction compared to
immediately prior to it, and the groups showed marginally significant differences on anxiety; the
PG group displayed higher overall anxiety compared to the FG group. However, there was no
interaction between time and condition on anxiety scores. Whereas the results showed only
mixed support for the hypotheses in terms of reaching a priori statistical significance levels, it is
important to note that all findings were in the expected direction. Limitations include the
inclusion of participants with extreme baseline God concepts, a low dose for priming God
concepts, the use of manipulation check questions that did not allow for excluding participants
who were not adequately primed, and difficulties with coding neutralizing behavior.
AN EXPERIMENTAL EXAMINATION OF RELIGIOUS TEACHING 
AND THOUGHT-ACTION FUSION

BY
JOHANNA YOUNCE
©2019 Johanna Younce

A THESIS SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE
MASTER OF ARTS

DEPARTMENT OF PSYCHOLOGY

Thesis Director:
Kevin D. Wu
ACKNOWLEDGEMENTS

First and foremost, I would like to thank my thesis director and mentor, Dr. Kevin Wu, for offering me substantial support without which I would not have completed this document and for encouraging me, guiding me, and reviewing previous drafts of this document to make this version the best it could be. I would also like to thank the undergraduate research assistants who assisted me in collecting the data for this study: Megan Boehning, Dylan Pencakowski, Vince Pozzie, Jordan Kerbis, Ashlyn Thurman, Madison Ballard, and Alejandra Arce. Finally, I would like to thank my parents, Mark and Hanna Younce, for their financial and emotional support and for teaching me the value of science and education.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>1</td>
</tr>
<tr>
<td>Scrupulosity</td>
<td>2</td>
</tr>
<tr>
<td>Obsessional Beliefs</td>
<td>6</td>
</tr>
<tr>
<td>Religion and TAF-Moral</td>
<td>10</td>
</tr>
<tr>
<td>Thought Induction Task</td>
<td>15</td>
</tr>
<tr>
<td>Priming Concepts of God</td>
<td>22</td>
</tr>
<tr>
<td>The Current Study</td>
<td>25</td>
</tr>
<tr>
<td><strong>METHOD</strong></td>
<td>30</td>
</tr>
<tr>
<td>Participants</td>
<td>30</td>
</tr>
<tr>
<td>Measures</td>
<td>31</td>
</tr>
<tr>
<td>Procedure</td>
<td>36</td>
</tr>
<tr>
<td><strong>RESULTS</strong></td>
<td>40</td>
</tr>
<tr>
<td>Data Cleaning</td>
<td>40</td>
</tr>
<tr>
<td>Preliminary Analyses</td>
<td>43</td>
</tr>
<tr>
<td>Manipulation Check</td>
<td>43</td>
</tr>
<tr>
<td>Primary Analyses</td>
<td>45</td>
</tr>
</tbody>
</table>

LIST OF TABLES........................................................................................................................................... v

LIST OF APPENDICES........................................................................................................................................... vi
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. DISCUSSION</td>
<td>47</td>
</tr>
<tr>
<td>Major Findings</td>
<td>47</td>
</tr>
<tr>
<td>Limitations</td>
<td>49</td>
</tr>
<tr>
<td>Future Directions</td>
<td>54</td>
</tr>
<tr>
<td>Conclusion</td>
<td>57</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>58</td>
</tr>
<tr>
<td>APPENDICES</td>
<td>66</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Skew and Kurtosis Standardized Scores</td>
<td>42</td>
</tr>
<tr>
<td>2. Descriptive Statistics and Mean Comparisons</td>
<td>44</td>
</tr>
</tbody>
</table>
# LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. DEMOGRAPHIC QUESTIONNAIRE</td>
<td>66</td>
</tr>
<tr>
<td>B. SANTA CLARA STRENGTH OF RELIGIOUS FAITH QUESTIONNAIRE</td>
<td>68</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION

Obsessive-compulsive disorder (OCD) is a mental disorder characterized by obsessions (persistent, unwanted intrusive thoughts, urges, or images that cause anxiety or distress) and compulsions (repetitive behaviors or mental acts that the individual feels an urge to perform in order to reduce anxiety; American Psychiatric Association [APA], 2013). The 12-month prevalence of OCD is about 1% (APA, 2013; Kessler, Chiu, Demler, & Walters, 2005; Ruscio, Stein, Chiu, & Kessler, 2010), and an estimate of lifetime prevalence is 2.3%, based on DSM-IV criteria (Ruscio et al., 2010). OCD is a serious disorder, with suicide attempts occurring in up to 46.3% of those diagnosed (Angelakis, Gooding, Tarrier, & Panagioti, 2015). Additionally, those with OCD are more likely to have presentations that are classified as serious (as compared to mild or moderate) than those diagnosed with any anxiety disorder (50.6% versus 22.8%; Kessler et al., 2005). Any disorder with such high rates of suicidality and severity must be researched thoroughly in order to best understand and treat the disorder.

To be diagnosed with OCD, obsessions and compulsions must take up at least one hour per day or cause clinically significant distress or impairment (APA, 2013). Virtually all individuals with OCD experience both obsessions and some type of compulsion (Leonard & Riemann, 2012; M. Williams et al., 2011). Compulsions typically are performed to decrease the distress from obsessional thoughts or to prevent a feared event from occurring. They can include overt behaviors, such as hand washing, or covert mental activities, such as silent praying in
response to having a bad thought. Compulsions may be irrational; they need not be realistically connected to the feared event; even if they somehow are related, they are excessive or unreasonable (e.g., spending several hours washing one’s hands) and typically counter-productive (e.g., washing hands to the point of injury; APA, 2013). Many compulsions involve neutralizing strategies, that is, physical or mental rituals performed in order to negate the perceived effects of an intrusive thought. For example, someone with OCD might attempt to cancel out a thought about something bad happening to a loved one by picturing the loved one as safe and engaged in an enjoyable activity.

OCD is heterogeneous with respect to symptom themes. Although some individuals with OCD obsess about contracting disease and compulsively wash their hands, others may repeatedly flip a light switch a certain number of times to prevent harm from happening to a close friend. It is common to have multiple kinds of obsessions and compulsions. Typical symptom themes involve checking, contamination, ordering and symmetry, religion and morality, and harm (APA, 2013). Checking and contamination symptoms have been studied more than other types (M. Williams, Mungo, Franklin, & Faber, 2013), but it is difficult to generalize those findings to other obsessive-compulsive (OC) themes, given key differences. For example, there is evidence to suggest that individuals with differing symptom themes respond differently to treatment (M. Williams et al., 2013). Thus, explicit attention to other symptom themes is warranted.

Scrupulosity

Religious and moral obsessions and compulsions constitute a theme of OCD termed “scrupulosity.” Individuals presenting with scrupulosity symptoms may experience excessive doubt about whether they have sinned, unusual concerns about minor aspects of their religion,
and fear of punishment from God (Abramowitz & Jacoby, 2014; Nelson, Abramowitz, Whiteside, & Deacon, 2006). Examples of scrupulous obsessions include persistent fear that one has or will commit a sin (including unintentionally or even without realizing it), fears that a religious practice was not performed correctly, intrusive blasphemous images (e.g., Jesus on the cross with an erection), doubts about one’s religious piety or that one does not love God enough, and fear of God’s punishment (Abramowitz & Jacoby, 2014; Olatunji, Abramowitz, Williams, Connelly, & Lohr, 2007). Examples of scrupulous compulsions include neutralizing strategies such as excessive prayer, repetition of religious rituals until done “correctly,” excessive confession, and seeking reassurance from clergy (Abramowitz & Jacoby, 2014; Olatunji et al., 2007).

There also is heterogeneity in the presentation of scrupulosity itself. Some individuals with scrupulous OCD turn toward religious observance to relieve obsessional fear, but others avoid religious stimuli and environments as a way to alleviate their fears (Abramowitz & Jacoby, 2014). This means that an individual with scrupulosity symptoms might avoid going to religious services (an important part of many religions) because of an excessive concern about a perceived minor aspect of religious observance. For example, Abramowitz and Jacoby (2014) describe an Orthodox Jewish client who experienced intrusive images of herself desecrating the Torah (Jewish written law, operationalized within the Judaic Bible or as the first five books of the Old Testament of the Christian Bible) in her synagogue. To relieve her guilt and anxiety, she avoided the synagogue, called the synagogue to check, and sought reassurance from the rabbi that she had not acted on the intrusive thoughts by mistake. Many of the behaviors that scrupulous individuals fear are sinful (e.g., swallowing saliva during a fast) are quickly pardoned
by religious authorities and others of the same religion (Abramowitz & Jacoby, 2014). Thus, scrupulosity differs from typical religious observance.

In one study examining 425 individuals with OCD, religion was found to be the fifth most common obsessional theme, with 5.9% identifying it as their primary obsession (Foa et al., 1995). In a survey of individuals with OCD, it was found that those with primary religious obsessions were less able to recognize that their OC beliefs were unreasonable as compared to those with primary contamination, symmetry/ordering, hoarding, or somatic and “miscellaneous” obsessions ($\chi^2_{[7]} = 29.56, p < .01$; Tolin, Abramowitz, Kozak, & Foa, 2001). Those with primary religious obsessions reported higher distortions in the perception of their surroundings and their own bodies compared to those with primary contamination obsessions ($\chi^2_{[7]} = 15.88, p < .05$). Those with primary religious obsessions reported more magical ideation—that is, belief in invalid forms of causation—than those with primary contamination, symmetry/ordering, and somatic obsessions ($\chi^2_{[7]} = 24.82, p < .01$). Thus, compared with other types of OCD symptoms, scrupulosity is associated with poorer insight, higher perceptual distortion, and higher magical ideation. Of note, it has been shown that poor insight may result in poorer treatment outcomes (Foa, Abramowitz, Franklin, & Kozak, 1999).

In a study focusing on the effects of insight on treatment outcomes, 11 individuals with OCD participated in an intensive 3-week exposure and response-prevention treatment program (Foa et al., 1999). Participants’ insight was measured by a questionnaire that focused on their certainty that feared consequences would occur. Those with some amount of certainty that fears would not occur and those who were uncertain were placed in a low-fixity group, and those who endorsed being mostly or completely certain feared consequences would occur were in the high-
fixity group. Analyses revealed that the low-fixity group experienced a larger reduction of OC symptoms than did the high-fixity group, \( t_{(9)} = 8.27, p < .03 \) (Foa et al., 1999). Despite these relatively unique features of scrupulosity, it remains a comparatively understudied presentation of OCD.

**Religion and OCD**

A relationship between religiosity (strength of religious faith) and OC symptomatology has been established in the literature. Scrupulosity also has been positively associated with strength of religious faith \( (r = .36, p < .01; \) Abramowitz, Huppert, Cohen, Tolin, & Cahill, 2002). Religiosity has been associated with scrupulosity beyond depression, anxiety, and stress \( (\beta = .42, p < .001; \) Gonsalvez, Hains, & Stoyles, 2010). In a sample of Iranian Muslims, a highly religious group scored higher on a measure of scrupulosity than a low-religious group (Shams & Milosevic, 2013). Protestants also have been found to be more scrupulous when compared to people who are not religious (Nelson et al., 2006). Such findings may seem commonsensical, but it is important to establish the associations empirically.

Finally, the association between general OC symptomatology (as opposed to specific symptom dimensions, such as scrupulosity) and religiosity also has been examined. Rassin and Koster (2003) found a significant correlation between religiosity and OC symptoms among Catholics \( (r = .31, p < .05) \) and an even larger correlation among Protestants \( (r = .41, ns \) due to the small sample). Higher religiosity has been associated with more obsessional symptoms in Muslims \( (r = .14-.24) \), Protestants \( (F_{(2, 220)} = 3.95) \), and Christians \( (r = .30; \) Abramowitz, Deacon, Woods, & Tolin, 2004; Inozu, Karanci, & Clark, 2012; Inozu, Ulukut, Ergun, & Alcolado, 2014). Among adult Israeli Jews, participants who reported having grown up to become more
religious than their parents displayed more OC symptomatology than those who had grown up to become less religious than their parents ($t_{34} = 2.20, p = .03, d = .76$; Zohar, Goldman, Calamary, & Mashiah, 2005).

We know that religiosity, scrupulosity, and nonscrupulous OCD symptomatology all are intercorrelated, but it is critical to recognize that most people who are religious do not have OCD. The cognitive model of OCD offers a framework within which to explore other factors involved in these relationships.

**Obsessional Beliefs**

Research on OCD primarily is guided by a cognitive framework. The cognitive model of OCD emerged from Beck’s (1976) cognitive specificity model that initially was developed for emotional disorders (i.e., depression). This model posits that certain beliefs become incorporated into how individuals view themselves, the world, and the future and lead to the development of mental disorders. Furthermore, certain cognitive beliefs or processes are unique to certain disorders and produce specific mood states. For example, instead of suggesting that depression causes individuals to view themselves negatively, the cognitive specificity model asserts that a consistent negative view of oneself leads to depression and depressed feelings.

Although Beck (1976) offered a cognitive analysis of OC symptoms, Salkovskis (1985) argued that it did not properly differentiate OC problems from problems with anxiety in that the proposed cognitions involved in the disorders were very similar. Salkovskis (1985) expounded on Beck’s (1976) work by developing a cognitive theory specific to OCD. Underlying this model is that unwanted intrusive thoughts are a common experience (Rachman & de Silva, 1978), yet very few individuals develop OCD. These intrusive thoughts develop into pathological
obsessions due to the appraisal or interpretation of the thoughts—not because of the thoughts per se (Salkovskis, 1985). In other words, intrusive thoughts are stimuli that individuals react to or interpret in a certain, automatic way based on their underlying beliefs. Although most people are able to dismiss an intrusive thought without experiencing undue negative consequences, individuals who go on to be diagnosed with OCD experience more distress about these thoughts and an urge to respond to them because the thoughts are perceived as important or somehow indicative of real life. Salkovskis’s (1985) model proposed that the dysfunctional interpretation of intrusive thoughts involves exaggerated personal responsibility for harm that could come to oneself or others. This includes the belief that one is responsible for preventing harm to oneself or others and that failing to prevent harm would be catastrophic. It also includes exaggerated responsibility for one’s own thoughts and the content therein. Thus, individuals with OCD become distressed that they might cause harm; they perform neutralizing behaviors or other forms of rituals in an attempt to prevent the content of the intrusive thoughts from occurring.

Rachman (1997) added to this theory by suggesting that another dysfunctional belief in OCD is concerned with the significance that individuals give to intrusive thoughts. Specifically, he suggested that individuals with OCD engage in a cognitive bias termed thought-action fusion (TAF), in which one’s thoughts and actions are considered intimately connected and/or equivalent. This included the belief that having an intrusive thought about an event increases the probability that the event actually will occur and the belief that having an intrusive thought is morally equivalent to performing the behavior involved in the thought. For example, a caring father may have an intrusive thought about molesting his child while changing a diaper; he may go on to place personal significance on this thought, believe he is a child molester, become
distressed, and take measures to prevent this from happening (e.g., reject his diaper changing role, avoid being alone in a room with the child). He believes that having this thought reveals something about himself (i.e., that he is a child molester) because all thoughts are important and meaningful. Rachman (1997) also suggests that TAF is connected to exaggerated responsibility in that increased responsibility for a harmful event leads to higher estimates of the probability of the event. Conversely, TAF may lead to an increase in feelings about one’s personal responsibility for harmful events.

Further work on this model by the Obsessive Compulsive Cognitions Working Group (OCCWG, 1997) focused on identification and measurement of the dysfunctional cognitive beliefs that contribute to OCD. They originally identified six rational (i.e., non-empirical) belief dimensions that contribute to dysfunctional appraisals of intrusive thoughts, including beliefs about exaggerated responsibility and importance of thoughts. After further examining these belief domains using empirical (i.e., factor analytic) methods, it was established that the belief domains loaded onto three factors, not six (OCCWG, 2005). One of the three belief dimensions involves exaggerated responsibility and overestimation of threat. Individuals with OCD tend to believe that they are responsible for preventing harmful events associated with intrusive thoughts, and they tend to overestimate the probability that harm will occur. Perfectionism and intolerance of uncertainty is a second set of beliefs characterized by high, rigid standards and a strong need for certainty. Third, individuals with OCD tend to place high importance on and a need to control their thoughts (ICT). They tend to believe that the presence of intrusive thoughts is meaningful and representative of reality in some way, and they feel a strong need to get rid of intrusive thoughts. All of these dysfunctional beliefs are believed to be critical to the experience
of distress about intrusive thoughts, lead to increased attention on the thoughts, make the intrusion more accessible, and ultimately lead to obsessions / compulsive behavior in an attempt to reduce the intrusions or decrease the responsibility associated with them, which maintains the dysfunctional beliefs (Salkovskis, 1999).

**Thought-Action Fusion**

TAF has been the focus of many studies in the scrupulosity literature (Abramowitz & Jacoby, 2014; Berman, Abramowitz, Pardue, & Wheaton, 2010; Berman, Stark, Ramsey, Cooperman, & Abramowitz, 2014; Berman, Wheaton, & Abramowitz, 2013; Deacon, Vincent, & Zhang, 2013; Inozu et al., 2014; Nelson et al., 2006; Olatunji et al., 2007; Rassin & Koster, 2003; Siev & Cohen, 2007). TAF commonly is measured using the Thought-Action Fusion Scale (TAFS; Shafran, Thordarson, & Rachman, 1996). Within this instrument, there are two major forms of TAF: (a) TAF-likelihood, the belief that thinking about an event will make that event more likely to occur, and (b) TAF-moral, the belief that thinking about a behavior or event is morally equivalent to performing the action (Shafran et al., 1996). TAF-likelihood can be further split into TAF-likelihood other, which involves beliefs about events happening to others, and TAF-likelihood self, which involves beliefs about events happening to oneself. The factor structure of the TAFS may show a different number of factors depending on the nature of the sample. A two-factor solution (TAF-moral and TAF-likelihood) emerged for OCD samples, accounting for 61-71% of the variance. A two-factor solution also showed good simple structure in a non-OCD sample, accounting for 59-66% of the variance, but a three-factor solution (splitting TAF-likelihood into TAF-likelihood other and TAF-likelihood self) appears to fit the data best, accounting for 61-75% of the variance (Shafran et al., 1996). Of note, the three-factor
solution was supported in a study conducted with college students at Northern Illinois University (Bailey, Wu, Valentiner, & McGrath, 2014).

TAF is theoretically associated with ICT, but TAF may be distinct from ICT as a belief domain involved in OCD (Bailey et al., 2014). In a hierarchical regression analysis, TAF-moral as measured by the TAFS did not add significant variance beyond ICT as measured by the Obsessional Beliefs Questionnaire-44 ([OBQ-44]; OCCWG, 2005; $\beta = .10, ns$). However, the TAFS-Likelihood Other subscale and the TAFS-Likelihood Self subscale did add significant variance beyond ICT ($\beta = .19$ and .16, respectively). This suggests that some aspects of TAF (i.e., TAF-likelihood self and other) are not well-represented by the OBQ-44 ICT scale and may require separate measurement, such as via the TAFS.

Compared to both forms of TAF-likelihood, TAF-moral more consistently has been associated with religion and religiosity (Berman et al., 2010; Berman et al., 2014; Berman et al., 2013; Rassin & Koster, 2003; Siev & Cohen, 2007) and will be a focus of the current study.

Religion and TAF-Moral

Several studies have examined the relationship between TAF-moral and scrupulosity (Nelson et al., 2006) and between TAF-moral and religiosity (Berman et al., 2014; Rassin & Koster, 2003; Siev & Cohen, 2007). One study examined scrupulosity and TAF-moral in 71 individuals who were diagnosed with OCD at an outpatient anxiety clinic, 22 of whom had symptom presentations involving unacceptable obsessional thoughts with religious, violent, or sexual themes (Nelson et al., 2006). These participants completed questionnaires including the TAFS and the Penn Inventory of Scrupulosity (PIOS; Abramowitz et al., 2002), a measure of scrupulosity. There was a significant positive correlation between the TAFS-Moral subscale and
the PIOS, $r = .44$ (Nelson et al., 2006). Berman et al. (2014) asked 85 Christian undergraduate students to respond to several online questionnaires, including TAFS-Moral and the Santa Clara Strength of Religious Faith Questionnaire (SCSRFQ; Plante & Boccaccini, 1997b), a measure of religiosity. They found a positive correlation, $r = .48$ (Berman et al., 2014). Furthermore, it has been shown that, in a sample of Turkish Muslims, TAFS scores mediated the relationship between religiosity and OC symptomatology (Inozu et al., 2014).

It has been shown that the strength of the relationship between TAF-moral and religiosity may depend in part on religious affiliation. One study involved 100 undergraduate psychology students who identified as Catholic, Protestant, atheist, or “other” (Rassin & Koster, 2003). These participants completed the TAFS and a one-item measure of personal involvement in one’s religion on a 0-100 visual analogue scale (VAS). Across all participants, there was a positive correlation between TAFS-Moral and religiosity, $r = .53$. No correlation was found within the atheist ($r = -.26, ns$) or “Other” ($r = .40, ns$) groups, although this could be because of small sample sizes ($ns = 17$ and 8, respectively) or floor effects. The mean religiosity score for the “other” group was similar to the Catholic group ($Ms = 34.9$ and 26.6, respectively), but the atheist group’s mean religiosity score ($M = 6.3, SD = 9.5$) was significantly lower than both the Catholic and Protestant groups. Furthermore, the atheist group mean is extremely low on the scale and has a relatively small standard deviation. Thus, the lack of a relationship between TAF-moral and religiosity may be due to floor effects resulting from consistently low religiosity scores with minimal variance. One may expect an atheist’s religiosity score invariably would be zero; the non-zero endorsement may be due to certain atheists’ interpretations of the religiosity question, “Please indicate your personal involvement in your religion” (Rassin & Koster, 2003,
p. 363), as asking about their involvement with their atheist beliefs. The relationship between TAF and religiosity was strongest in the Protestant group \((r = .71)\) and moderate in the Catholic group \((r = .42)\); note, however, that the sample sizes were 15 and 60, respectively (Rassin & Koster, 2003).

Not only may the strength of the relationship between TAF-moral and religiosity change depending on religious affiliation, but the level at which individuals endorse TAF-moral may depend on religious beliefs, or the lack thereof. In one study, 73 undergraduate students who met inclusion criteria for the study (those who scored above 33 on the SCSRFQ and could be classified as highly religious Protestants or those who identified as atheists/agnostics on a demographic questionnaire) were asked to complete the TAFS (Berman et al., 2010). The researchers compared the highly religious Protestants and atheists/agnostics on their responses to the TAFS and found that the Protestant group was higher on TAFS-Moral, \(t_{(72)} = 4.11\) (Berman et al., 2010). This suggests that people who endorse a religious affiliation are higher in TAF-moral than those who identify as atheist/agnostic.

Differences in TAF-moral also appear between Christian and Jewish individuals. In one study, 85 psychology students were split into three groups based on religious affiliation: Christian (33), Jewish (22), and atheist/agnostic (30; A. Williams, Lau, & Grisham, 2013). All participants completed the TAFS and the SCSRFQ. Results showed that Christians had higher TAFS Total Scores than both atheists/agnostics and Jewish participants, \(F_{(2, 82)} = 14.46\). Additionally, the correlation between SCSRFQ and TAFS scores was significant within the Christian group \((r = .74)\), but it was not significant within the Jewish group. Additionally, there was no correlation between the two variables in the atheist/agnostic group. Another study
included 218 participants gathered by means of a snowball sampling method (Siev & Cohen, 2007). Participants were divided into four groups based on religious affiliation: Orthodox Jewish ($n = 83$), Conservative Jewish ($n = 54$), Reform Jewish ($n = 34$), and Christian ($n = 47$). All participants completed the TAFS and a 6-item self-report measure of religiosity. There also were differences in mean-level TAFS-Moral between the groups, $F_{(3, 210)} = 63.07$, $\eta^2 = .47$. Post hoc comparisons revealed that the Christian group scored significantly higher on TAFS-Moral than the Jewish groups, even after controlling for religiosity, $F_{(3, 203)} = 56.11$, $\eta^2 = .45$. Results also showed that there was a stronger relationship between religiosity and TAFS-Moral among Christians ($r = .44$) than among each of the three Jewish groups (all $rs < .14$, $n$s).

Additionally, differences in TAF-moral have been found between liberal and conservative branches of the same umbrella denomination (Deacon et al., 2013). When comparing 38 pastors of a relatively conservative Lutheran denomination (Lutheran Church—Missouri Synod) and 32 pastors of a more liberal Lutheran denomination (Evangelical Lutheran Church of America) on their responses to the TAFS, it was found that the conservative pastors scored higher on TAFS-Moral than the more liberal pastors, $t_{(68)} = 4.07$. However, it is unknown which characteristics may be responsible for these differences.

Some studies examined religiosity and ICT, a broader cognitive belief involving TAF. In a study comparing highly religious Protestants ($n = 132$), moderately religious Protestants ($n = 37$), and atheists/agnostics ($n = 51$), participants were asked to complete questionnaires including the original Obsessive Beliefs Questionnaire (OBQ; OCCWG, 2001), which includes separate subscales measuring both importance of thoughts and control of thoughts, as a part of a larger study (Abramowitz et al., 2004). The groups were formed using three items on a questionnaire
related to strength of religious affiliation, religious beliefs, and agreement with the teachings of one’s religion. Participants rated these items on a scale from 1 to 5, with 5 corresponding with “very strong.” Those who responded with a 5 on each item were placed in the highly religious group, those who responded with all 3s in the moderately religious group, and those who identified as atheist or agnostic formed a third group. All other participants were excluded from this study. The highly religious group scored higher on OBQ Control of Thoughts compared to the atheist/agnostic group ($F_{[2, 220]} = 10.79$). There were also differences between groups on OBQ Importance of Thoughts ($F_{[2, 220]} = 18.72$), and post hoc analyses revealed that the highly religious group scored higher on OBQ Importance of Thoughts compared to both the atheist/agnostic and moderately religious groups (Abramowitz et al., 2004).

Together, these findings demonstrate that both strength of religious faith and religious affiliation are associated with TAF-moral and related constructs (i.e., ICT). This suggests that religious teachings may contribute to TAF-moral. Cohen and Rozin (2001) conducted preliminary research comparing Protestants and Jews on their moral judgments of thoughts and actions. In a series of studies, participants read vignettes that attempted to distinguish between a fictitious person’s thoughts and actions. Jewish participants consistently were less condemning of immoral thoughts than were Protestants. However, these groups did not differ on their moral judgment of a person’s character based on an immoral action. Protestants also rated thoughts about immoral actions as more controllable than did Jews. There were no group differences in individualistic or collectivistic self-construal (Cohen & Rozin, 2001). These moral judgments seem to match with Jewish and Christian (including Protestant) doctrine. Jewish teaching focuses on moral action and resisting temptation, with no teachings about the moral nature of mental
states or thoughts (Cohen & Rozin, 2001). For Christians, however, the morality of thoughts is explicitly targeted, as in Jesus’ Sermon on the Mount: “You have heard that it was said, ‘You shall not commit adultery.’ But I tell you that anyone who looks at a woman lustfully has already committed adultery with her in his heart” (Matthew 5:27-28, New International Version).

To my knowledge, all published research examining religion and TAF-moral is correlational; this precludes determination of causality and directionality. However, experimental methods have been used to explore TAF using an intrusive thought induction task.

**Thought Induction Task**

The Sentence Task—a thought induction task—originally was created by Rachman, Shafran, Mitchell, Trant, and Teachman (1996) to study covert neutralizing behaviors in OCD that were difficult to examine experimentally. Their intention was to examine whether neutralizing behaviors have the same effects as overt compulsions. They studied 63 university students who displayed evidence of TAF-likelihood or TAF-moral during a telephone screening. Baseline anxiety and guilt were measured using a verbal analogue scale (VeAS) ranging from 0-100. If anxiety ratings were above 30, relaxation instructions were given until ratings were below 30. Next, the experimenter gave the following instructions: “Keeping in mind a friend or relative who is close to you [pause], I would like you to write out the following sentence on this piece of paper inserting the name of the person in the blank” (Rachman et al., 1996, p. 891). Participants were given blank paper with the sentence, “I hope ______ is in a car accident.” After writing down the sentence with the friend’s or relative’s name in the blank, participants were instructed to close their eyes and think about the situation for a few seconds. Levels of anxiety, guilt, perceived likelihood of the event occurring, perceived control over the event occurring,
perceived responsibility if the event were to occur, perceived moral wrongness of writing the sentence, and urge to neutralize were collected using the VeAS. These variables were chosen based on features of covert compulsions, taken from Rachman and Hodgson (1980).

One goal of the Rachman et al. (1996) study was to examine whether anxiety and urge to neutralize would decline naturally if neutralizing behaviors were delayed or prevented. Participants were assigned either to be given the opportunity to neutralize immediately after the first set of VeAS ratings (condition 1) or be faced with a 20-minute delay before being given the opportunity to neutralize (condition 2). For both groups, the opportunity to neutralize was introduced with the following instruction: “You may do whatever you wish to try to reduce or cancel the effects of writing the sentence” (Rachman et al., 1996, p. 892). All participants were asked to give their 0-100 VeAS ratings again after neutralizing.

Rachman et al. (1996) employed a manipulation check that compared VeAS anxiety ratings from before and after the thought induction task. The results showed that the manipulation was successful across both conditions, $t_{(62)} = 24.6, p < .001$. Additionally, mean guilt ratings from before and after the thought induction task increased, $t_{(62)} = 13.3, p < .001$. It was found that, after neutralization in condition 1, participants experienced declines in anxiety ($t_{[28]} = 11.94, p < .001$), guilt ($t_{[28]} = 6.87, p < .008$), perceived probability of the event occurring ($t_{[28]} = 3.09, p < .008$), feelings of responsibility if the event were to occur ($t_{[28]} = 3.71, p < .008$), perceived moral wrongness ($t_{[28]} = 3.52, p < .008$), and urge to neutralize ($t_{[28]} = 6.33, p < .008$). Neutralization was not associated with changes in feelings of control over the threat. Analyses comparing the two conditions at each time point revealed only a difference in urge to neutralize and only at time point 2 ($F_{[1,60]} = 7.42, p = .008$). Thus, urge to neutralize decreased more after
immediate neutralizing than it did after waiting for 20 minutes, even though anxiety and the other variables decreased at the same rate in both groups. Together, these results suggest that the thought induction task is effective at mimicking obsessions in that it causes anxiety, guilt, and other negative feelings that are reduced after performing some form of compulsion (i.e., neutralization) and in that anxiety and other variables decline over a period of time when compulsions are prevented.

The relationships between scores on the TAFS-Likelihood Other subscale and post-task VeAS ratings also were examined in the study (Rachman et al., 1996). There were significant correlations between TAFS-Likelihood Other and anxiety ($r = .26, p < .001$), feelings of control ($r = .32, p < .05$), and feelings of responsibility if the event were to occur ($r = .38, p < .05$). However, TAFS-Likelihood Other was not correlated with guilt, moral wrongness, or urge to neutralize. The TAFS-Moral subscale was not correlated with any VeAS ratings, including moral wrongness ($r = .23, ns$). This may suggest that reactions to the thought induction task are related to TAF-likelihood, and that TAF-moral was not related to this task.

Rachman et al.’s (1996) study was partially replicated by van den Hout, van Pol, and Peters (2001). The same protocol was used for the thought induction task, but the participants were not selected for elevated TAF. The results revealed a post-task increase in anxiety ($F_{[1, 77]} = 210.2, p < .001$). After immediate neutralization, group 1 showed decreases in anxiety ($t_{[39]} = 9.3, p < .001$) and urge to neutralize ($t_{[39]} = 6.7, p < .001$). Thus, the authors concluded that the main findings from Rachman et al. (1996) can be replicated in participants not selected for TAF. However, scores on the TAFS were not related to greater increases in anxiety due to the thought induction task or greater decreases in anxiety and urge to neutralize after the opportunity to
neutralize was given (van den Hout et al., 2001). This suggests that TAF may not be the only factor influencing responses to the task. However, they did not distinguish between subscales of the TAFS, opting to use the total score in their analyses, which may explain the lack of association between TAF and responses to the Sentence Task in their study.

Other researchers have explored how responses to the thought induction task relate to scores on the individual subscales of the TAFS. Berman, Abramowitz, Wheaton, Pardue, and Fabricant (2011) used a modified version of Rachman et al.’s (1996) thought induction task to examine it as an in vivo measure of TAF. To begin the task as modified by Berman et al. (2011), baseline anxiety was measured with a VAS ranging from 0 to 100. Participants then were instructed to write the name of a close, living relative on a notecard. After that, participants were asked to write one of the following sentences, filling in the blank space with that person’s name: “I hope I have sex with _____” or “I hope _____ is in a car accident today.” The sentences were counterbalanced. Participants then were told to think about the event occurring, and then used 0-100 VASs to rate their current level of anxiety, perceived likelihood of the event occurring due to their thinking and writing about it, and perceived moral wrongness of thinking and writing about the thought. Participants then were told that they could do anything they wish to neutralize or reduce/cancel the effects of thinking and writing about the sentence; neutralizing behaviors were recorded. After going through this process for the first sentence, it was repeated for the second sentence.

Berman et al. (2011) found significant correlations between in vivo ratings and TAFS scores. For the car accident sentence, in vivo ratings of anxiety were positively correlated with TAFS-Likelihood Other \( (r = .37) \) and with TAFS-Likelihood Self \( (r = .21) \). In vivo likelihood
ratings were correlated with TAFS-Likelihood Other ($r = .26$). In vivo moral wrongness ratings were not correlated with TAFS-Moral ($r = .15$, $ns$). For the incest sentence, in vivo anxiety was correlated with TAFS-Moral ($r = .23$), TAFS-Likelihood Other ($r = .33$), and TAFS-Likelihood Self ($r = .23$). The relationship between in vivo anxiety and TAFS-Likelihood Self despite neither sentence referring to the self may reflect the known association between TAFS-Likelihood Self and TAFS-Likelihood Other ($r = .47$, $p < .001$; Rassin, Merckelbach, Muris, & Schmidt, 2001). In vivo moral wrongness was positively correlated with TAFS-Moral ($r = .26$).

For both the car accident and the incest sentences, TAFS-Moral predicted neutralizing behavior ($r_{pbo} = .20$ and .21, respectively). In vivo ratings were not associated with measures of anxiety ($rs = -.02$-.12, $ns$) or depression ($rs = -.08$-.12, $ns$). These results confirm Rachman et al.’s (1996) findings that ratings in response to the car accident sentence may not be related to TAF-moral, as measured by the TAFS. However, the results from Berman et al. (2011) show that neutralizing behavior in response to the car accident sentence is related to TAF-moral.

The Sentence Task also has been used to compare religious and non-religious individuals. Berman et al. (2010) asked highly religious Protestants and atheists/agnostics to complete this task, including the in vivo ratings and a dichotomous variable assessing whether or not participants performed neutralizing behaviors. For the car accident thought, there was no difference in ratings of moral wrongness or anxiety, but the Religious group showed higher likelihood ratings, $t_{(71)} = 1.96$. For the incest thought, there was no difference in anxiety or likelihood, but the Religious group gave higher ratings of moral wrongness than the atheist/agnostic group, $t_{(71)} = 3.29$. For both the car accident and incest thoughts, the Religious group was more likely to engage in neutralizing behaviors ($\chi^2_{(1)} = 4.98$ and 5.00, respectively).
The Berman et al. (2010) study introduced the variable of religion to the literature involving the Sentence Task and showed that individuals with different beliefs about religion have different responses to this task. Specific religious teachings have not yet been experimentally manipulated in this literature, but responses to experimental manipulation of other variables have been researched. In one study, the effects of an educational intervention on participant responses to the original version of the thought induction task were examined (Zucker, Craske, Barrios, & Holguin, 2002). Participants were 72 undergraduate students who scored at least one standard deviation above the mean on the TAFS. They were randomly assigned to an experimental or control condition. In the experimental condition, 36 participants heard an educational message about TAF, explaining that intrusive thoughts are normal and do not make one a bad person nor do they increase the likelihood that something will happen. The control group (n = 36) heard a message about stress and the importance of stress management. Participants then completed the Sentence Task, using the car accident sentence. For in vivo anxiety rated using a 0-100 VeAS, analyses revealed that anxiety increased for both groups after performing the task, $F_{(1, 35)} = 29.07, p < .001$. Additionally, the control group showed a greater increase on in vivo anxiety post-task than the experimental group, $F_{(1, 35)} = 4.43, p < .05$. Anxiety, as measured by the State Trait Anxiety Inventory (STAI), State version (Spielberger, Gorsuch, & Lushene, 1970), also increased, $F_{(1, 65)} = 14.29, p < .001$. Interaction analyses using STAI scores were nonsignificant, yet they showed a trend in the same direction as the analysis using the VeAS anxiety scores, $F_{(1, 65)} = 2.77, p = .10$. When examining TAF, the authors found a significant main effect of Time (baseline, post-message; $F_{(1, 63)} = 10.98$) and a significant Time x Group interaction, $F_{(1, 63)} = 9.56$. There was a significant decrease of TAF post-message for the
experimental group ($t_{32} = 3.61$), but not for the control group. Compared to the control group, the experimental group tended to have less post-task in vivo anxiety and less urge to neutralize, but these trends were nonsignificant. There was no difference between the groups on post-task guilt (Zucker et al., 2002).

A similar manipulation was utilized in another study with a focus on TAF-moral (Teachman & Clerkin, 2007). In this study, the thought induction task was not used; instead, the authors examined implicit and explicit morality of intrusive thoughts. After completing baseline measures of obsessive beliefs, depression, and in vivo anxiety, 91 undergraduate students and community volunteers were shown a list of many intrusive thoughts that had been reported by healthy individuals and were asked to reflect on the thoughts that they had experienced. Participants were randomly assigned to a Morally Significant (MS; $n = 30$), Meaningless (M; $n = 30$), or Control condition ($n = 31$). In the MS condition, participants were given information stating that intrusive thoughts are morally significant and indicative of a person’s character. The M group was given a similar statement, but it gave the opposite conclusion: intrusive thoughts are meaningless, fleeting, and unrelated to a person’s true character. The Control group was not given any such information. All participants then were instructed to again think about the intrusive thoughts that were most relevant to them for two minutes, and those in the experimental conditions were told to think of them “in light of whether they meant anything about a person’s virtue” (Teachman & Clerkin, 2007, p. 1008). Finally, participants completed post-task questionnaires which measured explicit responses (i.e., TAF-moral, self-esteem, mood, and beliefs about the personal meaning of one’s thoughts) and implicit association tests (IATs) that measured implicit responses in the form of automatic associations between concepts or stimuli.
In the study, three IAT tasks were used to measure the strength of association between oneself and “immoral” / “moral” (IAT self-immoral), between oneself and “dangerous” / “harmless” (IAT self-dangerous) and between unwanted thoughts and “important” / “meaningless” (IAT unwanted thoughts-important).

The manipulation check was successful; following the instruction, TAFS-Moral was higher in the MS group than the combined M and Control groups, $t(88) = 2.27$ (Teachman & Clerkin, 2007). Planned weighted contrasts did not show differences in instruction condition on any of the other variables, explicit or implicit. Baseline obsessional beliefs were associated with state self-esteem ($r = -.59$), TAFS-Moral ($r = .53$), and evaluations of the significance of intrusive thoughts ($r = .57$), but the instructional condition had no effect. However, implicit responses were predicted by the interaction between instruction condition and obsessive beliefs, including importance of thoughts.

Teachman and Clerkin’s (2007) finding that implicit outcomes only were predicted by the interaction between instructional condition and pre-existing obsessional beliefs highlights the limitations of brief manipulations. Namely, brief manipulations alone may be unlikely to influence implicit feelings and beliefs, and responses to brief manipulations can be influenced by prior beliefs. In a similar manipulation of salience of religious teaching, then, it would be important to use participants who already endorse affiliation with the religion being examined.

**Priming Concepts of God**

In order to warrant the priming of God concepts in OCD research, it must first be established that certain God concepts are associated with OCD. Some studies have examined this and found that higher scrupulosity and TAF-moral are associated with more negative or
“micromanaging” (the belief that God is deeply concerned with all thoughts and actions) views of God among Protestants (Deacon et al., 2013; Siev, Baer, & Minichiello, 2011).

In one study, 72 individuals with scrupulous OCD completed a measure of severity of scrupulous symptoms and concepts of God (Siev et al., 2011). The positive concept of God included beliefs that God is peaceful and compassionate, whereas the negative concept of God represented a view of God as punitive. Severity of scrupulosity was positively correlated with negative concepts of God ($r = .35$) and was uncorrelated with positive concepts of God ($r = -.05$, ns). The authors asserted that the nature of scrupulous fears implies that scrupulous individuals view God as punitive and harsh (Siev et al., 2011).

In their study comparing conservative and liberal Lutherans, Deacon et al. (2013) also compared the two groups on their endorsement of a “micromanaging” view of God. The conservative group, who scored higher on TAFS-Moral, also scored higher than the liberal group on micromanaging views of God, $t_{(68)} = 4.61, p < .001, d = 1.11$. There was no difference between the two groups on their endorsement of positive views of God ($t_{(68)} = .49$, ns), corroborating Siev et al.’s (2011) findings. These findings suggest that priming different concepts of God may influence individuals’ TAF-moral and scrupulous beliefs.

Although manipulating religious beliefs likely cannot be achieved in the context of a one-time laboratory-based experiment, it is possible to utilize a priming task to manipulate the salience of specific concepts of God. Unfortunately, there is a dearth of literature that attempts to do this, let alone within the context of OCD research. However, a few studies have attempted to prime certain aspects of God or religion outside of the context of OCD (Carpenter & Marshall, 2009; DeBono, Shariff, Poole, & Muraven, 2017; Shariff & Norenzayan, 2007).
Most of the literature involving priming God concepts does not seek to prime specific views of God, but simply a general sense of God and spirituality. Many studies, including Fergus and Rowatt (2015), have utilized Shariff and Norenzayan’s (2007) word unscrambling task that requires participants to unscramble 10 short sentences that contain God-related words, such as spirit, sacred, and prophet. Although used by many researchers, this manipulation is too general to effectively study the influence of more specific aspects of religion such as punishing and forgiving concepts of God.

Carpenter and Marshall (2009) created a procedure for priming the idea of a loving God in order to examine its influence on moral behavior. The researchers assigned 142 Christian participants to either a no-prime or religious-prime condition. In the priming condition, participants read a sheet of nine Bible verses that addressed themes such as love for God and God’s love for humanity as well as some neutral filler verses (which were screened for themes of reward or punishment). Subsequently, the authors used a moral hypocrisy paradigm to examine whether participants would behave in congruence with their beliefs. Morally congruent behavior was not influenced by priming condition ($\beta = .29, ns$) or by self-reported intrinsic religiosity ($\beta = .24, ns$), but moral congruence was influenced by the interaction between priming condition and intrinsic religiosity ($\beta = .42$, log-odds ratio = 1.52, $p < .05$). Post hoc analyses showed that morally congruent behavior was higher in the religious-prime condition, but only for those higher in intrinsic religiosity ($\beta = .86$, log-odds ratio = .24, $p < .01$; Carpenter & Marshall, 2009). This finding shows that the loving God prime successfully influenced behavior only in combination with the presence of a specific participant characteristic (i.e., intrinsic religiosity).
DeBono et al. (2017) primed two different concepts of God by having 32 Christian undergraduates read a Bible passage focused on either a forgiving God or a punishing God and then write about either God’s forgiving or punishing nature, respectively, for 10 minutes. For their manipulation check, the researchers used four questions created for their study that asked about how much the participants focused on God being forgiving, compassionate, vengeful, and punishing. The manipulation check indicated that participants in the forgiving condition thought more about God’s forgiveness and compassion \((t_{30} = 7.76, p < .001, d = 2.83)\), and those in the punishing condition thought more about God’s vengeance and punitiveness \((t_{30} = 7.32, p < .001, d = 2.67;\) DeBono et al., 2017).

The results of the checks from the Carpenter and Marshall (2009) and DeBono et al. (2017) studies suggest that comparing two experimental conditions may be better than comparing one experimental condition with a control condition, since the task that used two different primes passed the manipulation check but the task that compared a prime with a no-prime group failed its check. This may be because of Carpenter and Marshall’s (2009) choice of the loving God concept as their only prime, or it may be that using two contrasting primes, rather than one prime and a control, increases the chances of the groups being different.

The Current Study

The purpose of the current study was to clarify the relation between religion and TAF-moral, which is implicated in OCD (Bailey et al., 2014; Shafran et al., 1996). Previous research has shown that individuals with differing religious beliefs experience different levels of TAF-moral and display differences in strength of relationship between religiosity and TAF-moral (Berman et al., 2010; Deacon et al., 2013; Rassin & Koster, 2003; Siev & Cohen, 2007; A.
Williams et al., 2013). Specifically, Protestant Christians (perhaps especially those with more conservative beliefs) evidence higher TAF-moral and a stronger relationship between their religiosity and TAF-moral. The current study sought to examine how one particular religious domain (i.e., concept of God as forgiving or punishing) that is implicated in scrupulous OCD may contribute to TAF-moral. This study is the first to experimentally examine the effects of the salience of religious teaching on reactions to an in vivo TAF-moral task.

The current study used the priming task developed by DeBono et al. (2017) to manipulate the salience of specific God concepts. This priming task was chosen because (a) the primary results of DeBono et al. (2017), for which the priming task was created, revealed a large effect size ($d = .73$), and (b) the comparison between a forgiving versus a punishing God is appropriate for the current study because scrupulous obsessions commonly involve fear of God’s punishment (Abramowitz & Jacoby, 2014; Olatunji et al., 2007).

Protestant Christians (herein “Protestants”) most commonly have been examined in research on religion and OCD, and studies using this religious group most consistently have displayed significant correlations between religiosity and general OC symptomatology (Rassin & Koster, 2003), between religiosity and obsessional symptoms (Abramowitz et al., 2004; Inozu et al., 2012), and between religiosity and TAFS-Moral (Berman et al., 2014; Rassin & Koster, 2003; Siev & Cohen, 2007; A. Williams et al., 2013). When Protestants have been compared to other groups (e.g., Catholics, atheists/agnostics), they consistently have scored higher on TAFS-Moral (Abramowitz et al., 2004; Berman et al., 2010; Siev & Cohen, 2007; A. Williams et al., 2013). Additionally, the priming task in the current study utilized passages from the Christian Bible. Therefore, the current study selected a sample of Protestants.
Because there are mean-level differences in TAF-moral among participants from different religions and also different denominations within a given religion, it remains possible that different religious teachings may promote or weaken TAF-moral beliefs. However, no prior research has employed experimental methods to examine this, so it was unknown whether people who have higher TAF-moral tend to affiliate with certain religions or denominations that are consistent with established views. To examine causality, the DeBono et al. (2017) priming task was utilized in the current study.

Instead of including the TAFS as a measure of TAF-moral, this study employed the thought induction task as modified by Berman et al. (2011). Rather than explicit endorsement of items, the behavioral task may be better at assessing implicit reactions and thus may be more meaningful in light of findings from Teachman and Clerkin (2007), which suggest that experimental manipulations meant to manipulate the salience of a belief are more likely to have an effect on implicit measures of variables than on explicit measures. In this study, the car accident sentence stem was used because it has been effective in eliciting the urge to neutralize and feelings of anxiety and moral wrongness, and these feelings that were caused by the task using this stem have been shown to be reduced after neutralizing, thereby mimicking obsessions (Rachman et al., 1996). TAFS-Moral scores have predicted neutralizing behavior with the use of the car accident sentence (Berman et al., 2011). There also tends to be greater variability in moral wrongness ratings when using the car accident sentence (SD = 31.30) compared to the incest sentence (SD = 22.53; Berman et al., 2011). Furthermore, the variability in moral wrongness ratings for highly religious Protestants is much greater for the car accident sentence (SD = 33.70) compared to the incest sentence (6.08; Berman et al., 2010). In Berman et al.
(2010), there may have been a ceiling effect for moral wrongness ratings of the incest sentence among highly religious protestants, as evidenced by a high mean and a small standard deviation for moral wrongness ($M = 98.47, SD = 6.08$). Since this study utilized Protestants, and since greater variability is desirable for detecting differences, the car accident sentence was chosen. The thought induction task was used to reveal how the priming manipulation may affect in vivo reactions to an obsession-like thought that has been shown to induce TAF-moral.

The current study examined whether priming punishing God concepts (as compared to priming forgiving God concepts) leads to greater in vivo ratings of anxiety and perceived moral wrongness in response to a negative thought induction task.

**Hypotheses**

Three primary hypotheses were proposed for this study.

**Hypothesis 1**

Participants in the punishing God condition will report higher perceived moral wrongness following the thought induction task compared to participants in the forgiving God condition.

**Hypothesis 2**

Participants in the punishing God condition will display greater increases in anxiety from pre- to post-sentence task compared to participants in the forgiving God condition.
Hypothesis 3

Participants in the punishing God condition will be more likely to engage in neutralizing strategies following the thought induction task compared to participants in the forgiving God condition.
CHAPTER 2

METHOD

Participants

A total of 118 participants were recruited from introductory psychology courses at Northern Illinois University. Participants were screened using a one-item question assessing OCD diagnostic history, a one-item demographic question on religious affiliation, and the SCSRFQ in a mass testing procedure. Participants were included if they (a) denied any current or past OCD diagnosis, (b) identified as Protestant Christians (“Protestants”) and (c) scored higher than one standard deviation below the mean on the SCSRFQ. Those who met inclusion criteria were invited via email to participate in the laboratory-based portion of the study in exchange for course credit for a research exposure requirement. Of the 118 individuals who participated in the study, 19 were excluded from analyses for the following reasons: two were due to an invitation error, one due to suspected marijuana use, two due to an administration error, one discontinued the study due to a possible impending OCD diagnosis, five did not complete the SCSRFQ correctly, and eight refused to complete the Sentence Task. This left a final sample size of 99. This sample had a mean age of 19.18 (SD = 1.62), were 61.6% female, 37.4% male, and 1% “other.” Approximately half of the participants (49.5%) were African American/Black, 37.4% were Caucasian/European American, 5.1% Asian American, 5.1% multiracial, and 3% “other.” Most of the sample (93.9%) identified as non-Hispanic/Latino.
Measures

Demographic Questionnaire

In the mass testing procedure, one question was utilized to measure and screen for religious affiliation. Other demographic questions were asked to assess sex, age, and racial/ethnic identification in the sample. See Appendix A for demographic questions.

OCD Diagnostic History

One item was used to screen for diagnostic history of OCD. It was included in both the mass testing procedure and in the lab pre-manipulation. Participants were asked to respond yes or no to the following: “Have you ever been diagnosed with obsessive-compulsive disorder (OCD)?” Individuals who responded in the affirmative were not included in the study. In the mass testing procedure, 2.4% reported a diagnostic history of OCD and 7.4% did not respond to the item; both resulted in exclusion. In the lab pre-manipulation, one individual reported an impending OCD diagnosis and participation was discontinued.

Baseline God Concept

One item was used for the current study to assess participants’ views of God pre-manipulation. Participants responded to the question, “Please indicate the response that most accurately represents your view of God” on a 6-point scale: 1 (very punishing), 2 (punishing), 3 (somewhat punishing), 4 (somewhat forgiving), 5 (forgiving), and 6 (very forgiving).
Religiosity

The Santa Clara Strength of Religious Faith Questionnaire (SCSRFQ; Plante & Boccaccini, 1997b) is a self-report questionnaire of strength of religious faith that can be applied independent of religious affiliation. There are 10 items, with response options on a 4-point scale ranging from 1 (strongly disagree) to 4 (strongly agree). A total score is generated by summing the responses for each item. Factor analyses in several studies have supported a one-factor model (Cummings et al., 2015; Dianni, Proios, & Kouthouris, 2014; Freiheit, Sonstegard, Schmitt, & Vye, 2006; Lewis, Shevlin, McGuckin, & Navratil, 2001; Plante, Vallaey, Sherman, & Wallston, 2002). For the purposes of this study, the SCSRFQ was used to screen out potential participants for very low religiosity.

The SCSRFQ demonstrates good internal consistency, with a Cronbach’s alpha of .95 in undergraduate students (Plante & Boccaccini, 1997b). The three-week retest reliability coefficient was $r = .93$ for healthy adults (Sherman et al., 2001). The SCSRFQ has correlated strongly with spirituality as measured by the revised Spiritual Experience Index (SEI-R; Genia, 1997) and the spiritual support subscale of the SEI-R ($rs = .76$ and .89, respectively; Freiheit et al., 2006). Plante and Boccaccini (1997a) found that it also correlated positively with the Age Universal Religious Orientation Scale (AUROS; Gorsuch & Venable, 1983) intrinsic religiousness subscale ($r = .87-.90$) and with the extrinsic religiousness subscale of the AUROS ($r = .64-.73$). Discriminant validity has been established between the SCSRFQ and the depression subscale of the Symptom Check List-90-Revised (Derogatis, 1977; $r = -.20$; Plante & Boccaccini, 1997b), negative religious coping as measured by the Brief Religious Coping Scale (Pargament et al., 1998; $r = .00, ns$), the positive affect and negative affect subscales of the
Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988; $rs = .18$ and $.00$, $ns$, respectively; Freiheit et al., 2006), the Self-Righteousness Scale (Falbo & Belk, 1985; $rs = -.14$-.03, $ns$), the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960; for university students, $rs = -.02$-.09, $ns$; for individuals in an alcohol and drug recovery program, $r = .21$, $p < .01$), and the Taylor Manifest Anxiety Scale – Short-Form (Bendig, 1954; $r = -.16$; Plante, Yancey, Sherman, Guertin, & Pardini, 1999). See Appendix B for the SCSRFQ.

In Vivo Measures

Anxiety and moral wrongness each were measured using a computerized VAS ranging from 0 (not at all) to 100 (extremely high). The Adaptive Visual Analog Scales (Marsh-Richard, Hatzis, Mathias, Venditti, & Dougherty, 2009) were used to record these ratings. Participants were instructed to use the VAS to answer the following questions, adapted from Rachman et al. (1996): (1) How much anxiety do you feel right now? (2) How morally wrong was it to think about and write out the sentence? The first question had not been changed, but “think about and” was added to the second question. Asking about both thinking and writing has been implemented by other researchers who have used this paradigm (Berman et al., 2010; Berman et al., 2011; Berman et al., 2013). Different forms of these questions have been used repeatedly in studies that involve the thought induction paradigm developed by Rachman et al. (1996).

Convergent validity evidence is mixed, but it has been shown that in vivo anxiety ratings in response to the car accident version of the Sentence Task are correlated with TAFS-Likelihood Other and with TAFS-Likelihood Self ($r = .20$; Berman et al., 2011). Furthermore, in vivo ratings of moral wrongness have correlated positively with neutralizing behavior ($rs = .25$ and .21, respectively; Berman et al., 2011). Discriminant validity is supported for anxiety and
depression, both of which were not significantly correlated with in vivo anxiety (rs = .10 and .12, respectively, ns) and in vivo moral wrongness (rs = .11 and .07, respectively, ns).

**Priming Task**

The priming task was developed by DeBono et al. (2017), and all instructions and parameters reported herein follow their procedure. Participants were handed a Bible and instructed to write about either God’s forgiving nature or God’s punishing nature for 10 minutes. To begin, participants in the FG condition were directed to James 3:17, which reads, “But the wisdom that comes from heaven is first of all pure; then peace-loving, considerate, submissive, full of mercy and good fruit, impartial and sincere.” Then, as in DeBono et al. (2017), the experimenter told participants in the FG condition, “Please read the following passage in the Bible. We want you to really think about what this passage means to you and how it applies to your belief in God. We especially want you to think about how God is eternally forgiving. After you absorb the passage, we would like you to write about the passage for 10 minutes. Write about your interpretation of the Bible passage and how it conveys God as always compassionate and forgiving.”

Participants in the PG condition were directed to Deuteronomy 29:20, which reads, “The Lord will never be willing to forgive him; his wrath and zeal will burn against that man. All the curses written in this book will fall upon him and then the Lord will blot out his name in heaven.” Then, as in DeBono et al. (2017), the experimenter told participants in the PG condition, “Please read the following passage in the Bible. We want you to really think about what this passage means to you and how it applies to your belief in God. We especially want you to think about how God is eternally punishing. After you absorb the passage, we would like you
to write about the passage for 10 minutes. Write about your interpretation of the Bible passage and how it conveys God as always vengeful and punishing.”

Both groups then were given 10 minutes to write about their respective topics and their writing was collected. After completing this task, participants were asked four questions about how much they had focused on God being “forgiving,” “compassionate,” “vengeful,” and “punishing.” These questions were used as manipulation checks as they had been used in DeBono et al. (2017). They are as follows: “I focused on God being forgiving for the Bible task,” “I focused on God being kind during the Bible task,” “I focused on God being vengeful for the Bible task,” and “I focused on God punishing people for the Bible task,” rated on a scale from 1 (not at all true) to 7 (absolutely true; A. DeBono, personal communication, April 12, 2017).

**Thought Induction Task**

Participants completed the thought induction task (Rachman et al., 1996) as adapted by Berman et al. (2011). Participants first were asked to rate their baseline anxiety using the 0-100 VAS.

Next, as in Berman et al. (2011), participants were asked to think of a close, living blood relative and to write that person’s full name on a note card. Once the name was written, the note card was propped up by a computer monitor in front of the participant. Then the experimenter showed participants an unpleasant thought written out in a sentence on the computer monitor in order to trigger TAF beliefs: “I hope ______ is in a car accident.” Participants were instructed to copy the sentence on another note card, filling in the name of the previously identified close relative in the blank. Once the participant was finished writing, participants were asked to close
their eyes and think about the event occurring for a few (60) seconds (N. C. Berman, personal communication, April 6, 2017).

A prior study that used the car accident version of the Sentence Task at NIU reported that 6% of participants refused to write the sentence (Bailey, 2016). For participants who hesitated (e.g., “I’m not sure about this” or “Do I have to write that?”), the experimenter informed the participant that it is not a test of courage and the participant is free to refuse. For participants who decided not to complete the task, the refusal was coded, and the participant was debriefed.

After thinking about the event, participants were asked to rate their anxiety and perceived moral wrongness using the 0-100 VAS. The experimenter then said, “You may now do anything you wish to reduce or cancel the effects of writing or thinking about the sentence.” The function of this prompt was to give participants the opportunity to perform neutralizing behaviors if they felt compelled to do so. Following the specification made by Cougle, Purdon, Fitch, and Hawkins (2013), the experimenter then left the room for 1 minute; upon returning to the room, the experimenter asked whether the participant used “any coping behaviors in the last few moments.” All coping behaviors reported as well as any physical alterations to the note card were recorded by the experimenter.

Procedure

Selection

Participants were recruited from students enrolled in Psychology 102 at NIU. They were screened using a mass survey at the beginning of each academic term via the religious affiliation demographic question, the OCD diagnosis question, and the SCSRFQ.
Criterion 1

Only those who indicated a Protestant Christian affiliation were invited to participate. This group was chosen because (a) the priming task in this study involves quotes from the Christian Bible and (b) Christians have been shown to have higher levels of TAF-moral than other religious groups (Cohen & Rozin, 2001). Protestants and Catholics are the most prevalent Christian groups in the geographical area where this research was conducted (Association of Statisticians of American Religious Bodies, 2010), but the relationship between TAF-moral and religiosity is stronger among Protestants than among Catholics (Rassin & Koster, 2003). This suggested that Protestant Christians may be most sensitive to the manipulations that were used in this study.

Criterion 2

Participants also were screened for extreme low scores on the SCSRFQ. In previous research, mean scores on the SCSRFQ typically were near 29.5 (range from 26-33) in undergraduate populations, with standard deviations around 7 (range from 6-8; Plante, 2010). Those who scored lower than one standard deviation below the local mean score were not invited to participate. This criterion was implemented in order to exclude participants who may not respond as strongly to the priming manipulation due to low religiosity (Carpenter & Marshall, 2009; Gervais & Norenzayan, 2012). However, it was not so restrictive as to eliminate all variability in religiosity within the sample or to unduly reduce the likelihood of securing sufficient participation to reach the targeted sample size.
Criterion 3

In order to minimize potential harm, all participants were asked whether they have a current or prior diagnosis of OCD. Participants responded to this question during the mass testing procedure and at the beginning of the lab session. Those who reported an OCD diagnosis were not included in the study. The 195 participants who met all three criteria were invited to participate via email.

Tasks

Upon arrival, participants were provided with a written informed consent document and, after an opportunity to review it and ask questions, were asked to provide written informed consent. Using block randomization to increase the chances of having equal group sizes, participants were randomly assigned to either the FG condition \((n = 51)\) or the PG condition \((n = 48)\). Participant assignment used random numbers that were determined prior to participant arrival.

After giving informed consent, participants completed the baseline God Concept (BGC) question and the OCD diagnostic history question. Participants completed their assigned condition of the priming task. Then participants completed the thought induction task, gave their VAS ratings, and had the opportunity to perform neutralizing behaviors.

Debriefing

Upon completion of the study, all participants were provided with a written debriefing form as to the purpose of the study. Participants were given the opportunity to ask questions of
the experimenter. Then they were given proof of their participation and a list of local-area
counseling resources they could contact if anything they had done in the study bothered them or
if they had any concerns about intrusive thoughts. They were thanked for their participation and
dismissed.
CHAPTER 3
RESULTS

Data Cleaning

To examine patterns in missing data, Little’s MCAR test was conducted, and it revealed that the data were missing completely at random, $\chi^2(45) = 46.31$, $p = .418$. SCSRFQ data were missing for three participants: two failed to respond to 10% of the questionnaire and one failed to respond to 20%. Since these percentages were above 5%, they were excluded from analyses using the SCSRFQ (Schafer, 1999). No other data were missing.

Outliers were examined by plotting the data from each measure using box plots. There were three extreme outliers in the BGC data; these outliers were winsorized to 3.0 SD from the local mean. No other extreme outliers were identified.

It was proposed that normality of the data would be tested using both the Kolmogorov-Smirnov and Shapiro-Wilk tests of normality, and that if either test was significant, transformations would be used. Further, it was proposed that skew and kurtosis of the data also would be examined, and that data would be considered significantly skewed and/or kurtotic if standardized scores for skew and kurtosis exceeded $\pm 2.58$ (Field, 2005). However, since the project was proposed, I have learned more about alternative approaches to such situations. First, since it has been shown that kurtosis is more important than skew when using an ANOVA (Glass, Peckham, & Sanders, 1972; Khan & Rayner, 2003), it is more relevant to examine skew and kurtosis as separate entities; this suggests the use of skew and kurtosis standardized scores, as opposed to overall tests of normality that do not distinguish between the two. Another way to consider this issue is that focusing on these scores allows for an analysis of the degree of
normality, rather than a categorical test that emphasizes statistical significance. As such, skew and kurtosis scores were calculated by assigned condition (see Table 1 for all standardized scores). Since skew and kurtosis scores for BGC were high—even after winsorizing the three outlier data points—the BGC data were square-root transformed. The skew and kurtosis standardized scores after the transformation were as follows: FG skew = -3.59, FG kurtosis = 1.69, PG skew = -3.51, and PG kurtosis = 1.29. After the transformation, the distribution of BGC scores was closer to a normal distribution, and the standardized scores were closer to the scores for the other variables (see Table 1). However, the standardized scores for skew exceeded ±2.58 for the transformed BGC variable, the moral wrongness VAS variable, and the FG pre-task anxiety VAS variable. Whereas it was proposed that transformations would be performed for any variable found to exceed ±2.58, this may not be necessary for these variables and doing so may in fact be contraindicated. To begin, the BGC data already were transformed and still displayed skewness. Due to the type of test being used in the main analysis, if pre-task anxiety VAS is transformed, then so must post-task anxiety VAS. This can lead to problems, since transforming approximately normal variables actually can result in non-normality (Field, 2013). In general, the use of data transformation in the face of non-normal data is discouraged (Field, 2013; Glass et al., 1972). In addition, there is not agreement on which standardized score should be used as the criterion for significance in a given range of sample sizes. To illustrate, Kim (2013) posited that ±3.29 should be the criterion for samples with ns between 50 and 300; Ghasemi and Zahediasl (2012) proposed a criterion of ±1.96 for small samples and ±2.58 for sample sizes of 200 or more; and Mayers (2013) proposed a criterion of ±2.58 for ns between 51 and 100. Still, using the most liberal criterion of ±3.29, both BGC and FG pre-task anxiety VAS would be considered
significant. The reason why the remaining skewed variables need not be transformed is related to
the types of tests being used in the main analyses. Evidence suggests that the \( t \) test and ANOVA
are robust against modest (and at times extreme) violations of normality (Glass et al., 1972;
Havlicek & Peterson, 1974; Khan & Rayner, 2003; Sullivan & D’Agostino, 1992), as long as the
equal variances assumption is met, the distributions are skewed in the same direction for \( t \) tests
(Havlicek & Peterson, 1974), and the distribution is not significantly kurtotic for ANOVA (Khan
& Rayner, 2003). Since all standardized scores for skew did not exceed or were relatively close
to \( \pm 2.58 \) and all standardized scores for kurtosis did not exceed \( \pm 2.58 \), it was decided that no
further transformations would be used.

Table 1

*Skew and Kurtosis Standardized Scores*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Punishing Condition(^a)</th>
<th>Forgiving Condition(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Skew</td>
<td>Kurtosis</td>
</tr>
<tr>
<td>Baseline God Concept</td>
<td>-4.86</td>
<td>4.20</td>
</tr>
<tr>
<td>Pre-task anxiety VAS</td>
<td>2.46</td>
<td>-0.68</td>
</tr>
<tr>
<td>Post-task anxiety VAS</td>
<td>-0.67</td>
<td>-1.19</td>
</tr>
<tr>
<td>Moral wrongness VAS</td>
<td>-3.50</td>
<td>1.11</td>
</tr>
<tr>
<td>SCSRFO</td>
<td>-1.39</td>
<td>.45</td>
</tr>
</tbody>
</table>

*Note.* SCSRFO = Santa Clara Strength of Religious Faith Questionnaire.
\(^a\)\( n = 48 \). \(^b\)\( n = 51 \).
Levene’s test was used to examine the data for homogeneity of variance among the two groups on all variables. The test was nonsignificant for BGC ($p = .264$), pre-task anxiety VAS ($p = .543$), post-task anxiety VAS ($p = .584$), and SCSRFQ ($p = .275$); it was significant for moral wrongness VAS ($p = .026$). Thus, the more robust Welch’s test was used in analyses involving moral wrongness VAS.

Preliminary Analyses

Descriptive statistics were calculated for BGC, pre-task anxiety VAS, post-task anxiety VAS, moral wrongness VAS, SCSRFQ, and the manipulation check questions (see Table 2). Two $t$ tests were used to compare SCSRFQ scores and BGC scores between the two groups. There were no significant group differences on SCSRFQ scores ($t_{[94]} = 1.59, p = .115$) nor on BGC scores ($t_{[97]} = .66, p = .512$). Notably, both groups displayed very high BGC mean scores ($M$s = 5.77-5.81 on a 6-point scale), reflecting concepts of God that lean heavily toward the forgiving end of the scale in both groups. Since no group differences were found on these variables, neither SCSRFQ nor BGC was used as a covariate in the primary analyses. That noted, this issue will be revisited for the potential role it may have played in the main study findings.

Manipulation Check

The four questions used for the manipulation check were analyzed. The questions about whether participants focused on God being “forgiving” and “compassionate” served as the manipulation check for the FG condition; the questions about God being “vengeful” and “punishing” were the manipulation check for the PG condition. Responses for the two questions in each set were summed, and then a $t$ test was used to examine whether the groups differed on
these variables. As expected, it was found that those in the PG condition thought significantly more about a punishing God compared to those in the FG condition, $t_{(78)} = 6.08, p < .001$. Those in the FG condition did not report thinking significantly more about a forgiving God compared to those in the PG condition, $t_{(83)} = 1.30, p = .197$. This indicates that the manipulation only differentiated the groups regarding the PG prime, not regarding the FG prime. This may imply that the Bible task manipulation did not perform as expected, since the FG group did not think about a forgiving God more than did the PG group. Any findings from the group comparisons in the primary analyses should be framed as differences due to priming a punishing God (vs. not priming a punishing God) and not due to priming a forgiving God.

Table 2

*Descriptive Statistics and Mean Comparisons*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Punishing Condition$^a$</th>
<th>Forgiving Condition$^b$</th>
<th>$t / \chi^2$</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$M / %$</td>
<td>$SD$</td>
<td>$M / %$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Baseline God Concept</td>
<td>5.77</td>
<td>.33</td>
<td>5.81</td>
<td>.27</td>
</tr>
<tr>
<td>Pre-task anxiety VAS</td>
<td>23.61</td>
<td>22.06</td>
<td>18.31</td>
<td>20.77</td>
</tr>
<tr>
<td>Post-task anxiety VAS</td>
<td>56.64</td>
<td>27.19</td>
<td>45.22</td>
<td>28.09</td>
</tr>
<tr>
<td>Moral wrongness VAS</td>
<td>74.08</td>
<td>27.01</td>
<td>66.61</td>
<td>34.73</td>
</tr>
<tr>
<td>SCSRFQ</td>
<td>27.58</td>
<td>6.35</td>
<td>29.75</td>
<td>6.98</td>
</tr>
<tr>
<td>FG Manipulation Check</td>
<td>10.94</td>
<td>3.58</td>
<td>11.75</td>
<td>2.47</td>
</tr>
<tr>
<td>PG Manipulation Check</td>
<td>7.13</td>
<td>3.71</td>
<td>3.31</td>
<td>2.32</td>
</tr>
<tr>
<td>Neutralizing Behavior</td>
<td>79.17%</td>
<td>--</td>
<td>78.43%</td>
<td>--</td>
</tr>
</tbody>
</table>

*Note. SCSRFQ = Santa Clara Strength of Religious Faith Questionnaire. Neutralizing behavior is represented by the percentage of participants who responded *yes.*

$^a_n = 48. \ ^b_n = 51.

*p < .05. **p < .001
Primary Analyses

Hypothesis 1

Hypothesis 1 was that participants in the PG condition would report higher perceived moral wrongness following the thought induction task compared to participants in the FG condition. An independent-samples $t$ test was used to compare the PG and FG groups on moral wrongness VAS scores. Whereas the PG group scored higher than the FG group, as was predicted, the magnitude of the difference was not statistically significant, $t_{(94)} = 1.20, p = .234$.

Hypothesis 2

Hypothesis 2 was that participants in the PG condition would display greater increases in anxiety from pre- to post-Sentence Task compared to participants in the FG condition. A 2 x 2 mixed ANOVA was used to examine the main effects of Time (pre- vs. post-Sentence Task) and Condition (PG vs. FG) and the Time x Condition interaction on anxiety VAS scores. A significant main effect was found for Time ($F_{(1, 97)} = 144.02, p < .001, \eta^2_p = .598$) and the effect was in the expected direction (i.e., post-task anxiety VAS scores were higher than pre-task anxiety VAS scores). The effect for Condition also was in the expected direction, with the PG group showing generally higher anxiety VAS scores than the FG group, but this effect was not statistically significant, $F_{(1, 97)} = 3.78, p = .055, \eta^2_p = .037$. The interaction effect also was not significant, $F_{(1, 97)} = 1.50, p = .223, \eta^2_p = .015$. 
Hypothesis 3

Hypothesis 3 was that participants in the PG condition would be more likely than participants in the FG condition to engage in neutralizing strategies following the thought induction task. The PG group did show a higher percentage of participants reporting a neutralizing behavior than did those in the FG condition, but the difference was minimal and the chi-square test was nonsignificant, \( \chi^2(1) = .008, p = .929 \) (see Table 2).
CHAPTER 4

DISCUSSION

This was the first study to examine causal links between religion and TAF-moral. Previous research has established that different religious backgrounds are correlated with different levels of TAF-moral (Berman et al., 2010; Deacon et al., 2013; Siev & Cohen, 2007; A. Williams et al., 2013) and with differing strengths of association between religiosity and moral TAF-moral (Rassin & Koster, 2003; Siev & Cohen, 2007; A. Williams et al., 2013). However, no previous experimental studies have been conducted on the subject, precluding causal inferences. The purpose of this study was to begin exploring this promising area of research by examining whether two opposing God concepts—via religious teachings within the Protestant Christian faith—may influence TAF-moral.

Major Findings

Results for Hypothesis 1 revealed that participants’ ratings of how morally wrong it was to write down and think about the sentence in the Sentence Task did not differ between the two groups. The means were in the expected direction (the PG group mean was higher than the FG group mean), but the difference was nonsignificant. This may suggest that thinking about God as punishing has insufficient influence on whether an individual believes that having an intrusive thought (e.g., in this case, writing down and thinking about the car accident sentence) is morally wrong (i.e., causes increases in TAF-moral).
The results gave partial support for Hypothesis 2. It was clear that the Sentence Task produced increased anxiety from pre- to post-task, thus providing evidence that the Sentence Task performed as expected in mimicking a distressing intrusive thought in participants. The comparison between the two groups on anxiety levels (i.e., the main effect of Condition) was in the expected direction but was not statistically significant. The interaction between Time and Condition also was not significant, suggesting that participants’ anxiety levels from pre- to post-task did not depend on condition. These findings may indicate that focusing on a punishing concept of God (a) does not increase overall anxiety and (b) does not increase the effect that distressing intrusive thoughts have on anxiety. However, the main effect of Condition was close to reaching the a priori established alpha value of $p < .05$ ($p = .055$). Additionally, although not included among the formally proposed analyses, an independent-samples t test revealed that post-task anxiety VAS scores were significantly higher in the PG group compared to the FG group, $t(97) = 2.05, p < .05$. The pre-task anxiety VAS mean also was higher in the PG group, but this difference was not significant. Thus, there appears to be some difference in groups regarding anxiety, and it may suggest that the Bible task manipulation increased overall anxiety in the PG group compared to the FG group, especially after completion of the Sentence Task. Even though the groups showed marginally significant differences on overall anxiety, the interaction term in this analysis was unambiguously nonsignificant.

Results for Hypothesis 3 showed that the frequency of participants who performed a neutralizing behavior was not significantly higher in the PG condition than in the FG condition. This may suggest that neutralizing behavior in response to a distressing intrusive thought is not differentially influenced by whether an individual thought about a punishing God concept.
Notably, endorsement of neutralizing behavior was high in both groups (78.4% and 79.2%), demonstrating that the discomfort produced by the Sentence Task was strong enough to fairly consistently lead to a behavioral response.

Limitations

The limitations to this study will be discussed using Cronbach and Meehl (1955) as a framework. That seminal paper on construct validity defined a theoretical network as consisting of proposed relationships between a test and a construct, between one construct and another, and between constructs and observable phenomena. They posited that when a prediction is not confirmed by a research study, then any of these proposed relationships may be faulty. Thus, null findings should be considered in three contexts: (1) the test does not measure the construct variable, (2) the theoretical network which generated the hypothesis is incorrect, and/or (3) the experimental design failed to test the hypothesis properly.

Relationship to Construct Variable

Some of the findings may owe to a failure to properly measure the relevant study constructs. In part, the null findings may have been due to the way the manipulation check questions were structured. The Bible task and the manipulation check questions were identical to those used by DeBono et al. (2017), but in the current study, the check questions did not indicate that the Bible task functioned properly. This may be due to a failure of the Bible task to prime the two God concepts properly, but it also may be due to the structure of the check questions. The questions measured participants’ focus on both a forgiving and a punishing God in a parallel fashion, as opposed to explicitly pitting the two against each other. Instead of measuring the
degree to which participants in both groups focused on both a forgiving and a punishing God separately, one could ask participants to choose whether they focused more on a forgiving or a punishing God. This could allow for removal of all participants who did not focus on the God concept that corresponded with their respective priming condition. Although the essays written during the Bible task manipulation were not formally examined or coded, the need for a mechanism to exclude participants who did not follow the instructions was reflected in multiple participants’ essays; many participants in the PG condition included commentary on how they disagreed with the passage or how the passage did not apply to them. This likely undermined the effectiveness of the priming manipulation but was not caught by the check questions, since the questions only functioned to compare the overall groups, not to identify individual participants who may not have been influenced by the manipulation.

The null finding regarding neutralizing behavior may be a function of the difficulty of coding this variable. It is possible that the method of determining whether participants engaged in neutralizing behaviors in this study was flawed and did not properly distinguish true neutralizing behavior. For example, when given the opportunity to neutralize, participants were left alone and their actual behavior was not observed by the experimenters. Instead, the data were gathered by asking participants what they had done, which allows for problems with participants (intentionally or unintentionally) not reporting their behavior accurately or completely. Additionally, any alterations to the notecards used in the Sentence Task were recorded as neutralizing behaviors, even if the participant neglected to verbally report the behavior. This strategy assumed that these alterations were forms of neutralizing behavior and may have created a lower threshold for labeling a behavior as neutralizing. The high frequency that participants
neutralized in both groups (78.4% to 79.2%) provides some evidence for this interpretation. Although not a part of the proposal for this study, an additional VAS was collected post-Sentence Task which measured participants’ urge to neutralize using the same scale as the other VAS items. When neutralization was measured in this way, it was found that urge to neutralize was higher in the PG group ($M = 78.52$, $SD = 25.77$) than in the FG group ($M = 62.28$, $SD = 37.69$), $t_{(97)} = 2.49$, $p < .05$. This information may add to the concern that neutralizing behavior was not properly represented by the measurement method in this study (although it was similar to the method used in other studies involving the Sentence Task). It instead may imply that self-report of urge to neutralize may be a more useful measure for discriminating between groups because it uses a continuous scale rather than a dichotomous scale.

**Theoretical Network**

Parts of the theoretical network which generated study hypotheses have more prior evidence than others. For example, it has been shown using the Sentence Task across a number of studies that harm-related intrusive thoughts are distressing and create anxiety, guilt, urges to neutralize, and inflated interpretations of moral wrongness and likelihood of their occurrence (e.g., Rachman et al., 1996; van den Hout, 2001). There also is evidence that people from different religious traditions, including different branches of the same denomination, display different levels of TAF-moral (e.g., Deacon et al., 2013). However, there is little to no prior research that clarifies what constructs might be involved in the teachings of these different religious traditions that might be increasing members’ TAF-moral beliefs. Thus, it is possible that the punishing versus forgiving God concept dichotomy may not be the construct that
influences TAF-moral. Other types of God concepts or other types of religious teachings could be the explanation for the differing levels of TAF-moral in different religious groups.

**Experimental Design**

In reflecting on the design of the current study, there are multiple possibilities for why the experimental design may have failed to test the hypotheses properly. The null finding for the main effect of condition on anxiety may have been indicative of inadequate power, given that the finding was in the expected direction and the $p$ value was very close to .05 ($p = .055$). The power analysis that was completed to determine the target sample size for this study used a higher effect size ($d = .71$) than the effect size found for the main effect of Condition ($\eta^2_p = .037$, which is equivalent to a $d$ of .39) and thus was underpowered for the detection of the smaller than anticipated effect. However, it is unlikely that power was a major factor in the other null findings (i.e., group differences in moral wrongness, the interaction between Time and Condition, and group differences in frequency of neutralizing). This is because, when the effect sizes for these null findings were converted to Cohen’s $d$ for direct comparison, it was apparent that their effects were very small ($ds = .02-.24$). Thus, even if a larger sample size were to provide statistically significant findings of this same magnitude, the meaning of that finding essentially would be unchanged.

The null findings regarding all hypotheses may have been due to vulnerabilities in the design caused by the Bible task manipulation. The priming task may have been unsuccessful in manipulating which concept of God was forefront in participants’ minds while completing the Sentence Task. This study was the first to employ this Bible task in the OCD literature, and only the second beyond DeBono et al. (2017). Thus, although its effectiveness was demonstrated in
the context of the DeBono et al. (2017) study, the validity of the manipulation has not been established outside of its use in that study, and it may not have achieved the desired effect. This assertion is partially supported by the manipulation check findings, which showed that individuals in the PG condition thought more about a punishing God than those in the FG condition, but participants in both conditions were highly focused on God being forgiving. Thus, although the PG manipulation had some effect on whether the participants thought about a punishing God, the FG manipulation was unsuccessful in increasing participants’ focus on God as forgiving. Although not a planned analysis, each manipulation check question was examined individually to see whether one check question may have found differences that its counterpart did not. Independent-samples t tests revealed that both FG check questions were nonsignificant \( (t_{97} = 1.40, p = .166 \text{ for “forgiving,” and } t_{87} = 1.05, p = .299 \text{ for “kind”}) \) and both PG check questions were significant \( (t_{75} = 5.38, p < .001 \text{ for “vengeful,” and } t_{90} = 5.70, p < .001 \text{ for “punishing”}) \), showing a pattern consistent with the analyses using their aggregates. This comparable level of focus on a forgiving God across the two conditions was reflected by some participants in the PG condition who, as noted previously, wrote about their disagreement with the punishing God concept represented in the Bible passage they read. It appears that a one-time reading of a single Bible passage did not have the intended effect, likely owing to too small of a “dose” to display potential effects. The dose may have been particularly relevant given the sample in the current study. Examination of the BGC variable revealed that participants’ baseline God concepts were markedly on the “forgiving” end of the scale. Thus, this sample in particular may have needed a much larger dose in order to “move the needle” for priming this construct.
Perhaps instead of screening out those with extremely low religiosity, as was done in the current study, it may have been more useful to screen out those with extreme baseline God concepts (i.e., those who responded with a 1, *very punishing*, or a 6, *very forgiving*, on the BGC item). The goal of excluding those with extremely low religiosity was to increase the chance that participants would be affected by the priming manipulation. However, given this study’s findings, excluding those with extreme baseline God concepts may have been more effective in increasing those chances.

**Future Directions**

This area of research is ripe for exploration using experimental methods so that we may begin to understand whether and what types of religious teaching or thinking increase individuals’ TAF-moral beliefs. It will be useful for researchers in this area to attend to the following suggestions for future directions.

It may be appropriate to change the manipulation check for the Bible task to an item that will force participants to choose whether they thought more about one God concept or the other. This may allow for a better test of whether the manipulation was successful and will allow for the exclusion of participants for whom it was not successful, as discussed.

More research is needed to find the best way of coding neutralizing behavior within the Sentence Task paradigm. It generally seems to have been done the same way (i.e., by asking participants what they did after the time was up) in prior research, yet there are many potential problems with this method and it may not properly represent individuals’ experiences, as discussed. Other modes of collecting these data may involve using video recording to directly observe behavior; however, it is common in the study of OCD for participants to perform covert
neutralizing behaviors (i.e., those that are not visible to the experimenter). It also may be beneficial to provide participants with a better explanation of what the researchers are asking (i.e., rather than simply asking if they used any coping behaviors) so that participants may be better reporters of their experiences and responses.

Future research should involve an additional pre-screen to exclude individuals with BGC scores on the extreme ends of the spectrum. This may function as a replacement for the religiosity screener, as noted. Given that many of the participants in the current sample would have been excluded using a BGC screener, it may be important to explore samples with different characteristics. For example, samples from other parts of the country or samples of Catholics instead of Protestants may show greater variability in their baseline God concepts. Use of Catholic participants may be useful for another reason: there is agreement on which versions of the Bible are accepted. Between different denominations and individuals within the Protestant tradition, there are different ideas about which versions of the Bible should be accepted. This came up as an issue for at least one participant in the current study, who questioned the Bible verse because the participant did not believe it came from the “correct” Bible. This could be avoided by using a more homogeneous sample.

Researchers should consider developing a priming task with more dose. This could involve instructions that put emphasis on writing about the God concept respective to the participant’s condition even if the participant disagrees. It also could involve more Bible verses or perhaps a recorded sermon that is representative of a certain God concept. These both may increase the dose of the prime and the sermon may increase the chance that participants are
engaged in the prime the entire time (as opposed to the potential risk that the participant writes for only 3 of the 10 minutes they are asked to be writing).

In this study, only one out of many different types of God concepts (i.e., forgiving and punishing) was examined. The God concepts that were primed in this study were chosen in order to mimic the commonly held belief in individuals with scrupulous OCD that God is punishing (Abramowitz et al., 2002) and to provide a comparison (i.e., the FG condition). However, other God concepts and other types of religious teaching (e.g., parts of the New Testament of the Bible that may be interpreted to encourage TAF-moral) still should be examined as causal factors in the relationship between religious tradition and TAF-moral. Future research also should examine other methods similar to the Bible task developed by DeBono et al. (2017) for priming different concepts of God. This could involve finding passages focused on other religious teachings that may influence TAF-moral, such as passages that seem to encourage TAF-moral directly (e.g., part of Jesus’ Sermon on the Mount: “You have heard that it was said, ‘You shall not commit adultery.’ But I tell you that anyone who looks at a woman lustfully has already committed adultery with her in his heart” (Matthew 5:27-28, New International Version).

Finally, future research also may use quasi-experimental methods to examine whether individuals with a punishing baseline concept of God show different reactions to the Sentence Task compared to those with a more forgiving baseline concept of God. In this sample, very few individuals reported a view of God that was more punishing than forgiving in response to the baseline God concept item (one participant responded very punishing, and two responded punishing). Thus, this may need to be examined in a different population that may be more likely to have a punishing concept of God.
Conclusion

In conclusion, this study offered a first examination of the potential causal links between religious teaching/tradition and TAF-moral by examining punishing/forgiving God concepts as a causal factor in increased TAF-moral. Although the current study had notable limitations, it provides preliminary evidence that having a punishing view of God may affect a person’s anxiety in response to a distressing intrusive thought. This promising area of study should continue to be explored so that we may better understand the factors that influence TAF-moral and so that we can take preventative action for the development of unhealthy levels of TAF-moral.
REFERENCES


APPENDIX A

DEMOGRAPHIC QUESTIONNAIRE
Mass testing.

Do you identify as Protestant Christian, such as Lutheran, Methodist, Baptist, or Pentecostal?

1 = Yes  
2 = No

Have you ever been diagnosed with OCD?

1 = Yes  
2 = No

After screening.

1. What is your age?  ________

2. What is your sex?

1 = Male  
2 = Female  
3 = Other

3. What is your racial identification?

1 = African American/Black  
2 = Asian/Asian American  
3 = Caucasian/European American  
4 = Native American  
5 = Multiracial  
6 = Other

4. What is your ethnicity?

1 = Hispanic/Latino  
2 = Not Hispanic/Latino
APPENDIX B

SANTA CLARA STRENGTH OF RELIGIOUS FAITH QUESTIONNAIRE
SCSRFQ

Please answer the following questions about religious faith using the scale below. Indicate the level of agreement (or disagreement) for each statement.

1 = strongly disagree 2 = disagree 3 = agree 4 = strongly agree

1. My religious faith is extremely important to me.
2. I pray daily.
3. I look to my faith as a source of inspiration.
4. I look to my faith as providing meaning and purpose in my life.
5. I consider myself active in my faith or church.
6. My faith is an important part of who I am as a person.
7. My relationship with God is extremely important to me.
8. I enjoy being around others who share my faith.
9. I look to my faith as a source of comfort.
10. My faith impacts many of my decisions.