RELIGIOUS AFFILIATION AND CONCEPTIONS OF THE MORAL DOMAIN

Sydney Levine
Massachusetts Institute of Technology; Harvard University

Joshua Rottman
Franklin & Marshall College

Taylor Davis
Purdue University

Elizabeth O'Neill
Eindhoven University of Technology

Stephen Stich
Rutgers University

Edouard Machery
University of Pittsburgh

What is the relationship between religious affiliation and conceptions of the moral domain? Putting aside the question of whether people from different religions agree about how to answer moral questions, here we investigate a more fundamental question: How much disagreement is there across religions about which issues count as moral in the first place? That is, do people from different religions conceptualize the scope of morality differently? Using a new methodology to map out how individuals conceive of the moral domain, we find dramatic differences among adherents of different religions. Mormon and Muslim participants moralized their religious norms, while Jewish participants did not. Hindu participants in our sample did not seem to make a moral/non-moral distinction of the same kind. These results suggest a profound relationship between religious affiliation and conceptions of the scope of the moral domain.

Keywords: morality, religion, moral domain

What impact does religion have on moral psychology? There is already a fruitful exchange between the field of moral psychology and the study of religion, with each field informing the other. Indeed, past research has explored many aspects of the relationship between religion and morality (Graham & Haidt, 2010; Miller, 2001; Norenzayan, 2014; Shweder, Mahapatra, & Miller, 1987), including how religion influences beliefs about what actions and behaviors are right and wrong (Srinivasan, Kaplan, & Dahl, 2018), how religion influences whether mental states can be morally condemned (Cohen & Rozin, 2001), and how religion influences whether some acts are unforgivable (Cohen, Malka, Rozin, & Cherfas, 2006). Other research has looked at how religious worldviews inform moral reasoning (Jensen, 1997) and moral-life course narratives (McKenzie & Jensen, 2017), whether religious individuals are more or less prosocial than non-religious individuals (McKay & Whitehouse, 2015), and the relationship between religiosity and value prioritization (Schwartz & Huismans, 1995).

We continue in this tradition of gaining insights into moral psychology by looking at the relationship between religion and the moral mind. Our interest is in the question of how religious affiliation shapes conceptions of the scope of morality: What counts as right or wrong in a specifically *moral* way, as opposed to what counts as right or wrong in a legal way, or in an epistemic way, or in some other way (e.g., with respect to good manners)? For example, some religions require their adherents to pray multiple times a day while some religions do not, so it is clear that adherents of different religions differ on whether they think it is *wrong* not to pray multiple times a day. However, our interest is in whether these adherents disagree about whether such norms are (or are not) wrong in a specifically *moral* way.

Why study the relationship between religion and the scope of the moral domain? Answering this question will help us gain insights into the various ways that the moral domain may be structured in the mind. Looking at religious affiliation in relation to conceptions of morality reveals theoretically important differences in the way that the moral domain can be structured. Thus far, these differences have gone unrecognized; indeed, the most prominent theoretical approaches in moral psychology currently do not address such differences. Two leading theories, Social Domain Theory (SDT) and Moral Foundations Theory (MFT), both treat religious individuals as a homogeneous group when it comes to this question, claiming that religious affiliation in general either does or does not have an effect on which norms count as moral, without considering potential differences among different religions. These theories make opposite claims about the impact of religious affiliation on the moral domain. SDT proposes that for all people, regardless of religious affiliation, moral norms are those related to justice, welfare, and rights (Killen & Smetana, 2005; Nucci, 1981; Nucci & Turiel, 1993; Turiel, 1983; for discussion, see Fessler et al., 2015; Kelly, Stich, Haley, Eng, & Fessler, 2007; Sousa, Holbrook, & Piazza, 2009; Stich, Fessler, & Kelly, 2009). By contrast, MFT proposes that conceptions of morality are contingent on cultural factors, such as religion (Haidt & Graham, 2007; Haidt & Joseph, 2004), and that for religious people morality involves a wider range of norms, extending well beyond concerns about harm (i.e., welfare) and fairness (i.e., justice and rights) (Graham & Haidt, 2010). Note that at its root, this debate does not concern a disagreement about how religious affiliation affects which norms people care about. Rather, at issue is a disagreement about how religious affiliation affects which types of norms *count as moral*, as opposed to nonmoral. While the claims of SDT entail that religious affiliation

has no effect on the scope of morality at all, the claims of MFT hold that religion functions to broaden the scope of morality.

Part of the dispute is conceptual—the disagreement could be resolved if there existed some widely accepted, generally agreed-upon definition of "morality" in the field. Unfortunately, however, there is no such accepted definition of morality in moral psychology. Indeed, Stich (2017, 2018) has recently conducted a thorough review of the history of attempts to provide a principled, theoretical definition of morality in psychology and philosophy. It shows not only that there currently is no general agreement about how the moral domain should be delineated, but also that there never has been, despite many earnest attempts over many decades.

Accordingly, it is important to note that some researchers who have approached the question of how religion affects the scope of morality have done so in ways that fail to grapple with the central conceptual debate about how moral norms should be identified. Some have proceeded by simply interpreting judgments about wrongness in general as indicative of moral judgments, which clearly does not address the question of how moral norms are distinguished from nonmoral norms (for discussion, see Berniūnas, 2014; Machery. 2012; Machery & Stich, 2013). Others have proceeded by looking at whether subjects treat religious norms as having features that SDT associates with prototypical moral norms (e.g., Vainio, 2011), which begs the question against those who adopt other ways of delineating the scope of the moral domain. For instance, Nucci (1985) examined whether religious youth treat their own religious norms more like prototypical moral or conventional norms by assessing whether subjects treat the norms as unchangeable and culture-independent. Since proponents of MFT reject the claim that being unchangeable and culture-independent are properties that distinguish moral norms from nonmoral ones, they have no reason to accept the findings about the impact of religion on the moral domain that result from this approach.

Thus, in order to circumvent the stalemate faced by theory-driven accounts of the moral domain, we favor methods for identifying moral norms, and distinguishing them from nonmoral norms, that do not appeal to the principles of any theory, and thus avoid begging the question against *other* theories. For while theorists cannot agree about how to identify moral norms, it remains possible to study how, in different groups, ordinary people identify moral norms, using a data-driven approach.[1] For instance, Obeid, Argo, and Ginges (2017) look at the relationship between degree of identification with religion and the types of issues that subjects explicitly judge to be moral issues. A similarly direct approach is employed by Buchtel and colleagues (2015) and Dranseika and colleagues (2018), who found dramatic differences in moral concepts across cultures (see also Padilla-Walker & Jensen, 2016, for a questionnaire that explicitly asked subjects about moral values). However, no previous studies have used this direct approach to examine how members of different religions make explicit judgments about which issues count as moral. Our work presents the first data comparing how observant members of four different religious groups, as well as secular individuals in the United States, conceptualize the scope of the moral domain.

Our main goal, then, is to examine the influence of different religious affiliations on conceptions of the scope of the moral domain instead of conceiving of people as

---

1. We hasten to add that we do not expect there to be a single, universal folk concept. Rather, we expect the concept to vary across groups and perhaps even individuals.

either religious or non-religious. We generate a dataset that speaks to whether people of different religions conceive of the moral domain differently. These data need to be explained by any theory that offers a view on how religious affiliation impacts conceptions of the moral domain. To the extent that SDT and MFT do so, their proponents will need to revise and expand their theories to be able to accommodate the diversity represented across religions.

To that end, we developed a novel technique ("the Moral Questionnaire") that is theoretically neutral concerning what counts as *morally* wrong, as opposed to wrong in another way. To do this, we asked participants to categorize a wide range of norms as moral or non-moral, and we used hierarchical cluster analysis to recover implicit differences in conceptions of morality. Studies 1 and 2 validated the technique. Study 3 examined the relationship between religious affiliation and conceptions of morality.

STUDY 1

The goal of Study 1 was to demonstrate the capacity of the Moral Questionnaire to characterize the conception of morality in two populations that are thought to distinguish moral from non-moral norms in a particular way. Philosophers have explicitly theorized about a distinction between moral and non-moral norms (e.g., Wallace & Walker, 1970). Likewise, an analogous distinction is claimed to be present in non-philosopher, adult liberals living in the U.S. (Graham, Haidt, & Nosek, 2009; Huebner, Lee, & Hauser, 2010; Sousa & Piazza, 2014).[2] Both of the populations we sampled tend to consider norms related to harm and rights as being moral and other sorts of norms as being non-moral. The Moral Questionnaire will be validated if it can recover the moral/non-moral distinction that has already been documented in these two populations.

METHOD

Note on Demographic Data. This study's contribution relies on sampling subjects from religiously observant populations. As with any study that selects subjects from a particular demographic group, it is vital to gather and report substantial demographic data on the samples that were obtained. Due to space constraints, we cannot present all of those data in this article. However, throughout the article we refer to supplemental figures and tables (e.g., "Fig. S1" or "Table S1") that contain useful demographic information and more fine-grained analysis of the data than we can report here. Any figure or table labeled with an "S" (as well as all other supplementary materials) can be found at <u>osf.io/em9ga/</u>.

---

2. Some hypothesize that the distinction between moral and nonmoral norms on the basis of harm and rights may in fact be universal, while others suggest that this distinction is confined to liberal U.S. resident adults. The distinction between moral and nonmoral norms could also be categorical or could be a matter of degree. While Graham and colleagues (2009) asked participants to judge whether various actions or situations are right or wrong, they also claim to describe "moral differences" across the political spectrum. As we will argue in the general discussion, this seems to assume that these differences are about issues that liberals and conservatives would classify as being a moral rather than a non-moral matter.

Participants. Philosophers were recruited by advertising the study on philosophy blogs, including The Leiter Reports (http://leiterreports.typepad.com/blog/2009/10/a-survey-of-philosophers-about-moral-judgments.html) and the now defunct Experimental Philosophy blog (http://philosophycommons.typepad.com/xphi/2009/10/the-survey-says.html). Philosophers were not paid for participation. Non-philosophers were recruited on Amazon Turk and received a small amount ($1.00) in exchange for their participation. Participation was limited to U.S.-resident Amazon Turk workers. All participants were over 18 years of age.

A total of 1,256 participants took part in the study through the link posted on philosophy blogs. Of these 1,256 participants, 701 were philosophers, and the remaining 555 were excluded from analyses. Eleven additional participants were excluded because they did not complete the survey, leaving a sample of 690 participants. The mean age of the philosophers was 34.4 years (*SD* = 11.7) and 78.8% of them were men. Complete demographic information is reported in Table S1. A total of 228 non-philosophers were recruited on Amazon Turk; 29 were excluded because they did not complete the survey, leaving a sample of 199 participants. The mean age of the non-philosophers was 33.5 years (*SD* = 12.0) and 43.2% of them were men. The sample leaned non-religious (mean = 2.2, *SD* = 1.3, on a 1–5 scale of religiosity with 1 being "not at all" religious and 5 being "totally" religious) and liberal (mean = 3.3, *SD* = 1.8 on a 1–7 scale with 1 being very liberal and 7 being very conservative). Complete demographic information is reported in Table S1.

Participants (*n* = 690 professional philosophers, *n* = 199 non-philosopher U.S. residents) completed an online survey. The survey consisted of 20 norms drawn from the *Atlas of World Cultures* (Murdock, 1981) and the eHRAF Database (Human Relations Area Files Database), except for the Western norms, which were generated on the basis of authors' own intuitions (see Table 1). Each norm was followed by two questions in a fixed order: (1) *The agreement question* concerned whether people in the participant's own country should abide by the norm, and (2) *the moral judgment question* concerned whether the participant's judgment in response to the agreement question was a moral judgment or not.

Participants were asked to respond to the two questions on a 7-point Likert scale, labeled as shown in the example below.

(1) In Vanuatu, many people think that men should wear colorful shirts on public holidays. Using the scale below, please indicate whether YOU think that men IN YOUR COUNTRY should wear colorful shirts on public holidays.

−3	−2	−1	0	1	2	3
Strongly Disagree			Neither Agree nor Disagree			Strongly Agree

(2) Now consider the judgment you just made. Is that a moral judgment or some other kind of judgment?

−3	−2	−1	0	1	2	3
Clearly NOT a Moral Judgment			Not a Clear Case			Clearly IS a Moral Judgment

TABLE 1. List of Norms Used in Study 1

Label	Full Text of Norm
Pain to Animals	In England, many people think that people should not inflict pain on animals merely for amusement.
Returning Lost Wallet	In Norway, many people think that if a person finds a stranger's wallet in the street and the wallet has the owner's ID and phone number, the finder should attempt to contact the owner and return the wallet to him.
Honor Killing	Many people in the Middle East think that if an unmarried woman is involved in a sexual relationship, she should be harshly punished and perhaps killed by her own relatives.
Cutting Hands	Many members of the Kuna tribe in Panama think that thieves should be punished by having their hands cut off and tied around their neck.
Punishing Family	In the Philippines, many Ifugao people think that the entire family of a murderer is responsible for the murderer's action, and that a family member who did not commit the crime may be executed instead of the offender himself.
Forced Sexual Act	In Papua New Guinea, many Sambia people think that young boys should perform oral sex on older boys and drink their semen.
Bus	Many Canadians think that on a crowded bus young people should offer their seats to elderly people.
Circumcision	Many people in Turkey think that boys should be circumcised near the age of puberty.
Forced Marriage	Many people in India think that marriages should be arranged by parents and that young people should marry the partner chosen by their parents.
Imposed Modesty	Many people in Saudi Arabia think that a woman should always dress modestly and cover her head when in public.
Birth Control	Many People in Honduras think that married couples should not use artificial birth control.
Head Shape	Many Tlingit people think that the head of an infant should be compressed immediately following birth to give it what they consider a pleasing form, in which the eyebrows are drawn up and the nostrils stretched asunder.
Colorful Shirt	In Vanuatu, many people think that men should wear colorful shirts on public holidays.
Thai	In Thailand, many people think that Thai food should be eaten with a fork, not with chopsticks.
Sunscreen	Many Australians think that fair-skinned people should use sunscreen when they are on the beach.
Crossing Street	Many people in Scotland think people should look both ways before crossing the street.
Oyster	Many Americans think people should not eat raw oysters in June, July or August.
Pasta	Many people in Italy think that adults should not eat pasta with their fingers.
Professor First Name	Many Canadians think that university students should not address a professor by his or her first name unless the professor has said that they should do so.
Judge	Many people in France think that when a judge enters a courtroom the people seated in the courtroom should stand up.

The norms were presented in 5 different orders, and participants were randomly assigned to one of these possible orders. For a full list of the norms used, see Table 1. Participants then answered a series of demographic questions (see Tables S2 and S3).

Crucially, asking subjects to answer the first question (about whether people in their country should follow the norm) was simply a way to get them to generate a judgment, so we could then ask them if that judgment counted as moral or not. The second question (the moral judgment question) was the main target of our analyses. Answers to this question revealed how our subjects conceptualized the scope of the moral domain.

RESULTS

Cluster Analysis. Cluster analysis is a family of algorithms that divide items (participants in an experiment, survey questions, organisms, cities, etc.) into groups or "clusters" (e.g., Everitt, Landau, Leese, & Stahl, 2011). In general, cluster analysis algorithms divide items into clusters by maximizing their similarity within groups and minimizing it across groups. They differ from one another, among other things, because they involve different strategies for maximizing similarity within groups and minimizing it across groups. Clustering can be hierarchical (clusters are sub-clusters of larger clusters) or partitional (i.e., non-hierarchical). Hierarchical cluster analysis outputs are typically represented by dendrograms, which group together items in a hierarchical manner. Visually, a dendrogram describes a hierarchy of increasingly inclusive clusters.

In our article, we use hierarchical clustering, so as to group clusters into larger clusters, and report the results by means of dendrograms. We also use between-groups linkage as our default clustering method, and squared Euclidean distance as our similarity measure, but we test the robustness of our results by using several clustering algorithms.[3] Between-groups linkage defines the distance between two clusters, *A* and *B*, as the average of the distance between all pairs of *x* and *y* such that *x* belongs to *A* and *y* to *B*.

Cluster analysis has been extensively used in psychology, biology (including genetics), and sociology. One of its main uses was the classification of organisms into taxa (Sneath & Sokal, 1973). In psychiatry, cluster analysis has been used to identify types of psychiatric syndromes such as depression (Everitt, Gourlay, & Kendell, 1971; Kendell & Jablensky, 2003; Paykel, 1971). In psychology, cluster analysis has been used to identify clusters of emotion concepts and to compare these clusters across cultural contexts (e.g., Church, Katigbak, Reyes, & Jensen, 1998; Fontaine, Poortinga, Setiadi, & Markam, 2002).

Cluster analysis suffers from a few limitations. First, it is not an inferential statistical method. There is thus no power analysis associated with it, although rules of thumb have been proposed in the literature. We have followed these recommendations (see supplementary materials). Second, identifying the optimal number of clusters in a hierarchical cluster analysis can be difficult. As is commonly done, we identified the stage at which the difference between the distance coefficients (a measure of the

---

3. Between-groups linkage combined with squared Euclidean distance is a common setting for hierarchical clustering, and it is indeed the default setting in SPSS. Alternative clustering methods such as single and complete linkages are not appropriate since they do not take into account all the components of a given cluster. Other methods such as Centroid linkage or Wards' method would also be appropriate. We take Study 2 to validate the use of between-groups linkage combined with squared Euclidean distance, but we examined the robustness of our analyses in Study 3.

within-cluster similarity) at two consecutive clustering stages suddenly increased, indicating a sudden increase in the heterogeneity of the clusters. Third, there are several clustering methods to choose from: As noted, we examined the robustness of our results.

The analysis was conducted with SPSS. The variables (i.e., the answers to Question 2 for all 20 norms) were clustered. The clustering method chosen was between-groups linkage, and the similarity measure was squared Euclidean distance. The data were not standardized.

Results. For both groups, at the highest-level clusters, norms related to unjustified harm and infringement on freedom were separated from norms related to prudence, authority, and ways of life (see Figure 1). Participants judged norms in the former cluster as significantly more moral than norms in the latter cluster (philosophers: $t(18) = 11.88$, $p < .0001$; non-philosophers: $t(18) = 7.62$, $p < .0001$; see also Supplemental Results and Tables S4–S7 for individual item analyses). These results suggest that the Moral Questionnaire succeeds at identifying the distinction between moral and non-moral norms that has been independently documented in philosophers and lay U.S. residents.

Convergent evidence that the Moral Questionnaire is picking up the moral/non-moral distinction made by philosophers and by U.S. residents comes from the mean moral judgment ratings that subjects gave the items in each group. On our scale, a score of 0 represented the midpoint, labeled "Neither moral nor non-moral," while a score of 3 was labeled "Clearly IS a moral judgment," and a score of –3 was labeled "Clearly NOT a moral judgment." For both philosophers and non-philosophers, the cluster containing the items related to harm and rights was rated as significantly above the midpoint on the moral judgment question (philosophers: $t(7) = 11.66$, $p < .001$, non-philosophers: $t(7) = 8.65$, $p < .001$), while the cluster containing the other norms was rated as significantly below the midpoint (philosophers: $t(11) = –8.05$, $p < .001$, non-philosophers: $t(11) = –4.05$, $p = .002$). (See supplemental materials for item-by-item analysis of the moral judgment question and the agreement question for philosophers and non-philosophers, Tables S4–S7.)

STUDY 2

Study 2 was designed to provide further convergent evidence that the Moral Questionnaire is picking up on the moral/non-moral distinction made by philosophers and liberal U.S. resident adults. Do norms that are grouped into the "moral cluster" by the Moral Questionnaire have "moral" characteristics? We tested whether the norms grouped into the moral cluster in Study 1 have the distinctive features that theories of moral judgment attribute to moral norms. Although there is no consensus about what makes a norm a moral norm (Stich, 2018), most agree that violations of moral norms elicit negative third-party judgments (Graham et al., 2013), are coupled with negative affective reactions such as disgust or anger in third-party observers (Haidt, Koller & Dias, 1993; Prinz, 2006), and elicit guilt or shame in transgressors (Prinz, 2006; Tangney, Miller, Flicker, & Barlow, 1996); further, moral norms cannot be nullified by an authority figure (Smetana, 1981). We tested whether norms in the moral cluster were associated with these characteristics of moral norms more than norms in the non-moral cluster.
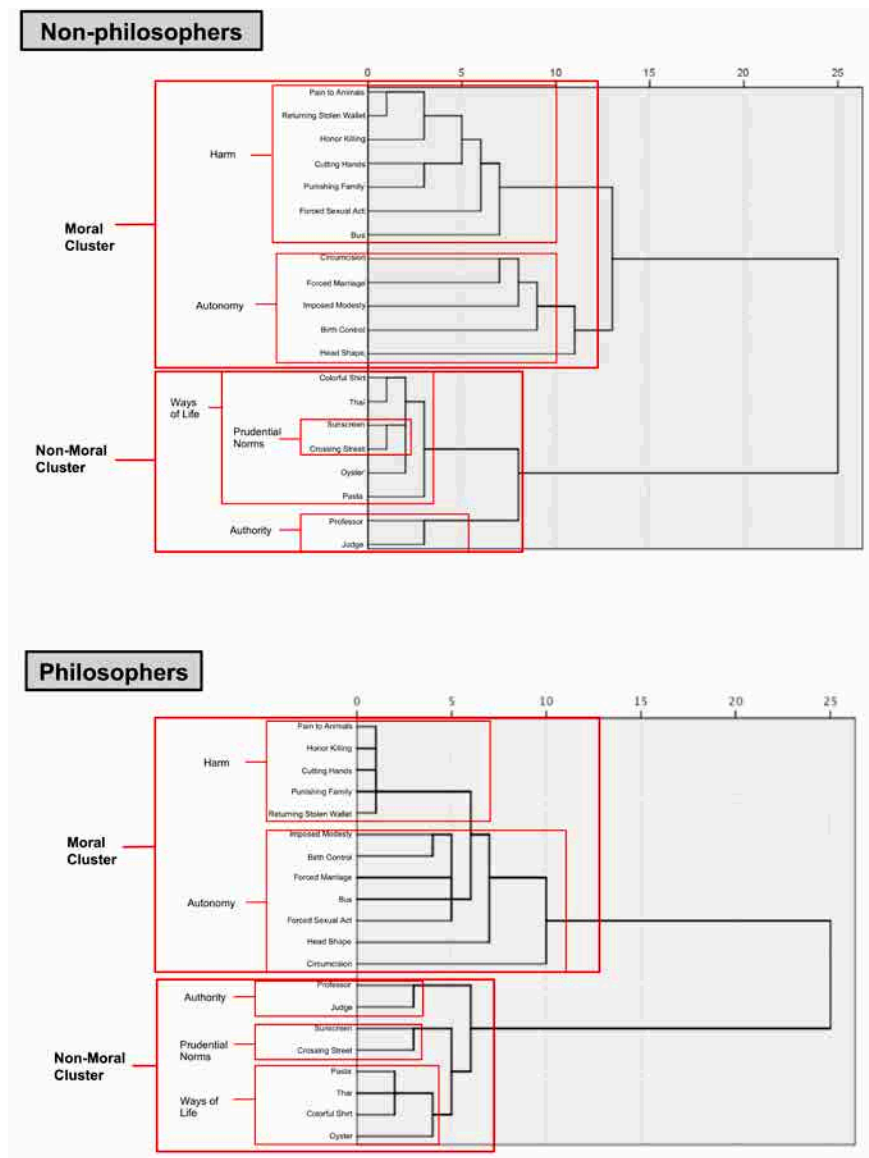
FIGURE 1. Dendrogram depicting participants' answers to the moral judgment question in Study 1. Upper panel describes results from the non-philosophers. Lower panel describes results from philosophers. Across both groups, the same 12 norms are grouped into the moral cluster and 8 norms into the non-moral cluster. Sub-clusters differ slightly across the groups.

METHOD

A new sample of U.S. resident participants (n = 143) was shown 16 of the 20 norms used in Study 1 (see Table S8).

The survey consisted of 16 of the 20 norms used in Study 1 presented in random order. When non-philosophers in Study 1 on average agreed with the application of the norms in their own country (as determined by the mean answer to *the agreement*

question; see Table S1), participants in Study 2 were asked to imagine the norm being *violated*. When non-philosophers in Study 1 disagreed with the application of the norm in their own country, participants in Study 2 were asked to imagine that someone *followed* or *enforced* the norm.

Three norms from Study 1 (Oyster, Thai, and Colorful Shirt) were not used in Study 2 because they were not clearly endorsed or condemned by the non-philosopher participants in Study 1 (see Table S5); agreement ratings for these norms were not significantly different from the midpoint ("neither agree nor disagree"). The other norms were either significantly endorsed or condemned by the non-philosopher participants in Study 1 (see Table S5). One additional norm (Cutting Hands) was excluded because the dependent variables used in Study 2 ask participants to make judgments about individual norm violators, as opposed to government/organizational institutions, and it was not clear what individual would act on this norm. Table S6 lists which norms were classified as moral and which as non-moral, based on the cluster analysis of the non-philosopher data from Study 1.

Participants were randomly assigned to one of three conditions, and asked to imagine that the norm violator or enforcer was (A) someone living in the participant's community ($n = 49$), (B) someone living in the location where the norm is prevalent ($n = 46$), or (C) someone living in a different country ($n = 48$). Participants judged the actor on 10 measures: wrongness of action, seriousness of transgression, harmfulness of action, shame felt, guilt felt, blameworthiness, punishment deserved, disgust induced in observer, anger induced in observer, and how bothered an observer would be.

RESULTS

Statistical Analysis. Statistical analysis was initially conducted with SPSS. Norms were first divided for analysis into two categories: moral and non-moral. This division was based on the cluster analysis of non-philosophers in Study 1 (see Figure 1; Table S8). Average participant judgments were computed for norms in each category. Means are graphed in Figure 2. Means with judgments broken down by condition (participant's community, community of the norm, different country) are graphed in Fig. S1. Norms were then further divided into five sub-categories based on the sub-clusters that emerged in Figure 1 (non-philosopher participants). Average participant judgments were computed for norms in each sub-category; these means are graphed in Fig. S2.

Study 2 ratings were additionally analyzed in R (Version 3.6.2) with a linear mixed model fit by REML. Fixed effects were the between-participants Condition, the Item Category (moral or non-moral, classified based on the cluster analysis of the non-philosophers in Study 1), the Question Asked, and the interaction between Item Category and Question Asked. Random intercepts were included for Item and Participant. Data were analyzed with the following syntax (using the lme4 and lmerTest packages): lmer(Rating ~ Condition + Category*Question + (1 | Item) + (1 | SubjectNumber), data = Study2).

The linear mixed model indicated that responses on the 10 dependent variables (DVs) were higher for items that were classified as moral compared to those that were classified as non-moral, $B = 1.80$ ($SE = 0.58$), $p = .008$. Using the "how wrong?" question as the baseline, there were few significant interactions between the Question Asked and the moral
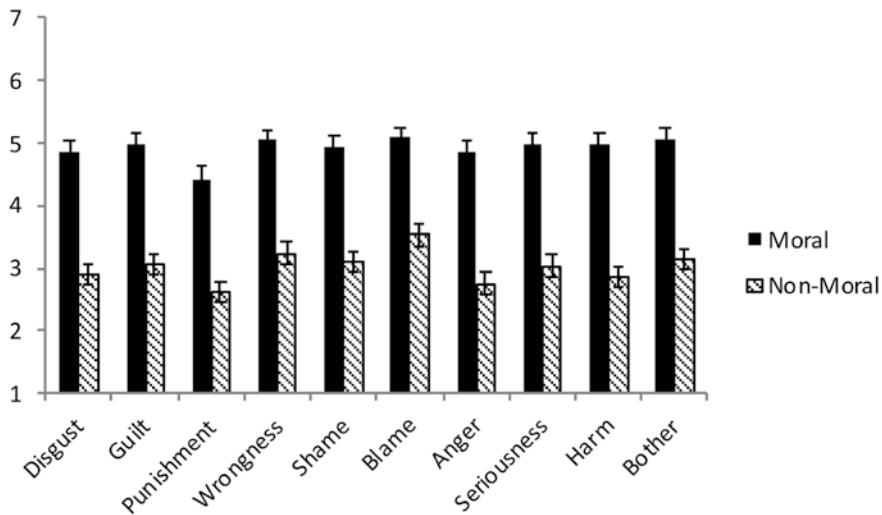
FIGURE 2. Norms falling in the moral cluster are rated higher than norms falling in the non-moral cluster on each of the 10 distinctive properties of moral norms in Study 2. Error bars show standard errors of the mean. Data are collapsed across the three conditions (see Fig. S1 for data broken down by condition).

status of the item. The only interactions that were uncovered were for "how blameworthy is the person?" ($B = 0.25$, $SE = 0.11$, $p = .024$), "how angry would you feel toward the person?" ($B = -0.29$, $SE = 0.11$, $p = .009$), and "how harmful is the person's action?" ($B = -0.32$, $SE = 0.11$, $p = .003$). (All other interactions: ps > .10.) This indicates that moral norms differed from non-moral norms to a similar extent across most of the dependent variables, suggesting that the moral/non-moral distinction does not generally influence ratings more powerfully for certain broad dimensions such as condemnatory emotions or moral principles (see Figure 2). Using Condition A (participant's community) as the baseline, Condition B (community of the norm) produced significantly different ratings ($B = -0.34$, $SE = 0.11$, $p = .002$), whereas Condition C (different country) did not ($B = -0.08$, $SE = 0.17$, $p = .639$) (see Fig. S2). Overall, this analysis shows that participants treated norms from the moral cluster in Study 1 as having the distinctive characteristics that many theorists independently ascribe to moral norms (Graham et al., 2013; Prinz, 2006; Smetana, 1981): Violations were judged as warranting a significantly more negative reaction in the transgressing agent (shame, guilt) and in a third party observer (blame, disgust, anger, and bother), and they were treated as more wrong, harmful, and deserving of punishment than violations of norms from the non-moral cluster.

STUDY 3A

Having validated the technique of using the Moral Questionnaire, Study 3a tested the relationship between religious affiliation and participants' conceptions of the moral domain. Five target populations were selected: Hindus, Jews, Mormons, Muslims, and secular individuals.

These four religions were not chosen because of how common they are in the United States or in the world. Rather, our goal was to generate a dataset that might express some degree of variability in the relationship between religious affiliation and people's conception of morality. The religions were selected based on that goal and also based on our ability to locate religious informants who were able to generate lists of norms. Even with this small subset of religions, our results show substantial variability in the conception of the moral domain across religions, challenging current theories to modify their views to capture the data. It is possible, of course, that even more heterogeneity exists across religions than we have reported here. In addition, because a sample of 50 participants cannot be representative of a whole religious population, we cannot exclude the possibility that there are different conceptions of morality within a religion, even within the United States alone. This is particularly true for Hinduism, which is often described as a set of closely related religions or a synthesis of prior traditions that merged over hundreds of years (Flood, 2008). But we can at least conclude that for some religious groups, the relationship between religious affiliation and conceptions of morality cannot be explained by current theories.

METHOD

Two informants from each of the four religious groups, blind to the goals of the study, made lists of norms that are followed by members of their community. Ten norms from each religion were randomly selected to be included in the survey. Four moral and four non-moral norms from Study 1 were also included, plus one additional paradigmatically moral norm and one additional paradigmatically non-moral norm, creating a survey with 50 norms (see Table 2). As in Study 1, two questions were asked about each norm—(1) the agreement question and (2) the moral judgment question—followed by a demographic questionnaire including a religion-specific religiosity measure (see Fig. S3; Tables S9–S12).

We decided to collect 50 participants for each of the four religions of interest (Judaism, Mormonism, Islam, and Hinduism) as well as 50 secular participants. Data collection stopped when 50 participants completed the whole survey and met our exclusion criteria (see Supplemental Materials). Five additional Hindu participants were accidentally allowed to take the survey, so their data were included in the analysis.

Participants were recruited by a professional data collection service (Qualtrics). As in Study 1, we were primarily interested in the moral judgment question. Answers to the agreement question reveal whether people from different religions have the same norms. It is already well known that people from different religions follow different norms. Our interest, instead, is in whether people of different religions make different judgments about what counts as a *moral* issue. Additional information about the methods for Study 3a can be found in the supplementary materials.

RESULTS

Statistical Analysis. The plan for analysis followed that of Study 1. Statistical analysis was conducted with SPSS. The data were analyzed by means of hierarchical cluster analysis. The variables (i.e., answers to Question 2 for the 50 norms) were clustered.

The clustering method chosen was between-groups linkage, and the similarity measure was squared Euclidean distance.[4] The data were not standardized. We also computed the means of participant responses to the moral judgment question by participant group (see Tables S12 and S13; see also Table S15 for the means and SDs for each DV for each norm). We then identified the stage at which difference between the coefficients at two consecutive clustering stages suddenly increased, indicating a sudden increase in heterogeneity. For the secular, Jewish, and Mormon participants, we identified the two clusters that emerged at the highest hierarchical level. We averaged responses to all norms in each cluster to determine the average moral judgment rating for each cluster. We compared the means using two-tailed *t*-tests and treated the cluster with the highest rating as the "moral cluster" for comparison across religions. For the Muslim cluster, we identified three clusters (including a cluster composed of three non-moral norms). We conducted an ANOVA to determine if one cluster was rated as more moral than the others and conducted post hoc *t*-tests (equality of variance assumed) to find the most moral cluster. This cluster was used as a comparison point with the moral clusters of the other religions. Finally, no clear clusters emerged in the Hindu cluster diagram. This rendered the rest of the analysis plan moot for that group.

To compare the moral clusters across participant groups, we examined two primary questions: (1) Do the norms that appear in the secular moral cluster appear in the moral clusters of other groups? (2) Do the norms of a particular religious group appear in its moral cluster?

We were also interested in looking at the relationship between how strongly a norm is held and participants' assessments of whether that norm is moral. One way of doing this is to calculate the correlation between the agreement question and the moral judgment question. However, it is possible that a better predictor of the moral judgment question would be the strength of conviction with which a participant agreed or disagreed with a norm. That is, a participant's judgments falling on either end of the agreement scale (indicating that they strongly agreed or strongly disagreed with the norm) may predict whether the participant considered the norm moral. To evaluate this possibility, an average "conviction strength" was computed for each norm. Conviction strength was computed by taking the absolute value of the distance of each participant's judgment from the midpoint of the scale. Higher values thus indicated higher conviction strength. For example, participants rated their agreement with the statement that people in their community should follow each norm (as explained in the methods section above). Responses were coded on a scale from 3 (participants indicating that they strongly agree) to –3 (participants indicating that they strongly disagree), with 0 indicating that the participant neither agreed nor disagreed with the norm. If a participant's judgment for a particular question was –2, then that judgment would transform into a conviction strength judgment of 2 (indicating 2 units from the midpoint judgment). Conviction strength judgments therefore fell on a scale from 0 (little conviction) to 3 (high conviction). The correlation between average conviction strength and average moral judgment ratings was calculated, treating the norm as the unit of analysis (see Table S14).

---

4. We examined the robustness of our analyses using Ward's method combined with squared Euclidean distance (see footnote 1). The results were not substantially affected.

TABLE 2. List of Norms Used in Study 3a

Label	Full Text of Norm
Animal Pain	Many people in England believe that people should not inflict pain on animals merely for amusement.
Arabic	Many Muslims believe that prayer should be recited in Arabic.
Baptism	Many Mormons believe that people should receive baptism and confirmation in their church.
Basil Worship	Many Hindus believe that they should worship the basil plant.
Blessed Bread & Water	Many Mormons believe that they should consume blessed bread and water each Sunday at church.
Blessings for Sick	Many Mormons believe that the seriously sick or injured should receive blessings from two priesthood holders in addition to medical attention.
Church Donations	Many Mormons believe that ten percent of their income should be donated to the Church.
Clothes	Many Hindus believe that they should not wear anything but Indian clothes to a Hindu temple or event.
Colorful Shirts	Many people in Vanuatu believe that men should wear colorful shirts on public holidays.
Coming of Age Rituals	Many Hindus believe that specific rituals should be held to commemorate a son's coming of age.
Court	Many people in France believe that when a judge enters the courtroom, the people seated in the courtroom should stand up.
Cow Worship	Many Hindus believe that they should worship cows.
Curfew	Many Hindus believe that parents should enforce a strict curfew on their children.
Deities	Many Hindus believe that they should worship many deities.
Divorce	Many Jews believe that divorce should be reserved only for particular, concrete reasons.
Dress	Many Muslims believe that women should dress conservatively and wear a headdress in the presence of non-mehram (men other than their husband/brother/father).
Driving Left Side	Many people in South Africa believe that people should drive on the left side of the road.
Education	Many Jews believe that parents should financially ensure that their children are well educated.
Elder	Many Jews believe that they should listen quietly when elderly persons speak.
Emergency Provisions	Many Mormons believe that they should keep a store of food and provisions to sustain themselves in case of emergency.
Family Religious Study	Many Mormons believe that, as a family, they should spend time together daily to study their religious texts.
Foul Language	Many Mormons believe that they should not use foul language.
Groom Dowry	Many Muslims believe that a groom should pay money to his bride's family.
Hair Wash	Many Hindus believe one should not wash their hair on Saturdays and Tuesdays.
Halal	Many Muslims believe they should only eat Halal meat.
Home Worship	Many Hindus believe that one should worship their home in consideration of the microorganisms killed in the process of making the property.
Homosexuality	Many Muslims believe that no one should engage in homosexual acts.

TABLE 2. (*Continued***)**

Label	Full Text of Norm
Honesty	Many Mormons believe that they should be honest with and kind to their fellow beings.
Hug	Many Muslims believe that women should not hug men.
Kids	Many Jews believe that their family should have at least two children.
Killing	Many people in Honduras believe that people should not kill others for no reason.
Laws	Many Jews believe that they can break almost any religious law in order to save a human life.
Marry	Many Jews believe that they should marry another Jew.
Murder	Many people of the Kuna tribe in the Philippines believe that the entire family of a murderer is responsible for a murderer's action, and that a family member who did not commit a crime may be executed instead of the offender himself.
Nutritious Food	Many Mormons believe that they should eat nutritious food.
Pasta	Many people in Italy believe that adults should not eat pasta with their fingers.
Pray	Many Muslims believe they should pray five times a day.
Prayer	Many Jews believe that they should not disrupt another's prayer.
Professor First Name	Many people in Canada believe that university students should not address a professor by his or her first name unless the professor has said that they should do so.
Religious Services	Many Jews believe that one should not attend services of religions other than their own.
Respect	Many Jews believe that they should not bow out of respect to other people.
Returning Lost Wallet	Many people in Norway believe that if a person finds another person's wallet in the street and the wallet has the person's ID and phone number, the finder should attempt to contact the owner and return the wallet to him.
Right Hand	Many Hindus believe that they should use only their right hand for important actions like eating or serving others.
Serving Food	Many Muslims believe that when serving food or drink they should always serve the eldest guest first.
Sex	Many Muslims believe that a single boy and single girl should not be alone with one another unless their families are considering their marriage.
Shoes	Many Muslims believe that people should not leave shoes upside-down.
Temple Donations	Many Hindus believe that after visiting a temple, one should leave a donation.
Thief	Many members of the Kuna tribe in Panama believe that thieves should be punished by having their hands cut off and tied around their neck.
Vulgarity Around Rabbi	Many Jews believe that it is unacceptable to discuss vulgar topics around a rabbi.
Weekly Letters	Many Mormons believe that they should write weekly letters to family members who are serving missions.

With respect to the hierarchical cluster analysis, our questions of interest were: (1) Does each group make a clear moral/non-moral distinction? (2) Are there any norms deemed to be moral across all groups? (3) Does each religious group moralize its own norms? (4) Are norms unrelated to harm and fairness moralized by each group? (5) Is there a relationship between strength of conviction in a norm and whether that norm is considered moral?

The cluster with the highest average ratings for the moral judgment question was considered the "moral cluster" (see Tables S13–14). For secular participants (see Figure 3), the highest-level clusters divided the norms into a small moral cluster and a large non-moral cluster. Eight of the 50 norms—including all 4 moral norms from Study 1, the additional paradigmatically moral norm, one Mormon norm, one Muslim norm, and one Jewish norm—fell into the moral domain. These eight norms pertained to harm, justice, and rights.

For Jewish participants (see Figure 4), the highest-level clusters divided the norms into a small moral domain and a large non-moral domain. The moral cluster contained exactly the same 8 norms that constituted the moral cluster of the secular participants. This finding is consistent with Social Domain Theory (Nucci & Turiel, 1993), but is unexpected by Moral Foundations Theory, which has emphasized the idea that religiosity leads to a broader moral domain (Graham & Haidt, 2010). Seven of the 10 Jewish norms were located together in a non-moral sub-cluster, suggesting that Jews do regard their own norms as having a special status, even though they do not consider them to be moral norms.

In contrast, for Mormon participants (see Figure 5), the cluster of moral norms was substantially larger than the cluster obtained with secular and Jewish participants. Mormons treated as moral all of the same 8 norms deemed moral by Jews and secular individuals, but in addition, 7 of their own religious norms were included in the moral cluster. The Mormon moral cluster was also divided into sub-clusters of prototypical moral norms and religio-moral norms.

For Muslim participants (see Figure 6), the moral cluster contained 7 of the 8 norms identified as moral by each of the other groups, in addition to 9 of the 10 Muslim norms. Thus, Muslims also moralize their own religious norms, although unlike Mormons, they do not divide the moral domain into religio-moral and prototypical moral sub-clusters. Findings for Mormon and Muslim participants are consistent with Moral Foundations Theory, but are contrary to Social Domain Theory, as many of the Mormon and Muslim religious norms did not pertain to violations of justice or welfare.

For Hindu participants (see Figure 7), no clear moral or non-moral clusters emerged from our sample; the prototypically moral and Hindu religious norms were distributed throughout the diffuse clusters. It is possible that with a larger sample of Hindu participants, a clearer moral/non-moral distinction may have emerged; however, we failed to detect one in this sample. (Note that neither significance testing nor power analysis are available for cluster analysis techniques, as we point out in the Cluster Analysis explanation, see above.) What our results do show is that 50 individuals was a large enough sample to identify a moral/nonmoral distinction among Mormons, Muslims and Jews, but not among Hindus. Thus, if the distinction does exist in the Hindu population, it seems to be more vague and diffuse, and thus harder to detect, than in other populations. One possible explanation for this is that Hindu religious traditions are extremely diverse and that we sampled individuals from different Hindu
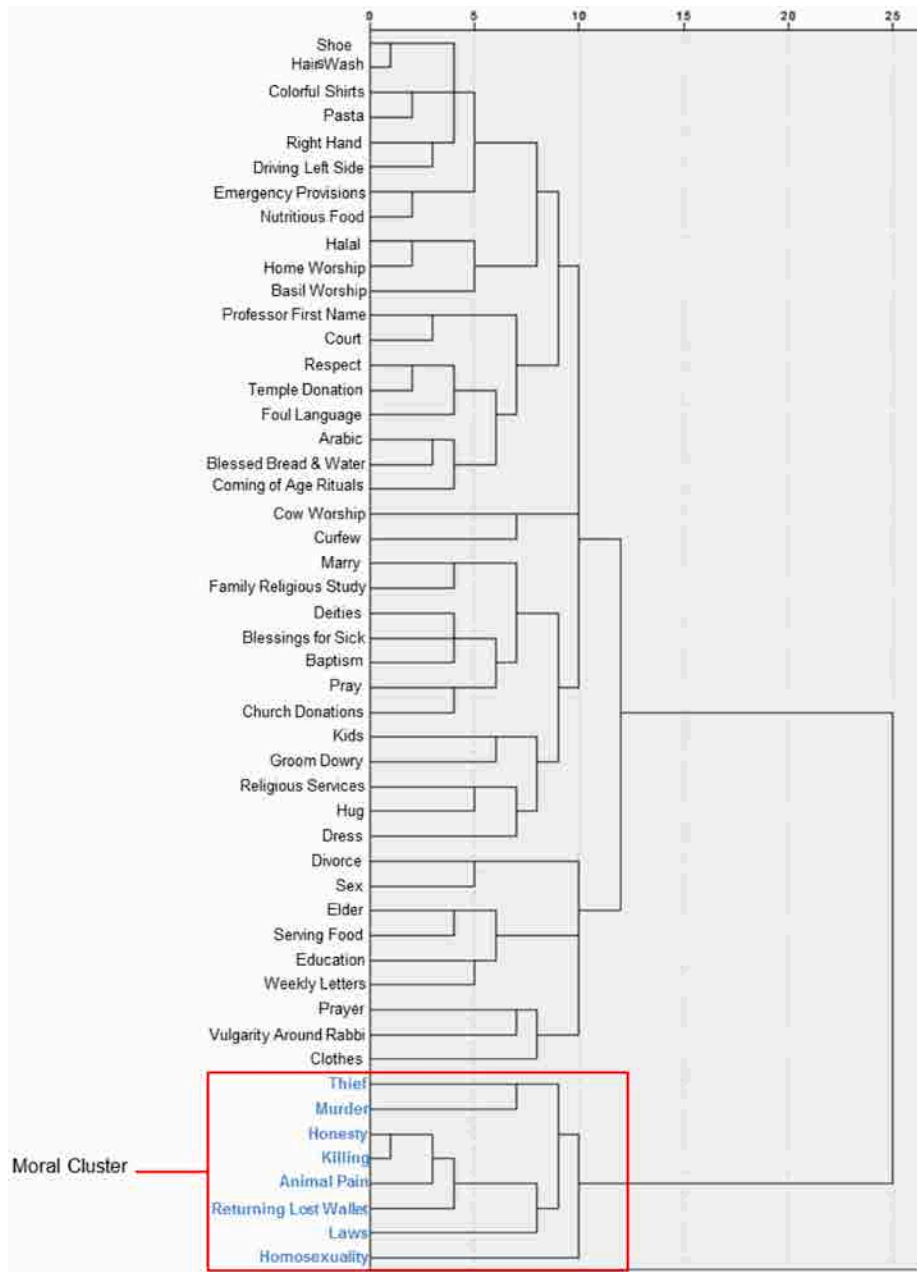
FIGURE 3. Dendrogram of secular participants' responses to the moral judgment question in Study 3a. The highest-level clusters divide the norms into a narrow moral domain (8 norms) and a broad non-moral domain. The norms in the moral cluster are noted with asterisks; these norms are highlighted on Figures 4–7 as well.
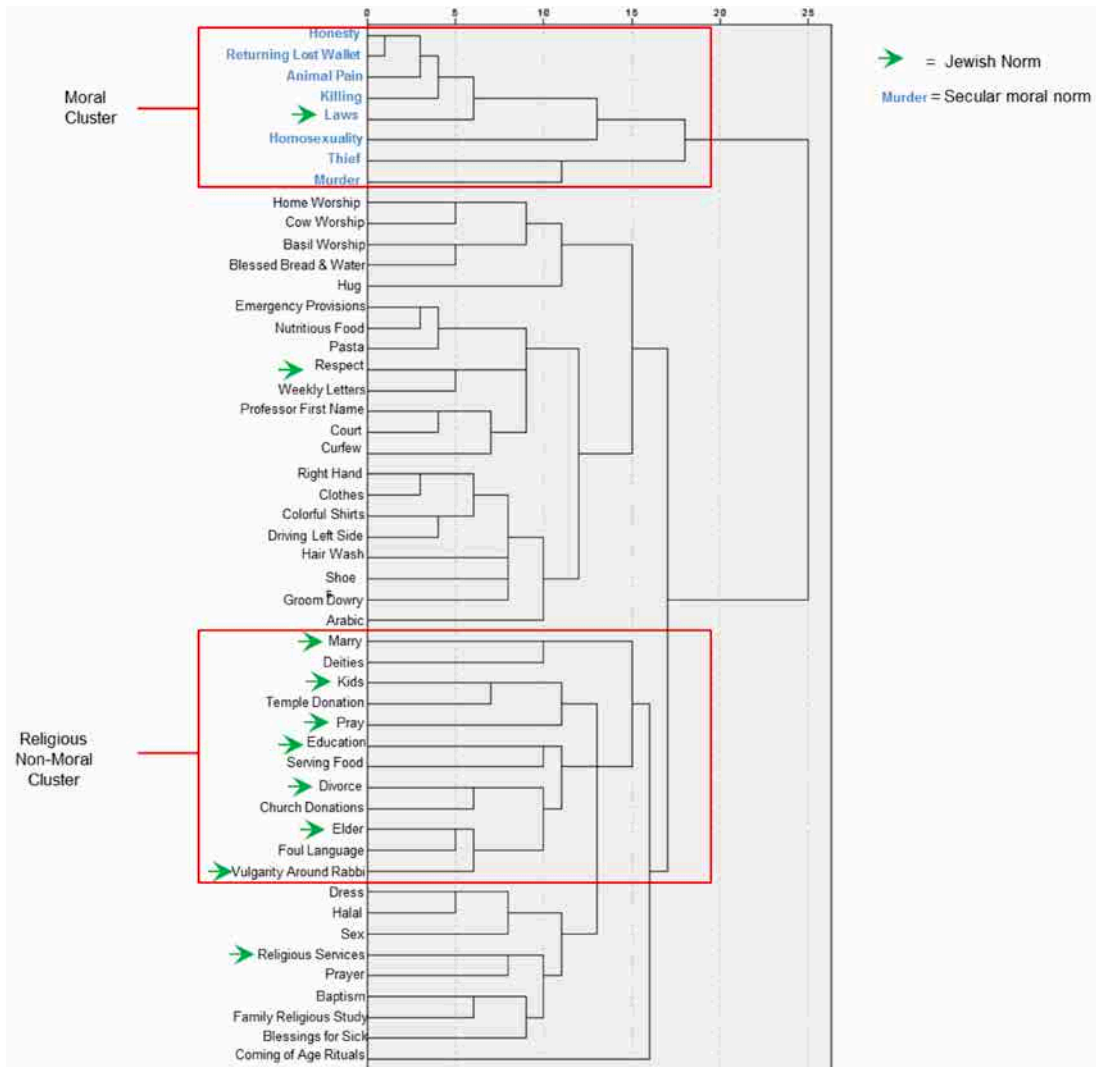
FIGURE 4. Dendrogram of Jewish participants' responses to the moral judgment question in Study 3a. The highest-level clusters divide the norms into a narrow moral domain (the 8 secular moral norms, noted with asterisks) and a broad non-moral domain. The non-moral domain also includes a religious sub-cluster.

sub-sects (Flood, 2008). Another possible explanation is that our Hindu participants did make a moral/non-moral distinction, but that the question we used to measure that distinction failed to detect it.

The findings from the five demographic groups are summarized in Figure 8.

For each group, strength of conviction (how strongly participants agreed/disagreed with a norm) was significantly correlated with ratings on the moral judgment question ($p < .001$; R^2s range from moderate to strong; see Table S15). Thus, strength of conviction is a partial contributor to determining what makes a judgment a moral
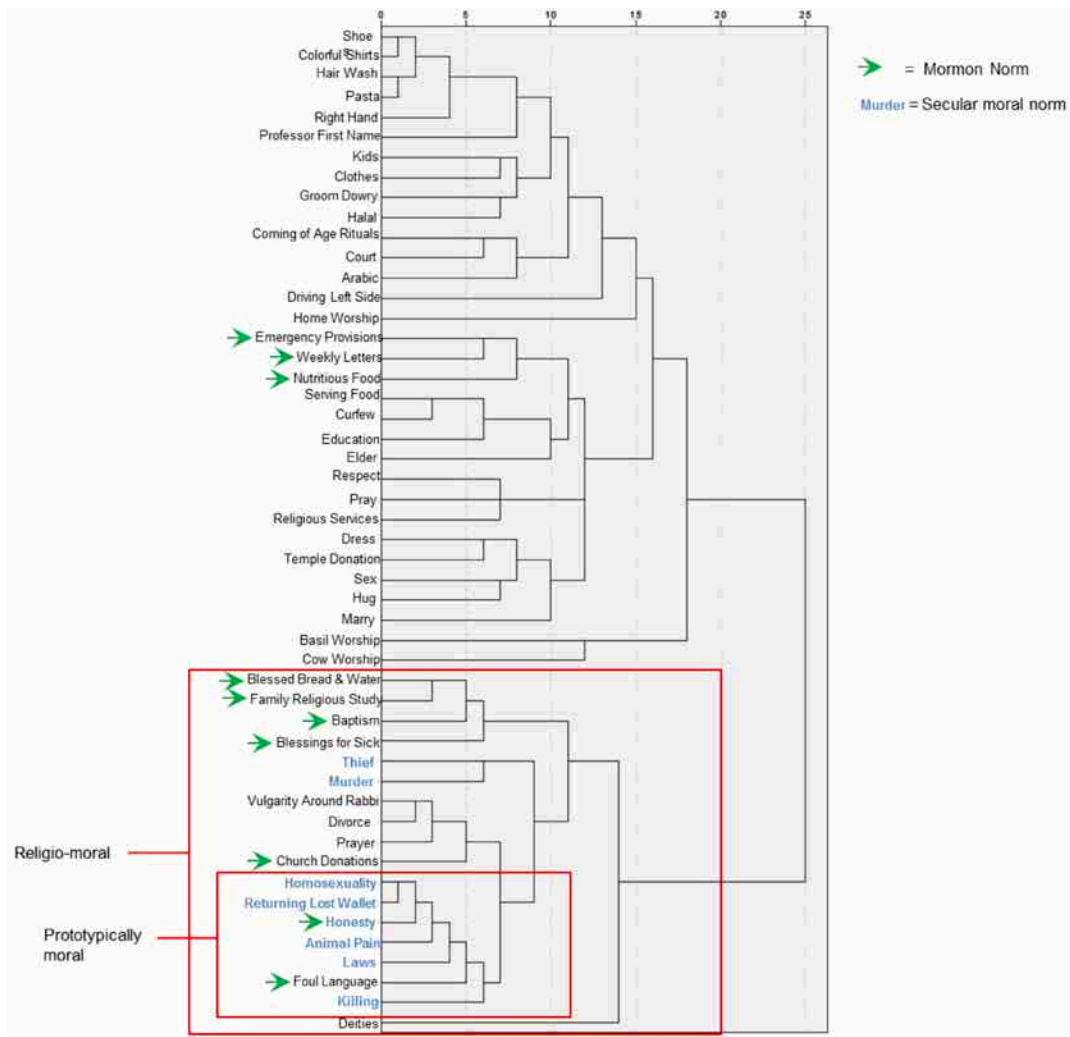
FIGURE 5. Dendrogram of Mormon participants' responses to the moral judgment question in Study 3a. The highest-level clusters divide the norms into a moral domain (divided into sub-clusters of prototypical moral norms and religio-moral norms) and a non-moral domain. Asterisks indicate the norms that were considered moral by the secular individuals.

judgment (Skitka, 2010), although some amount of variance remains to be explained by other factors.

STUDY 3B

To ensure that our findings from Study 3a were robust, we conducted a replication of the study with new samples of religious and secular participants. The study design, materials, methods, and statistical analysis plan were the same for Study 3b as for Study 3a. New samples of Jewish, Muslim, Hindu, Mormon, and secular individuals were recruited (see Tables S17–S19). The main findings were replicated (see Figs. S5–S9).
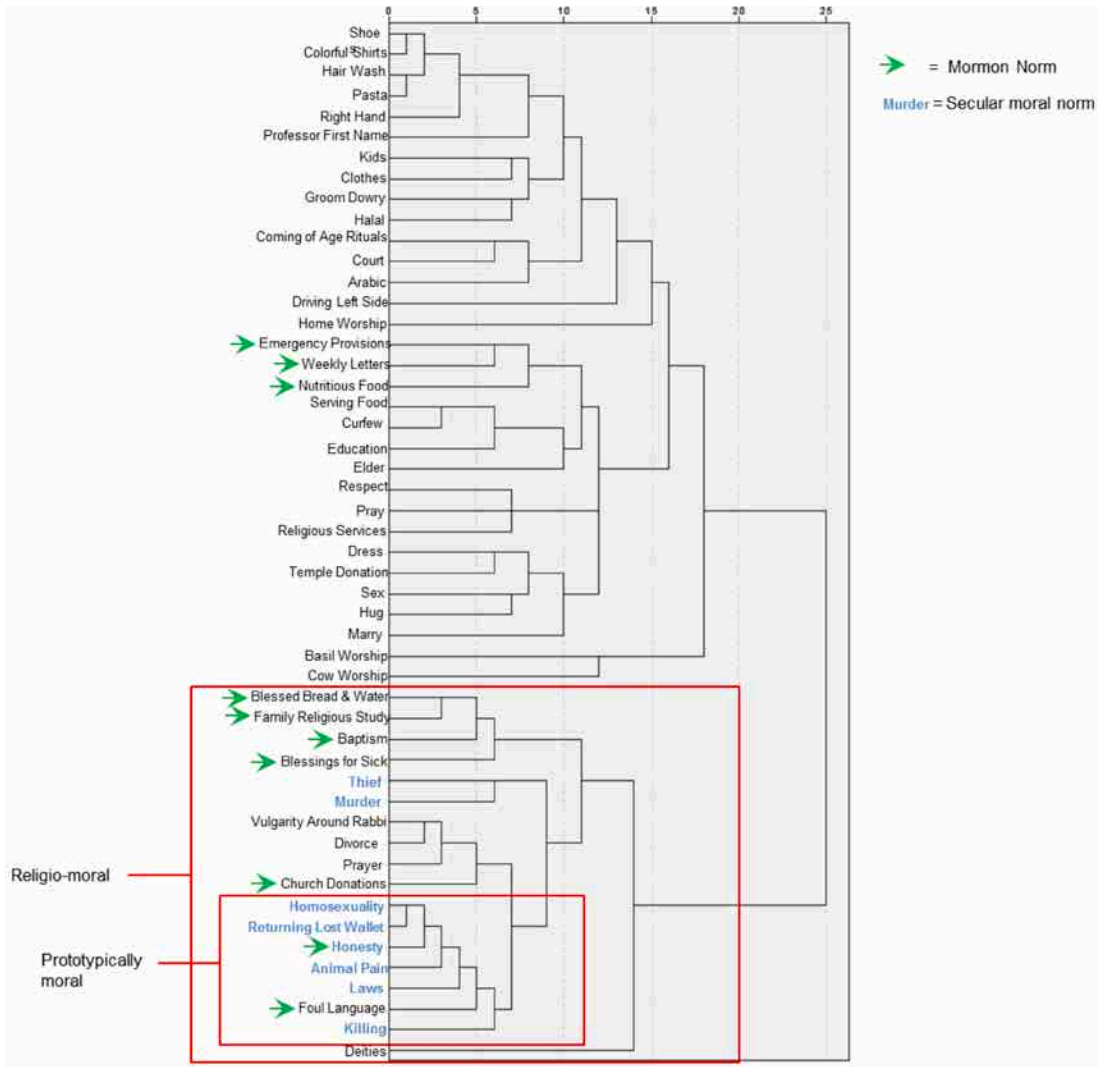
FIGURE 6. Dendrogram of Muslim participants' responses to the moral judgment question in Study 3a. The highest-level clusters divide the norms into a moral domain (containing many of the Muslim religious norms spread evenly through the cluster) and a non-moral domain. Asterisks indicate the norms that were considered moral by the secular individuals.

GENERAL DISCUSSION

We set out to generate a theory-neutral dataset that shows similarities and differences in the ways that members of different religions conceive of the moral domain. We suggest that these data need to be explained by any theory that wishes to describe the influence of religion on the conception of the moral domain. The main finding, in the most general terms, is that members of different religions conceive of the moral domain differently. When put crudely, the finding is perhaps unsurprising, but the
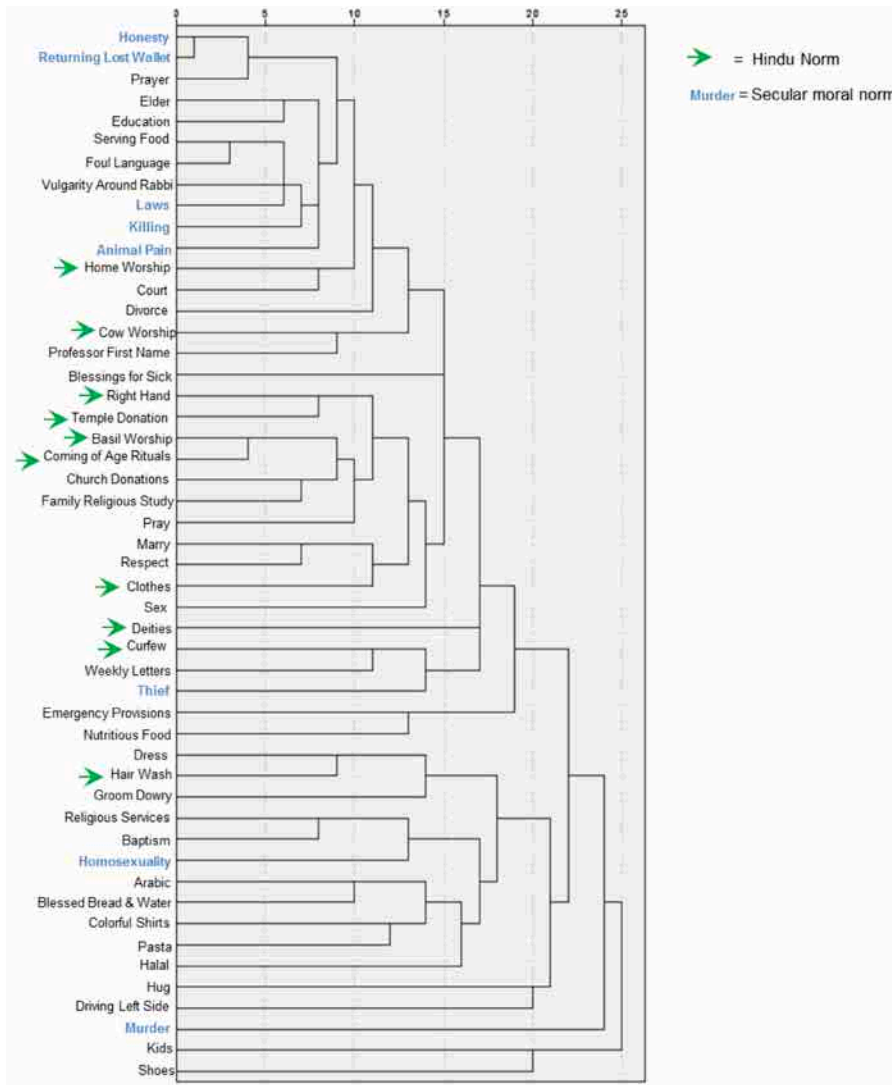
FIGURE 7. Dendrogram of Hindu participants' responses to the moral judgment question in Study 3a. No clear moral/non-moral distinction appears. Asterisks indicate the norms that were considered moral by the secular individuals.

leading theories that offer views on the impact of religion on the moral domain have treated religious individuals as a more-or-less homogenous group, not specifying how different religions might differently impact the moral cognition of their adherents.

We have found that it is neither true that secular people and religious people share a common conception of the moral domain nor that religious morality is expanded beyond secular morality in a uniform manner. Further, when participants in a group did make a moral/non-moral distinction, there was broad agreement that norms related to harm, justice, and rights counted as moral norms. However, some religious

			Secular	Jews	Mormons	Muslims	Hindus
Secular	Moral	Animal Pain					
		Killing					
		Murder					
		Returning Lost Wallet					
		Thief					
	Non-moral	Colorful Shirts					
		Court					
		Driving Left Side					
		Pasta					
		Professor First Name					
Jewish		Divorce					
		Education					
		Elder					
		Kids					
		Laws					
		Marry					
		Prayer					
		Religious Services					
		Respect					
		Vulgarity Around Rabbi					
Mormon		Baptism					
		Blessed Bread & Water					
		Blessings for Sick					
		Church Donations					
		Emergency Provisions					
		Family Religious Study					
		Foul Language					
		Honesty					
		Nutritious Food					
		Weekly Letters					
Muslim		Arabic					
		Dress					
		Groom Dowry					
		Halal					
		Homosexuality					
		Hug					
		Pray					
		Serving Food					
		Sex					
		Shoes					
Hindu		Basil Worship					
		Clothes					
		Coming of Age Rituals					
		Cow Worship					
		Curfew					
		Deities					
		Hair Wash					
		Home Worship					
		Right Hand					
		Temple Donations					

FIGURE 8. Norms from Study 3a, organized by religion. Boxes are colored if participants treated a norm as moral (i.e., the norm appeared in the moral cluster). Mormon and Muslim participants in our sample tend to moralize their own norms, whereas the Jewish participants did not. The Hindu participants in our sample seem to not make a moral/non-moral distinction in the sense that the other groups do. See Table S16 for individual item analyses for each norm.

individuals (such as the Mormon and Muslim participants) also moralized norms from their own religion that are not related to these themes. Meanwhile, others (such as the Jewish participants) acknowledged the special status of their own norms but did not moralize them. Yet others (such as the Hindu participants in our sample) seemed to make no distinction between the moral and the non-moral in the way that the other

groups did.[5] Our dataset therefore suggests that any theory about the lay conception of the scope of morality needs to explain why the Jewish participants in our dataset do not consider their own norms to be moral norms and why Mormons and Muslim participants do. To the extent that SDT and MFT make any predictions about how lay people decide whether a norm is moral, they too must find a way to explain these datasets.

Some proponents of SDT and MFT could deny that their theory makes any prediction about how lay people judge whether a norm is moral. However, if both theories really fail to address how the folk categorize norms as moral or nonmoral, it is to their detriment. MFT is supposed to be a theory of "morality": It is after all called the *Moral* Foundations Theory (emphasis added); its main instrument is called the *Moral* Foundations Questionnaire (Graham et al., 2009; emphasis added); and Graham and colleagues' classic article is entitled "Liberals and Conservatives Rely on Different Sets of *Moral* Foundations" (emphasis added). We are also told that "moral foundations theory . . . was originally developed to describe moral differences across cultures," including "moral differences across the political spectrum within the United States" (Graham et al., 2009, p. 1029). Perhaps Graham and colleagues use "morality" in a technical sense, which has little to do with what lay people mean by "morality" and with how lay people distinguish moral from nonmoral norms, but they haven't told us what this technical sense is, and it is not clear why we should care about morality so understood. And if "morality" is not to be understood in a technical sense, which norms, issues, and so on qualify as moral must depend on how lay people distinguish what is moral and what is not. Further, any attempt to describe *moral* differences across cultures or political affiliations must then aim at describing differences in what people across cultures or the political spectrum recognize as a moral issue. We thus find it hard to see how MFT can avoid being committed to specific predictions about what norms lay people will count as moral and about what demographic variables (e.g., religiosity or political affiliation) moderate how lay people make such judgments.

SDT could also be understood as avoiding any prediction about how the folk categorize norms as moral or nonmoral. Instead, SDT could be taken to provide a definition of morality for scientific purposes, which, its proponents claim, identifies an independent, theoretically interesting phenomenon that need not map onto folk judgments about which types of norms count as moral. Of course, scientific concepts often diverge from folk concepts for good reason. Folk concepts are sometimes deficient in some respect or other, and we don't need to bring their deficiencies into our scientific theories. SDT could claim that the folk concept of morality is deficient in some respect or other, and that instead of bringing the deficiencies of the folk concept into our scientific theory of morality SDT picks out an independent kind that is not captured by the folk concept. However, a significant body of research (reviewed in Stich, 2018) has shown that SDT's definition of morality does not pick out a coherent phenomenon at all. And if so, it is unclear what makes this phenomenon important from a scientific perspective. Thus, despite the fact that SDT may have operated with a stipulative, theoretical

---

5. Of course, it is entirely possible that our Hindu subjects distinguish other categories of norms that we did not test (such as prudential vs. religious norms, for instance). It is also possible that they draw a distinction similar in nature to one of the other religious groups, but that this distinction maps onto a word other than "moral."

definition up to this point, we think their theory could be improved by moving to a view of the moral domain that attempts to capture the folk concept of "moral."

Here, we offer a sketch of how a new theory might be developed that connects religious affiliation with the conception of the moral domain, given our findings. What characteristics of the religious traditions examined in Study 3 could explain the different conceptions of morality that we uncovered in our participants? Several ready explanations are not supported by the data. The findings cannot be explained by simple differences in political affiliation: The Jewish participants did not differ significantly in political attitudes from secular or Muslim participants, yet their conception of morality is identical to that of secular individuals and highly different from that of Muslims. Muslims and Mormons differed significantly on political attitudes, yet their conceptions of the moral domain are similar. (For statistical details and analysis of political attitudes, see Supplemental Materials, Study 3a: Supplemental Results Section and Table S10.) Levels of religiosity cannot explain the findings either: The Jewish participants in Studies 3a and 3b differed significantly on religion-specific religiosity scales (see Supplemental Materials, Study 3a: Supplemental Results Section and Tables S12 and S19), yet they conceived of the moral domain identically (see Figures 4 and S5). Finally, the data cannot be explained by a difference between Abrahamic and non-Abrahamic religions.

The pattern of findings may be partially explained by the difference between proselytizing religions (such as Mormonism and Islam) and non-proselytizing religions (such as Judaism). Proselytizing religions (which showed an expanded moral domain in our sample) seek to increase the size of the in-group, and they aim to do this by encouraging others to adopt their religious beliefs and practices (Arnold, 1913; Shepherd & Shepherd, 1996; Thiessen, 2011). Some Hindu subsects do proselytize (Hare Krishnas and Arya Samaji, for instance) while others do not. This could explain why no difference between moral and non-moral norms emerged from Hindu participants' answers. Whether the proselytizing feature results from an expanded moral domain or whether this feature is a partial cause of expanded moral domains remains an open question for future research.

One important limitation of our study is that religion is often confounded with other social and cultural factors, and those factors may have had a role in shaping our participants' conceptions of the moral domain, independent of religion. Above, we already mentioned that a few social/cultural factors cannot explain our findings (political affiliation and level of religiosity). However, the religious groups likely differed in their ethnicities, immigration histories, family structures, levels and sources of education, sources of news and entertainment, and so on. Moreover, while all participants were highly competent English speakers, they differed on whether English was their native language; in Study 3a, over 90% of the Jewish, Mormon and secular individuals spoke native English while only 66% of Muslims and 36% of Hindus did. (Similar percentages characterize the Study 3b population.) We conclude, therefore, that we find a correlation between religious affiliation and the conception of the moral domain and avoid making strong causal claims. Future research should explore whether it is religious beliefs or practices that create this effect, or whether it is cultural factors that are correlated with a certain religious life that do the work, or (most likely) whether both of these features make important contributions to the conception of the moral domain.

The Moral Questionnaire is just one way of investigating the question we have posed: whether religions impact what counts as a moral issue for their adherents. Future work

should include other methods and metrics (the role of intention, willingness to punish, etc.) as well as investigate what is common among the norms that a given religion thinks of as moral. We think that the fundamental difference of what counts as a *moral* issue could be at the core of why it so often seems that religious dialog involves the parties talking past each other.

REFERENCES

Arnold, T. W. (1913). *The preaching of Islam: A history of the propagation of the Muslim faith.* New York: Constable.

Berniūnas, R. (2014). Delineating the moral domain in moral psychology. *Problemos, Supplement*, 90–101.

Buchtel, E. E., Guan, Y., Peng, Q., Su, Y., Sang, B., Chen, S. X., & Bond, M. H. (2015). Immorality East and West: Are immoral behaviors especially harmful, or especially uncivilized? *Personality and Social Psychology Bulletin, 41*(10), 1382–1394.

Church, T., Katigbak, M. S., Reyes, J. A. S., & Jensen, S. M. (1998). Language and organization of Filipino emotion concepts: Comparing emotion concepts and dimensions across cultures. *Cognition & Emotion*, *12*(1), 63–92.

Cohen, A. B., Malka, A., Rozin, P., & Cherfas, L. (2006). Religion and unforgivable offenses. *Journal of Personality, 74*(1), 85–118.

Cohen, A. B., & Rozin, P. (2001). Religion and the morality of mentality. *Journal of Personality and Social Psychology*, *81*(4), 697–710.

Dranseika V., Berniunas R., & Silius, V. (2018). Immorality and *Bu Daode*, unculturedness and *Bu Wenming*. *Journal of Cultural Cognitive Science, 2*(1–2), 71–84.

Everitt, B. S., Gourlay, A. J., & Kendell, R. E. (1971). An attempt at validation of traditional psychiatric syndromes by cluster analysis. *British Journal of Psychiatry*, *119*, 399–412.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. Chichester: Wiley.

Fessler, D. M. T., Barrett, H. C., Kanovsky, M., Stich, S., Holbrook, C., Henrich, J., Bolyanatz, A. H., Gervais, M. M., Gurven, M., Kushnick, G., Pisor, A. C., von Rueden, C., & Laurence, S. (2015) Moral parochialism and contextual contingency across seven societies. *Proceedings of the Royal Society B: Biological Sciences, 282*(1813), 20150907.

Flood, G. (2008). *The Blackwell companion to Hinduism*. Oxford: Wiley.

Fontaine, J. R., Poortinga, Y. H., Setiadi, B., & Markam, S. S. (2002). Cognitive structure of emotion terms in Indonesia and The Netherlands. *Cognition & Emotion, 16*(1), 61–86.

Graham, J., & Haidt, J. (2010). Beyond beliefs: Religions bind individuals into moral communities. *Personality and Social Psychology Review*, *14*(1), 140–150.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47*, 55–130.

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029.

Haidt, J., & Graham, J. (2007). When morality opposes justice: Conservatives have moral intuitions that liberals may not recognize. *Social Justice Research, 20*, 98–116

Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus, 133*(4), 55–66.

Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology, 65*, 613–628.

Huebner, B., Lee, J. J., & Hauser, M. D. (2010). The moral-conventional distinction in mature moral competence. *Journal of Cognition and Culture*, *10*(1), 1–26.

Human Relations Area Files Database. Available online at http://ehrafworldcultures .yale.edu/ehrafe/

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). Harm, affect, and the moral/conventional distinction. *Mind & Language*, 22(2), 117–131.

Kendell, R., & Jablensky, A. (2003). Distinguishing between the validity and utility of psychiatric diagnoses. *American journal of psychiatry, 160*(1), 4-12.

Killen, M., & Smetana, J. (Eds.). (2005). *Handbook of moral development*. Mahwah, NJ: Psychology Press.

Machery, E. (2012). Delineating the moral domain. *Baltic International Yearbook of Cognition, Logic and Communication*, 7(1), 6.

Machery, E. (2018). Morality: A historical invention. In J. Graham and K. Gray (Eds.), *Atlas of moral psychology* (pp. 259–265). New York: Guilford.

Machery, E., & Stich, S. (2013). You can't have it both ways: What is the relation between morality and fairness? *Behavioral and Brain Sciences*, 36(1), 95.

McKay, R., & Whitehouse, H. (2015). Religion and morality. *Psychological Bulletin*, 141(2), 447.

McKenzie, J., & Jensen, L. A. (2017). Charting the moral life courses: A theory of moral development in US evangelical and mainline Protestant cultures. *Culture & Psychology*, 23(4), 433–460.

Miller, J. G. (2001). Culture and moral development. In D. Matsumoto (Ed.), *The handbook of culture and psychology* (pp. 151–169). New York: Oxford University Press.

Murdock, G. P. (1981). *Atlas of world cultures*. Pittsburgh: University of Pittsburgh Press.

Norenzayan, A. (2014). Does religion make people moral? *Behaviour, 151*(2–3), 365–384.

Nucci, L. (1981). Conceptions of personal issues: A domain distinct from moral or societal concepts. *Child Development*, 52, 114–121.

Nucci, L. P. (1985). Children's conceptions of morality, societal convention, and religious prescription. *Moral dilemmas: Philosophical and psychological issues in the development of moral reasoning*, 137-174.

Nucci, L., & Turiel, E. (1993). God's word, religious rules, and their relation to Christian and Jewish children's concepts of morality. *Child Development, 64*(5), 1475–1491.

Obeid, N., Argo, N., & Ginges, J. (2017). How moral perceptions influence intergroup tolerance: Evidence from Lebanon, Morocco, and the United States. *Personality and Social Psychology Bulletin*, 43(3), 381–391.

Paykel, E. S. (1971). Classification of depressed patients: A cluster analysis derived grouping. *British Journal of Psychiatry*, 118, 275–288.

Prinz, J. (2006). The emotional basis of moral judgments. *Philosophical Explorations, 9*(1), 29–43.

Schwartz, S. H., & Huismans, S. (1995). Value priorities and religiosity in four Western religions. *Social Psychology Quarterly, 58*, 88–107.

Shepherd, G., & Shepherd, G. (1996). Membership growth, church activity, and missionary recruitment. *Dialogue: A Journal of Mormon Thought, 29*(1), 33–57.

Shweder, R. A., Mahapatra, M., & Miller, J. G. (1987). Culture and moral development. In J. Kagan & S. Lamb (Eds.), *The emergence of morality in young children* (pp. 1–83). Chicago: University of Chicago Press.

Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass, 4*(4), 267–281.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more? *Journal of Personality and Social Psychology, 8*(8), 95–917.

Smetana, J. (1981). Preschool children's conceptions of moral and social rules. *Child Development, 52*, 1333–1336.

Sneath, P. H., & Sokal, R. R. (1973). *Numerical taxonomy. The principles and practice of numerical classification*. San Francisco: Freeman.

Sousa, P., Holbrook, C., & Piazza, J. (2009). The morality of harm. *Cognition*, 113(1), 80–92.

Sousa, P., & Piazza, J. (2014). Harmful transgressions qua moral transgressions: A deflationary view. *Thinking & Reasoning, 20*(1), 99-128.

Srinivasan, M., Kaplan, E., & Dahl, A. (2018). Reasoning about the scope of religious norms: Evidence from Hindu and Muslim children in India. *Child Development, 90*(6), e783–e802.

Stich, S. (2017). The moral domain. In K. Gray and J. Graham (Eds.), *Atlas of moral psychology* (pp. 547–555). New York: Guilford.

Stich, S. (2018). The quest for the boundaries of morality. In K. Jones, M. Timmons and A. Zimmerman (Eds.), *The Routledge handbook of moral epistemology* (pp. 15–37). New York: Routledge.

Stich, S, Fessler, D., & Kelly, D. (2009). On the morality of harm: A response to Sousa, Holbrook and Piazza. *Cognition, 113*, 93–97.

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology, 70*(6), 1256.

Thiessen, E. J. (2011). *The ethics of evangelism: A philosophical defense of proselytizing and persuasion.* Downers Grove: InterVarsity Press.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.* Cambridge: Cambridge University Press.

Vainio, A. (2011). Religious conviction, morality and social convention among Finnish adolescents. *Journal of Moral Education, 40*(1), 73–87.

Wallace, G., & A. Walker (Eds.). (1970). *The definition of morality.* London: Methuen.