The Effect of Religion on Muslims' Charitable Contributions to Members of a Non-Muslim Majority[1]

Rebecca B. Morton[2]   Kai Ou[3]   Xiangdong Qin[4]

July 31, 2018

[2]Professor of Politics, NYU NYC and NYU Abu Dhabi, and Director of Social Sciences Experimental Laboratory, NYU Abu Dhabi. Mailing address: Dept. of Politics, NYU, 19 West 4th Street, 2nd Floor, NYNY 10012. Email: rebecca.morton@nyu.edu.

[3]Assistant Professor, Department of Political Science, and faculty of the xs/fs Experimental Social Science Research Group at Florida State University, 551 Bellamy Hall, Tallahassee, FL 32306-2230. kou@fsu.edu.

[4]Professor of Economics and Director of the Smith Experimental Economics Research Center, Shanghai Jiao Tong University, 1954 Hua Shan Road, Shanghai 200030, P.R. China, xdqin@sjtu.edu.cn.

Abstract

We investigate the effects of religion on charitable contributions of Muslims who are in a minority to non-Muslims who are in a majority and to fellow Muslims. We find that religious thinking leads to significantly more charitable giving by 10%. The effect of religious thinking is dependent on the ethnic identity of the recipient: we find a significant effect on giving behavior towards relatively more privileged outgroup members (Han Chinese), but a small and generally insignificant effect towards ingroup members (fellow Muslims). With religious thinking, prosocial behavior towards outgroup members is significantly higher by 14%, which is mainly explained by the religiosity of Muslims. Our results have implications for our understanding of the influence of Islamic rules on Muslims' attitudes and behavior towards non-Muslims and for the design of fundraising mechanisms in Muslim communities.

1 Introduction

What motivates individuals to voluntarily contribute to charitable activities when the recipient is a member of an ethnic or racial group which is generally perceived as more privileged than their own? Considerable experimental research has examined motivations behind charitable contributions generally. To our knowledge there has been little experimental study of this type of giving behavior. The warm glow hypothesis provides a direction for research in which donations and giving enter directly into utility functions (e.g., Arrow, 1972; Andreoni, 1989, 1990). This hypothesis has been tested and justified in experimental studies (e.g., Andreoni, 1995; Crumpler and Grossman, 2008; Harbaugh, Mayr and Burghart, 2007; Null, 2011; Andreoni, Rao and Trachtman, 2017), and it is widely used in the design of fundraising strategies. In studies on promoting charitable giving and contributions to public goods from which individuals derive no direct consumption benefits, the effective strategies include the public display of identity (e.g., Andreoni and Petrie, 2004; Rege and Telle, 2004; Alpizar, Carlsson and Johansson-Stenman, 2008) and emphasizing the shared identity of receivers. These strategies are believed to increase the donors' utility value from giving. Yet, what if individuals do not share identities and in fact the recipients are viewed to have an identity which gives them greater status or power in the larger society?

Fundraisers and charities also appeal to religion, as religion functions as a moral compass and has been found to promote contributions to public goods (e.g., Benjamin, Choi and Fisher, 2016), and increase donations (e.g., Lambarraa and Riener, 2015). Religiosity and religious precepts can motivate individuals who have altruistic tendencies (both pure and impure) to give and donate because it contains messages which often emphasize others' benefits from one's behavior (e.g., Warr, 1982; Roberts, 1984; Bergstrom, Blume and Varian, 1986; Bernheim, 1986). Nevertheless, there has been to our knowledge no study of the extent that these religious messages can promote giving to those who are nonbelievers and even members of a group which is more privileged and in some ways limit the religious practices of the believers.

This paper reports the results of a "lab-in-the-field" experiment designed to isolate

and measure the mechanism of how Islamic rules influence Muslims' giving behavior to non-Muslims who are part of a majority privileged group. We conducted our experiment with non-western subjects (Muslims in China) in contrast to most previous studies which examined mainly western students. We designed a novel experiment that is based on a collective giving game. That is, a group of Muslim subjects vote to decide whether the group collectively gives an amount of money to a verified needy receiver. We manipulated the treatment by varying the amount of money to be transferred, the identity of the receiver (whether Muslim Chinese who is a member of the minority or Han Chinese who is a member of the majority and privileged group), whether subjects engaged in a religious thinking exercise prior to voting, and the privacy of the decision-making environment.

We find that in the absence of religious thinking exercise, the frequency of Muslims' giving behavior when the receiver is an in-group member (fellow Muslim) is not statistically distinguishable from Muslims' giving behavior when the receiver is an outgroup member (Han Chinese). Compared to the baseline, we find that religious thinking leads to significantly more charitable giving by 10%. The effect of religious thinking is dependent on the ethnic identity of the recipient: we find a significant effect on giving behavior towards relatively more privileged outgroup members (Han Chinese), but a small and generally insignificant effect towards in-group members (fellow Muslims). With religious thinking, prosocial behavior towards outgroup members is significantly higher by 14%, which is mainly explained by the religiosity of Muslims.

Our study has strong empirical and policy implications in several ways. First, our study provides new evidence on the extent that the religiosity of Muslims treats Muslim and non-Muslims/non-believers equally, which implies that religious thinking alone may not be a force towards violence towards non-believers. Second, our study improves our understanding of the influence of Islamic rules on Muslims' giving behavior that has implications for the design of fundraising mechanisms in Islamic countries and Muslim communities. Our results imply that methods that are effective to Westerners on promoting giving and donations may result in negative outcomes when applied to Muslims. In future research on fundraising campaigns in Muslim communities scholars and charities

2

need to consider the fundamental effects of religion and the effects of religiosity that we isolate in this study.

In the next section, we describe our experimental procedures and the design of treatments. In Section 3 we analyze the experimental results and in Section 4 we discuss the implications of the results for the effects of religious thinking and understanding charitable giving behavior towards others (in-and outgroup members).

2   Experimental Design

2.1   Location and Subjects

Our experiment was conducted in Yinchuan, the capital city of China's Ningxia Hui (Muslim) Autonomous Region using a total of 480 student subjects who were randomly recruited from different departments and colleges at a representative public university in the area. The Hui Chinese are Muslims and they practice Islam. The population is relatively homogeneous in terms of their social norms and moral standards, living in an autonomous region in a relatively closed community. However, they still clearly face competition from Han Chinese who also live in the region and are the dominant ethnic group in the region (exact figures of the Hui population are not easily obtained, but most estimates are that they are a third of the population).[1] It is essential to note that Hui Chinese and Han Chinese look distinctly different, and the cultural and religious differences between Hui and Han lead to different patterns in their dress, consumption of food and drink, and other dimensions that may cause conflicts. In the local university where we recruited subjects and conducted the experiment, the majority is Han Chinese, although this university has significantly more Muslim students than most other Chinese universities. Thus, we can clearly identify in- and out-group individuals and which group is arguably in the majority and more privileged.

In our experiment, 43% of the subjects were male and 57% female. We balanced gender in the process of our recruitment and implementation of the experiment, although subjects were randomly assigned to treatments and sessions. We used an improvised "lab-

---

[1]See Encyclopedia Britannica, http://www.britannica.com/place/Ningxia.

in-the-field" methodology; the experiment was conducted in a standard classroom. All the subjects who made decisions in the experiment were Muslims.

We investigate the effects of religious thinking on charitable contributions towards Han Chinese as compared to Muslims in our experiment. The receivers were chosen from a database of the "most needy students" verified by the university. The family background of these students has been investigated by the university. The "most needy students" have received tuition waivers and living allowances. Moreover, almost every student at the local university understands the program of aiding the "most needy students," so we could easily explain the experimental task when we introduced the role of the receiver and who they were during the experimental sessions. Students in the "most needy students" database were not allowed to participate in the experiment as decision-makers.[2] Only the decision-makers came to the experimental sessions; the receivers were not present. Decision-makers were allowed to participate in only one session.

Instructions were read aloud by the same experimenter in all sessions. The subjects were allowed to ask questions privately and to make sure that everyone had common knowledge of the decision tasks. On average, one session lasted for 90 minutes, and average earnings were about 49 RMB (1 US \$ = 6.2 RMB at the time of the experiment). Subjects also received an extra 10 RMB for showing up on time.

2.2   Preference Elicitation

We measured charitable contributions using a collective decision making task. That is, the Muslim subjects voted to decide whether a group that consisted of three voters collectively gave an amount of money to the receiver. The receiver had no other option than to take whatever the group allocated to him or her. We recruited 360 Muslim subjects as voters and an additional 120 subjects to serve as receivers. Half of the 60 receivers were Muslim and half were Han Chinese (non-Muslim). The receivers did not make decisions in the

---

[2] The receivers were asked first as to their willingness to participate in our experimental study. Once a receiver agreed to participate in the study, his or her information was recorded and (s)he was randomly matched with a 3-person team in one session as described below. After the experiment was concluded, the experimenters and the local coordinator gave the money that had been transferred by the corresponding team to the receiver.

Table 1: Payoffs in the Experiment

	A wins	B wins		
		$T = 3$	$T = 5$	$T = 7$
Each Voter	20	17	15	13
The Receiver	0	9	15	21

experiment.

Table 1 below presents the payoffs from the two options for the voters and recipients as a function of the transfer amount (hereafter, T). Voters did not receive any feedback during the session.[3] The collective giving game in each period proceeded as follows:

1. Each voter was endowed with 20 RMB for the period.

2. Each voter voted as to whether to transfer an amount $T \in \{3, 5, 7\}$ of his or her endowment to the receiver assigned to their group. Voters chose between two options: A or B. Option A meant not transferring the money, while Option B meant transferring the money to the recipient. Abstention was not allowed.

3. The result of the voting was determined by simple majority rule. That is, if two out of three voters chose Option A, Option A was imposed on the group and vice-versa. Each voter kept his or her 20 RMB and the recipient received nothing. If Option B was chosen, then each voter paid T from his or her endowment and the recipient received $3T$. Table 1 below presents the payoffs from the two options for the voters and receivers as a function of T.

Why did we choose this voting procedure and investigate prosocial behavior in a group decision-making setting? A multiple-dictator setting can be found in the literature (e.g., Dana, Weber and Kuang, 2007). The procedure of our experiment is similar to the team dictator game in Cason and Mui (1997) and the voting mechanism in List, Berrens, Bohara and Kerkvliet (2004). Our design has a number of advantages. First, it

---

[3]Note that the sessions in which we used double-blind privacy, as explained below, were conducted using paper and pencil and providing feedback during a session was difficult. As abstention was not allowed in the voting game, it is straightforward that voting nonsincerely is weakly dominated and, based on previous voting experiments, we expected that subjects did not require the repeated feedback and learning that may be necessary for equilibrium behavior to emerge in more complicated games.

reduced the number of receivers required for the study and thus increased the number of observations for a given subject pool.[4] Second, there is some evidence from Cason and Mui (1997) that there is greater prosocial behavior in the game, providing us a challenging environment to find significant effects of religious thinking. Theoretically it will be more difficult to identify a significant marginal effect caused by the religious thinking. We acknowledge that an individual may make giving decisions differently in a team-dictator game than the canonical 2-person dictator game and in some situations the group setting may decrease generous decisions (e.g., Luhan, Kocher and Sutter, 2009). Yet, in the voting game arguably people may be more prosocial since the probability that their choice affects the outcome is less and thus the subject can at lower cost express prosocial behavior. For example, Feddersen, Gailmard and Sandroni (2009) find that when the probability of being pivotal in voting declines, people are more prosocial in their voting choices.

2.3 Experimental Procedure

In each session, there were 18 voters and 6 receivers. The voters were randomly divided into three Sets of 6 participants. Within each Set of 6, the participants were randomly assigned to two Groups of 3 voters each. The 6 voters in each Set were fixed during a session, but in each period, the two groups were randomly re-drawn from the group of 6 participants within the Set.

How did a subject know the identity of receivers? In our experiment, subjects were provided with a printed sheet with basic demographic information about the six receivers assigned to their session. The printed sheet described the six receivers' gender, whether the recipients were Hui Muslim or Han Chinese, and where each was born, but it did not reveal the receivers' names. Moreover, in each session, there were always three male and female receivers on the list and all six were in the same ethnic group. The possible influence of gender of the receiver (Dufwenberg and Muren, 2006; Croson and Gneezy, 2009) was

---

[4]We faced significant logistic constraints when we recruited subjects. Although our subjects were recruited from a public university in which there were significantly more Muslim students than most other Chinese universities, only a small minority are needy students and thus eligible to be recipients and we were only able to recruit and manage a somewhat small subject pool when we conducted the experiment.

therefore controlled through our design. In addition, we required that the receivers agree to keep their identities confidential. Given that most preferred to be anonymous to the voters and that the receivers did not know which group they were assigned, it is unlikely that voters could identify the receiver to a specific person when they made decisions in the experiment. Subjects knew that each receiver participated in only one session. Each three-voter group was randomly matched up with a specific receiver prior to the beginning of a session but voters did not know the names of the six receivers who were assigned to their group.

As noted above, the principal experiment was conducted in a *double-blind* setting in which the subjects' decisions were anonymous to the experimenters as we explain below, which is different from the *single-blind* setting in which subjects' decisions are anonymous to other subjects but closely monitored/recorded by the computers and experimenters. Our procedure was designed to convince subjects that they were making decisions in a purely anonymous environment. Although the difference between double-blind and single-blind may seem subtle, for the subjects the experience was vastly different.

In the double-blind setting, upon arrival, two subjects were randomly selected in a public manner to serve as monitors and they were asked to wait in another room.[5] Monitors were used to ensure credibility and calculate subjects' payoffs at the end of a session. During the session, the monitors waited in another room that was close to the laboratory in which we conducted sessions. The monitors could hear the experiment but did not see the subjects or observe their choices. The subjects were assigned as voters and they randomly chose ID number cards.[6] The experimenters had marked the same number on two pieces of paper and had stapled them together in advance. Therefore, every voter received two ID number cards.

The voters were given a copy of the written instructions and 30 large sealed envelopes. Each of the large sealed envelopes had a number written on the front for each experimental

---

[5]We recruited 20 subjects for each session. 20 cards were placed publicly into a large envelope, and subjects saw that two cards were marked "Monitor" and the other 18 cards were marked "Subject." Subjects were asked to randomly draw their roles publicly. We used the two monitors to facilitate the calculation of payments at the end of the session.

[6]The ID numbers were randomly marked in symbols and letters such as "S+L" to prevent any feeling of superiority based on the number drawn.

period. The voters were asked to open the sealed envelope labeled number 1 in the first period, and the envelope labeled number 2 in the second period, and so forth, for 30 periods. We paid subjects for three randomly selected periods, one for each value of T.

Subjects were arranged to sit individually so that there was enough distance between them to guarantee the privacy of the subjects' decisions. In the designated large envelope for each period, there was one white standard letter-sized envelope and two ballot tickets. The ballot tickets were marked "Vote for Option A," and "Vote for Option B." The voters' ID numbers were not marked on the tickets, but their group numbers were marked on the tickets. If voting for Option A, the voters put "Vote for Option A" in the standard letter-sized envelope, "Vote for Option B" in the large envelope; if voting for Option B, the voters put "Vote for Option B" in the standard letter-sized envelope, "Vote for Option A" in the large envelope. After the voters made their choices, the experimenters collected the standard letter-sized envelopes into an opaque ballot box. The experimenters did not collect ballots until the voters had all completed choices, and they did not scrutinize the collected envelopes.

Only the votes in the standard letter-sized envelopes were used to determine the winning choices using simple majority rule discussed above. At the end of the session, we revealed the period numbers that were ex-ante written on board and calculated the payoffs for the voters in those periods.[7] Since we conducted the experiment in 30 periods, we changed T every fifth period using either Sequence 1: $\{3, 5, 7, 7, 5, 3\}$ or Sequence 2: $\{7, 5, 3, 3, 5, 7\}$.

Once we collected the standard letter-sized envelopes, we asked the voters for one of their two ID number cards and gave these materials to the monitor who then calculated the payoffs without knowing the identities of the voters. The monitors calculated the payoffs by opening the submitted standard letter-sized envelopes for the selected periods. Specifically, the monitors could see on the ballot tickets each voter's group number in

---

[7]We chose the period in advance for three reasons: First, given the prohibition against gambling in Islam, we did not want to use traditional techniques of randomly selecting periods to be paid. Second, we wanted to show subjects that everything in the experiment was transparent and the *ex-ante* selected periods were not modified by the experimenters based on the results of the experiment. Third, we wanted to make sure that subjects were incentivized for each value of T.

that period and his or her decision and could calculate the payoffs based on the winners in that group. In some treatments, the monitors also calculated the voters' earnings from answering a set of incentivized questions. The monitors received a flat payment equal to the maximum that could be earned in the experiment if they calculated and paid all the subjects correctly. That is, the monitors informed the experimenters the results of those selected periods - the chosen option of each 3-person group. The subjects then calculated their payoffs themselves and check whether or not they received the correct payment. The experimenters also calculated the total payoffs and gave the appropriate money to the monitors. If any subject had reported a wrong payment, the monitors forfeited their payments. However, no mistakes were reported.

Next, the monitors put the voters' payoffs in new white envelopes, sealed them and gave them to the experimenters. The voters' ID number cards were taped to the front of each standard letter-sized envelope so that the experimenters could check the second ID card in each voters' hand and give the sealed payoffs to him or her accordingly. The experimenters did not know how any particular voter's choices nor how much (s)he earned. The voters were asked to take their ID number cards with them when they left the laboratory, to ensure that the experimenters could not retrospectively match up subjects with their decisions and payoffs.

We conducted the decision task repeatedly without providing subjects feedback at the end of each round. The reason for this design is twofold. First, because of our experiment was conducted using paper and pencils, it was extremely complicated and time-consuming to provide feedback to subjects in the middle of the experiment. Second, we wanted to investigate whether giving behavior may decay over time as documented in other studies (e.g., Chuan, Kessler and Milkman, 2018). Giving is often a spontaneous act of generosity, but as time progresses, an individual may rely less on the unconscious and fast-thinking and rely more on sophisticated and analytical thinking. Note that in our design any decrease in giving in time, however, would not be due to learning from feedback of others' behavior, but from a decay in the initial effect of religious thinking. A dynamic process of the evolution of giving over time can help us explore this question.

3 Experimental Results

3.1 Results of the Baseline

To establish an appropriate counterfactual, we study Muslim subjects' charitable contributions to in-group and outgroup members in a baseline setting. Table 2 reports aggregate percentage votes for giving by recipient identity and value of T. In the statistical analysis that follows we use the mean of the aggregate vote shares by Set. As noted above, we designed a double-blind decision-making environment such that the experimenters could not match subjects' IDs and their choices. Hence, by using the Set Level averages in the analysis we have strictly independent observations while also controlling for repeated play in a session. Because the Set Level averages are numbers between 0 and 1 rather than dichotomous results, we conduct Wilcoxon-Mann-Whitney tests to investigate the difference of the Set Level means of voting for option B towards Hui and Han receivers.

Table 2: Individual Level Voting Averages for Giving by Treatment

Treatment	Total Obs.	Set Level Obs.	Average Votes for Giving			
			Total	T=3	T=5	T=7
Baseline/Pooled	2160	12	0.61	0.70	0.64	0.48
			(0.04)	(0.03)	(0.05)	(0.05)
Baseline/Han	1080	6	0.62	0.72	0.63	0.50
			(0.05)	(0.04)	(0.07)	(0.06)
Baseline/Muslim	1080	6	0.59	0.67	0.64	0.47
			(0.07)	(0.06)	(0.07)	(0.08)
Religious/Pooled	2160	12	0.70	0.80	0.71	0.60
			(0.02)	(0.02)	(0.02)	(0.04)
Religious/Han	1080	6	0.75	0.84	0.76	0.68
			(0.02)	(0.03)	(0.02)	(0.03)
Religious/Muslim	1080	6	0.65	0.77	0.67	0.53
			(0.03)	(0.03)	(0.04)	(0.05)

Note: Standard errors are clustered at the Set level and reported in parentheses. Each cluster consists of 180 observations.

We find a number of interesting results. First, compared to the share of voting for option B when the recipient is Hui Chinese, we do not observe a statistically significant difference in giving behavior when the receiver is Han Chinese. The average percentage of voting for giving is 62% when the receiver is Han and 59% when the receiver is Hui, which is not statistically significant ($P = 0.87$, $N = 12$). As in a number of religions,

Islam teaches that Muslims should treat non-Muslims with respect and that all humans are created from the same dirt.[8] In our experiment, we find no evidence of discrimination on giving between Muslims and non-Muslims. Second, we find that subjects are more likely to give when the amount of giving is relatively smaller. This suggests that when subjects make giving decisions, they may consider the cost of giving and fairness. In our design, we varied the transfer amount to investigate voters' fairness and cost concerns. That is, when $T = 5$ and option B is selected, both the voter and the receiver have the same earning. But when $T = 3$ or $T = 7$, there is inequality of earning between the voter and the receiver. Table 2 shows a clear ordinal relationship between the transfer amount and the averages of voting for option B. We perform Kruskal-Wallis statistical tests and find that the ordinal relationship is significant ($P = 0.01$, $N = 36$), which implies that as the transfer amount increases individuals are less likely to vote for Option B.

3.2 Religious Thinking and Effects

In this section, we investigate the extent to which religious thinking affects individual behavior in our experiment. In our Religious Thinking Treatment, subjects were asked to answer an incentivized quiz before engaging in the collective giving game. The quiz consisted of questions concerning common Islamic beliefs as explained in the Supplemental Appendix. These questions are found in one of the most influential and popular Chinese Muslim online communities. This source also provides training in learning Arabic and the Quran and spreading Chinese Islamic leaders' writings. We asked 20 questions, and subjects had 10 minutes to answer these questions. We designed questions which were straightforward, that is, if subjects knew the answers they should have no difficulty answering the questions. The effort wise cost is minimized. Subjects earned 50 cents local currency (RMB) for each correctly answered question. The average number of questions answered correctly was 9.34. None of these questions concerned inducements to moral behavior. Example questions are given below.

Example 1 *Is lottery allowed in Islam?*

---

[8]See, for example, the numerous examples from the Prophet's life in Tariq Ramadhan's *Footsteps of the Prophet.*

A.   Yes

B.   No

Example 2 *According to the Islamic calendar, when should Muslims fast? ____ [Fill-in-the-blank]*

Our religious thinking exercise functioned as a nudge that may cause one to think about religious issues. By performing a light but incentivized task, subjects may pay more attention to the religious information. Importantly, our method was indirect, which is closer to naturally occurring situations. That is, in daily life in Ningxia where we conducted the experiment, it is rare for individuals to be faced with strong religious messages. To control for possible demand effects, our study focuses more on the results of religious messages provided in a natural and indirect way. The meta-analysis of Shariff, Willard, Andersen and Norenzayan (2016) finds little difference in effect sizes from such indirect messages and more direct blatant messages. As religious priming has been found to increase prosocial behavior in other studies, we expect the following hypothesis.

Hypothesis 1 *Compared to the Baseline, subjects are more likely to vote for option B in the Religious Thinking treatment.*

Result 1 *Overall, the religious thinking exercise leads to significantly more giving behavior by approximately 10%. In particular, when the receiver is Han Chinese, religious thinking significantly increases the average of giving by approximately 14%. But when the receiver is Muslim, the increase is not significant.*

Support We conduct Wilcoxon-Mann-Whitney tests to investigate the difference of the Set Level means of voting for option B in different treatments. Muslim subjects vote for option B about 61% of the time in the baseline sessions but 70% in the religious thinking sessions, which rejects the null hypothesis at a 3 percent level ($P = 0.03$, $N = 12$). When we break down the effect by identity, we find that the aggregate effect of religious thinking is mainly driven by the higher level of giving to Han Chinese. That is, when the receiver is Han Chinese, Muslim subjects vote for option B 62% of the time in Baseline but 75%

of the time in Religious Thinking, which rejects the null hypothesis at a 2 percent level ($P = 0.02$, $N = 12$). But when the receiver is Muslim, we find the average percentage of voting for option B is 59% in Baseline and 65% in Religious Thinking, which cannot reject the null hypothesis ($P = 0.26$, $N = 12$).

How does the religious thinking exercise affect giving to an in-group member and an out-group member? It is an empirical question. Previous studies show that social identity plays a crucial role when individuals make economic decisions. However, in our baseline sessions, we do not find evidence of discrimination in giving between in-group and out-group members. We argue that this is the evidence on the extent that religiosity of Islam treats Muslims and non-Muslims equally. If it is the religiosity of Islam that results in a statistically indistinguishable averages of voting for option B, then the religious thinking exercise should make the religion's effect stronger. Hence, we expect the following hypothesis.

Hypothesis 2 *In the religious thinking sessions, Muslim subjects are equally likely to vote for option B when the receiver is Muslim as compared to when the receiver is Han Chinese.*

Result 2 *The religious thinking exercise has a significant effect on outgroup favoritism. It increases the average of voting for giving to outgroup members by approximately 10%.*

Support In the religious thinking sessions, we find that when the receiver is Han Chinese about 75% subjects vote for option B as compared to 65% when the receiver is Muslim. It suggests that in the religious thinking sessions, the averages of voting for option B is significantly larger which rejects the null hypothesis at the 3 percent level ($P = 0.03$, $N = 12$). It is essential to notice that in the baseline where the religious thinking exercise is not assigned, there is no significant difference between the average percentage of voting for option B when the receiver is Han Chinese and the results of voting for option B when the receiver is Muslim, and compared to the baseline, the religious thinking exercise only leads to more giving behavior when the receiver is Han. The differences of Set level means in the baseline are significantly different from the differences of Set level means in

the religious thinking treatment (2% vs 10%, $P = 0.02$, $N = 12$). Thus, we can conclude that religious thinking causes our Muslim subjects to favor Han Chinese.

Hypothesis 3 *The cost of transfer amount plays a crucial role in making voting decisions. More subjects are willing to vote for option B when the cost is relatively lower.*

Result 3 *As the transfer amount increases, less Muslim subjects vote for option B.*

Support We perform Kruskal-Wallis statistical tests and investigate the ordinal relationship demonstrated in Table 2. Overall, significantly more Muslim subjects vote for option B when $T = 3$ compared to when $T = 5$ or $T = 7$ ($P < 0.001$, $N = 72$). Such an ordinal relationship is both significant in the baseline as reported above and in the religious thinking treatment ($P < 0.001$, $N = 36$). When we break the effects down by the identity of the receiver, the results are more nuanced. In the baseline, the ordinal relationship is marginally to not significant whether the receiver is Han ($P = 0.07$, $N = 18$) or Hui ($P = 0.14$, $N = 18$). But in the religious thinking treatment, the ordinal relationship is always highly significant both for Han ($P = 0.016$, $N = 18$) and Hui ($P = 0.009$, $N = 18$).

Since charitable contributions and giving is part of Muslims' daily life and there are no strategic interactions in our experiment, one might think that subjects' preferences should not change over time. However, it is also the case that the effects of our religious prime might deteriorate over time as well. This leads to the following hypothesis.

Hypothesis 4 *Subjects' preferences may decline over time.*

Result 4 *When focusing on the choice of the first period, we find qualitatively identical results as compared to the findings based on repeated play.*

Support The choice in the first period is arguably an independent observation. Based on the results of the first period, we find that in the absence of the religious thinking exercise, about 69% of subjects vote for option B when the receiver is Hui (an in-group member) but 61% when the receiver is Han (an out-group member), but the difference is statistically indistinguishable ($P = 0.46$, $N = 12$). However, in the religious thinking sessions, about

14

92% of Muslim subjects vote for option B when the receiver is Han but 75% when the receiver is Hui, which is marginally significant ($P = 0.06$, $N = 12$). When the receiver is Han, Muslim subjects' voting for option B is significantly higher with religious thinking compared to the baseline ($P = 0.008$, $N = 12$), while we do not observe a significant effect caused by religious thinking when the receiver is Hui ($P = 0.10$, $N = 12$). Taken together, based on the results of the first period these additional investigations identify that religious thinking causes Muslim subjects to treat non-Muslims better.

We now examine whether subjects' willingness to vote for option B changes over time. We estimate the probability that a Muslim subjects votes for option B and consider a Probit model in which an individual's voting decision is a function of the *Period* in a session. The dependent variable is voting for option B. The standard errors are clustered at the Set level. For each cluster, there are 180 observations.

Table 3: Probit Estimations of Voting for Option B as a Function of Period

Treatment	Coefficient	Std. Error	z	$Pr > |z|$
Baseline/Han	-0.00	0.00	-2.57	0.01
Baseline/Muslim	-0.00	0.00	-1.61	0.11
Religious/Han	-0.00	0.00	-1.35	0.18
Religious/Muslim	-0.01	0.00	-3.01	0.00

Note: The reported coefficients are marginal effects.

We find that in general Period has a negative influence on an individual's likelihood of voting for option B, since the estimate is always negative. This result coincides with the finding of Luhan, Kocher and Sutter (2009) as we find individuals tend to engage in less prosocial behavior over time. We should also notice that in Baseline/Han and Religious/Muslim the estimate of the coefficient is significant, which suggests that Muslim subjects are significantly less likely to vote for option B as time goes. We then test whether the coefficients are significantly different by treatment. We find that the coefficient of each treatment is statistically indistinguishable by treatment.

While giving is often a spontaneous act of generosity, asking subjects to repeatedly make giving decisions in our experimental environment may make them think about their choices. Hence, the negative influence of Period might be caused by learning or analytical

thinking. When subjects rely less on the fast and spontaneous thinking process but more on an analytical thinking process, it may reduce prosocial behavior. Because the negative influence is systematic across treatments and the magnitude of the coefficients are statistically undistinguishable, although the negative effect is significant in some treatments, we argue that only minor and limited differences in our main findings are related to learning or analytical effects.

3.3 The Possibility of a False Discovery

Our main findings suggest that religious thinking appears to cause discrimination in favor of outgroup members, treating them better than receivers from their own group. With the stimuli of religious thinking, out-group favoritism is significant in our experiment. These results are somewhat counter-intuitive. Before moving on to the discussion of the results, it is important to check whether it is a false discovery.

We test multiple hypotheses for 13 total comparisons. Because the probability of false significance is higher when making such multiple comparisons, we use a nonparametric procedure described by Benjamini and Hochberg (1995) to explore the robustness of our findings. Define $q^* = 0.05$ as the desired minimum false discovery rate or FDR. If we rank the comparisons by their corresponding p-values, where 1 denotes the smallest and 13 the greatest and the rank is denoted by i, Benjamini and Hochberg (1995) show that rejection of only null hypotheses such that the p-value is less than $(\frac{i}{13})q^*$ (which we label the q_{FDR}) controls for the FDR at q^* when the test statistics are independent. We report the results of statistical comparisons in Table 4.

We find that religious thinking significantly increases the averages of voting for option B. The effect is primarily in favor of outgroup receivers. Overall, more charitable contributions are observed when the cost of giving is relatively smaller. The results of the multiple hypotheses testing show that our findings are robust to a false discovery check.

3.4 Identification of the Underlying Mechanism

Understanding the underlying mechanism and boundary conditions of our remarkable findings is of crucial interest before thinking of policy implications. To further justify the

Table 4: Statistical Comparisons of Treatment Effects

Category	Comparison	Treatment	Identity	p-value	q_{FDR}
Religious Thinking	Religious>Baseline			0.032*	0.035
	Religious>Baseline		Han	0.022*	0.023
	Religious>Baseline		Hui	0.26	0.05
	Hui=Han	Baseline		0.87	0.05
	Hui=Han	Religious		0.03*	0.031
	ΔBaseline $= \Delta$Religious			0.025*	0.03
Transfer Amount	T3=T5=T7			0.000*	0.004
	T3=T5=T7	Baseline		0.01*	0.02
	T3=T5=T7	Religious		0.00*	0.01
	T3=T5=T7	Baseline	Han	0.07	0.04
	T3=T5=T7	Baseline	Hui	0.14	0.04
	T3=T5=T7	Religious	Han	0.016*	0.019
	T3=T5=T7	Religious	Hui	0.01*	0.012

Note: * $q \leq q_{FDR}$, reject null hypothesis of no difference.

causal claim that religion causes Muslims to be nicer to outgroup members, we need to identify the motivations behind the reported treatment effects.

As discussed above, we argue that the underlying mechanism in the reported effects is the religiosity of Islam. To empirically examine this argument, we conducted additional investigations and replicated the reported treatments in a single-blind setting via a connected network. The experiment was programmed in z-tree (Fischbacher, 2007). The experimental procedure is similar to the double-blind setting but in a completely different environment using a computerized network.

Specifically, each voter was seated at a desk; no form of communication was allowed during the experiment and voters were placed such that their decisions were private. Voters were assigned experiment specific ID numbers. We varied the value of T by period according to a fixed sequence such that each value of T was applied for 20 periods in 10-period parts. We varied the order in which T changed by session to avoid order effects. T remained fixed for 10-period intervals. We used three sequences (one session each): Sequence 3 $\{7, 5, 3, 7, 5, 3\}$; Sequence 4 $\{3, 5, 7, 3, 5, 7\}$; and Sequence 5 $\{7, 5, 3, 3, 5, 7\}$. Because the computerized environment largely facilitated experimental work, we were able to conduct 60 periods in the single-blind sessions, unlike the 30 periods in the double-blind

sessions.[9] To prevent possible wealth effects and to increase the payoffs in each period, one period for every 20 periods for each value of T was *ex-ante* selected as a paid period, such that subjects were paid for 3 periods. The paid periods were written on a blackboard, and covered by opaque papers. When the entire experiment was concluded, the experimenters uncovered the papers to show subjects which periods were *ex-ante* selected. We did not provide feedback between periods so that the results of the computerized treatments are comparable to the paper and pencil treatments. Table 5 below summarizes the treatments studied in the experiment.

Table 5: Summary of Treatments

Treatment	Thinking Exercise	Recipient Identities	Sessions	Sets	Voters	Recipients
Baseline/Han/DB	None	Han	2	6	36	12
Baseline/Muslim/DB	None	Muslim	2	6	36	12
Religious/Han/DB	Yes	Han	2	6	36	12
Religious/Muslim/DB	Yes	Muslim	2	6	36	12
Baseline/Han/SB	None	Han	3	9	54	18
Baseline/Muslim/SB	None	Muslim	3	9	54	18
Religious/Han/SB	Yes	Han	3	9	54	18
Religious/Muslim/SB	Yes	Muslim	3	9	54	18

Note: DB refers to the double-blind setting, and SB refers to the single-blind setting.

The computerized environment is a single-blind setting in which subjects' decisions are closely monitored/recorded by the computers and experimenters. Islam teaches specifically that the "moral" giving behavior should only be conducted when the behavior cannot gain praise or recognition from others.[10] If the religiosity of Muslims explains our findings, it should be that Muslim subjects care about the Islamic prohibition of showing off giving behavior, and they should behave differently in the double-blind setting in which neither the experimenters nor the other subjects know their identities as compared to the single-blind setting where the experimenters know who made which decisions. Indeed,

---

[9]In one session there was a power outage in the last period and the data for that period were lost. When we compare results across privacy procedures, we restrict our comparisons to the first 30 periods of computerized sessions.

[10]See the prescription in Qur'an, 2:264. It is worth noting that a similar rule/obligation exists in the Christian Bible and other religions too, but to the best of our knowledge, recent literature on charitable giving and/or prosocial behavior with non-Muslim subjects does not provide evidence that this rule has behavioral consequences.

Lambarraa and Riener (2015) in a field experiment conducted in Morocco find evidence which suggests that publicizing giving actually reduces the prosocial behavior.

While we are interested in the role of privacy so that we can identify the underlying mechanism of our main findings reported earlier, our study is different and novel in two aspects. First, Lambarraa and Riener (2015) compare Muslims' giving behavior in a public environment to the behavior in a private environment, but in our experiment participants' choices are anonymous in both the double-blind and single-blind sessions. It is essential to note that the privacy level of the two treatments is extremely close, except that the experimenter knew individuals' choices in the single-blind sessions. Yet, the differences in the environments between the two treatments makes it clear that there are different degrees in privacy. If the difference results in a significant behavioral difference, then we find further support for the influence of religious beliefs on Muslims' giving behavior. Second, our study investigates the joint effect of religious thinking and ethnic identity which are not systematically studied in Lambarraa and Riener (2015). In addition to Muslims' moral giving behavior, we are also interested in the effect of religious thinking on charitable giving to nonbelievers.

We combine our data together and conduct a Probit analysis of individual voting for Option B as a function of the various treatments separated by the ethnic identity of the recipient, which enables us to better discern the robustness of our results reported above. In Table 6, we report the results of the marginal effects of voting for Option B. We cluster the data by the 30 Sets in which voters were grouped for each recipient type.[11]

Because we conducted 30 periods in the double-blind sessions and 60 periods in the single-blind sessions, in our analysis, we focus on the comparisons of the first 30 periods.[12] Each cluster consists of 180 observations. Session 1 of the Baseline/Muslim/SB was one period short because of a power outage in the last period. Given we focus our analysis on

---

[11]We also estimated two regression equations with the Set averages as the dependent variables (30 observations for each) as well as one large Probit with all treatments and the individual data clustered by the 60 total Sets, which yielded identical results.

[12]We analyzed the results using data of all the 60 periods in the single-blind setting too. These additional results are reported in Table 6. We arrive at the same conclusion regardless how we pool the data.

Table 6: Probits of Voting for Giving (null = *Baseline/SB/Recipient*)

Treatments	Total		$T=3$		$T=5$		$T=7$	
			Marginal Effects					
Panel A: Han Receiver (Clustered by 30 Sets)								
Baseline/DB	-0.074	-0.089†	-0.025	-0.042	-0.072	-0.075	-0.125*	-0.151**
Religious/DB	0.072*	0.055*	0.107*	0.087*	0.056†	0.052	0.056	0.029
Religious/SB	-0.105†	-0.116*	-0.081†	-0.104*	-0.115*	-0.095†	-0.118†	-0.149*
First 30 Periods	No	Yes	No	Yes	No	Yes	No	Yes
Number of Obs.	8640	5400	2880	1800	2880	1800	2880	1800
Pseudo R^2	0.013	0.016	0.014	0.020	0.013	0.012	0.014	0.020
Log likelihood	-5528.24	-3402.12	-1679.58	-1020.12	-1828.98	-1127.04	-1941.09	-1201.07
Panel B: Muslim Receiver (Clustered by 30 Sets)								
Baseline/DB	-0.121†	-0.132†	-0.099	-0.082	-0.068	-0.085	-0.196*	-0.229**
Religious/DB	-0.064	-0.074	-0.007	0.014	-0.047	-0.063	-0.139*	-0.172**
Religious/SB	-0.042	-0.064	-0.009	0.006	-0.040	-0.092	-0.077	-0.105
First 30 Periods	No	Yes	No	Yes	No	Yes	No	Yes
Number of Obs.	8622	5400	2862	1800	2880	1800	2880	1800
Pseudo R^2	0.010	0.010	0.006	0.010	0.002	0.010	0.014	0.022
Log likelihood	-5380.16	-3408.93	-1564.13	-1013.50	-1793.36	-1133.34	-1919.79	-1195.99

Significance Level: † sig. at 10%level, * sig. at 5%level, ** sig. at 1%level.

the first 30 periods, it has little influence on our analysis.

It is straightforward to see that none of the coefficients of the dummy variable *Religious/SB* are significant in Panel B, but they are negative and significant in Panel A when the recipient is an outgroup member. These results suggest that we find clear evidence showing that when choices are single-blind, religious thinking has little effect on giving behavior when the recipient is Muslim but significant negative effects when the receiver is Han Chinese. To estimate the effects of religious thinking when choices are double-blind, we test whether the coefficient on the treatment Religious/DB is significantly different from the coefficient on treatment Baseline/DB. Our religious thinking results when privacy is double-blind are consistent with previous results. We find robustness evidence that there is significantly higher prosocial behavior towards outgroup members overall with religious thinking and for the value of T=7 below the 1% level, for T=3 below the 5% level and for T=5 at the 6% level when choices are double-blind.[13] We find weaker

---

[13]For the overall comparison of the two coefficients the χ^2 statistic= 8.97, $P = 0.003$; for $T = 3$, it equals 6.55, $P = 0.010$; for $T = 5$, it equals 3.46, $P = 0.063$; for $T = 7$, it equals 8.24, $P = 0.004$. We also conducted the comparisons based on all the data, and they yielded qualitatively the same results.

effects of religious thinking on voting for giving to ingroup members: Religious thinking has no significant effect on overall prosocial behavior toward ingroup members but significantly higher prosocial behavior toward ingroup members at the 8% level when T=3 and choices are double-blind.[14]

Taken together, we find consistent and robust evidence showing that religious thinking in the double-blind setting results in significantly higher charitable contributions to outgroup members, but religious thinking also results in significantly lower charitable giving behavior in the single-blind setting. The opposite observations in different environments demonstrate that the effects of religious thinking on charitable contributions are highly related to whether giving behavior itself violates the religious prohibition of showing off good behavior. Again, the only difference between our single-blind and double-blind setting is whether or not the decision making environment is in line with religiosity of Islam. Even in the single-blind setting, subjects' choices are anonymous to other participants, which means there is no social pressure. The influence of extrinsic utility is largely controlled by the design of our experiment. Hence, we find strong evidence that supports our view of the underlying mechanism of our findings and we conclude that the effects of the religious thinking exercise on charitable contributions to non-Muslims and non-believers should be explained by the religiosity of Muslims.

4 Discussion and Implications

In this study, we conduct a novel experiment using a unique subject pool of Chinese Muslims and demonstrate that religious thinking has a salient and significant effect on Muslim subjects' giving decisions towards outgroup members. We use nonparametric statistical tests in our analysis and discuss our results at a conventional level of significance (0.05). According to a large number of behavioral scientists, new findings should be subject to more stringent tests. For example, Benjamin, Berger, Johannesson et al. (2018) propose

---

[14]For the overall comparison of the two coefficients the χ^2 statistic= 0.80, $P = 0.371$; for $T = 3$, it equals 2.99, $P = 0.084$; for $T = 5$, it equals 0.09, $P = 0.758$; for $T = 7$, it equals 0.49, $P = 0.484$. We also conducted the comparisons based on all the data, and they yielded qualitatively the same results. We also used the panel data method and estimated the models with random effects across sessions. We have qualitatively the same findings in these other models.

that scholars should change the default P-value threshold for statistical significance from 0.05 to 0.005 for claims of new discoveries. According to their standard we do not have a significant result. Nevertheless, after examining the findings carefully and conducting the false discovery check, we expect that our study provides suggestive evidence of religion's effect on charitable contributions to outgroup members and nonbelievers and we argue that this article has strong implications for our understanding of the influence of Islamic rules on Muslims' attitude and behavior towards non-Muslims and for the design of fundraising mechanisms in Muslim communities.

First, the tension between religious groups and western societies has become particularly salient post 9/11. A controversial topic that has been discussed is whether religion may be the source of conflict since it might motivate people to treat the other - "nonbelievers" worse. There are a number of historical cases of wars caused by the intolerance of other religious believers. Our results find that the effect of religion causes Muslim subjects to treat non-Muslims nicer, which is interesting and striking given that contrast. It implies that religion may not be a necessarily motivating factor in violence. Of course, our findings arise from a decision-making environment established in the laboratory and administered to a sample of Muslim subjects in a society in which they are the minority, so our results may be subject to the context of Muslims in this situation and religiosity of Islam. However, when some commentators in social media and policymakers connect Islam to terrorism, it is especially meaningful to conduct an experimental study as to whether Islam incites people to violence against non-Muslims.[15] Our study, while not directly addressing this question, makes remarkable contributions to this end.

Second, while our findings are somewhat counter-intuitive, they are consistent with a world in which group competition affects the benefits and costs of prosocial behavior, and these costs and benefits become less salient when engaging in religious thinking leading to different effects on prosocial behavior depending on the group membership of the recipient. Acting prosocially towards members within a group may be an analytical response to competition with outgroups; while acting prosocially towards outgroup members may be

---

[15]See discussions in https://www.npr.org/2017/06/19/532963059/when-is-it-terrorism-how-the-media-covers-attacks-by-muslim-perpetrators.

a more religiosity driven response.

Giving is part of Muslims' religious practice and daily life, it is a socially desirable behavior. Then our Muslim subjects might have extrinsic motivations to engage in charitable contributions that may be driven by the fear to be seen as faithless by others, and there might be social pressure when the majority of others vote for option B while one does not. It is essential to note that such extrinsic motivations have been largely controlled in our study because, in both the single-blind and the double-blind sessions, Muslim subjects' voting and giving decisions are anonymous to each other. Hence, we argue that what drives our findings is the intrinsic motivation to please Allah.

As noted above, Islam preaches that Muslims should treat non-Muslims with respect and that all humans are created from the same dirt. When the recipient is a member of the group that has a competitive relationship with one's own group, in the Religious Thinking treatment Muslim subjects' behavior may be dominated by the desire to obey Islamic rules. Religious thinking reminds them of the norm that Muslims should treat non-Muslims equally, while in-group favoritism is not promoted, and that explains why religious thinking has limited and not significant effects when the recipient is Muslim. It is worth noting that the intrinsic utility from giving to outgroup members seems larger in the double-blind setting as there is no social recognition in that environment which is in line with religiosity of Islam, so Muslim subjects tend to be more willing to give to outgroup members in private. Future research may explore whether Muslim subjects are more likely to vote for ingroup members with analytical thinking/training.

Third, our findings demonstrate that the joint effects of religious thinking and privacy mainly affect giving behavior towards outgroup members, and they have little influence on charitable giving towards ingroup members. It implies that future research of fundraising campaigns, especially in Islamic communities within a secular country, needs to be cautious when using religious and/or ethnic identity of the receiver to promote giving in the process of fundraising.

Finally, one important methodological implication of our results is that in experiments on religious people's charitable contributions, the level of privacy may introduce uninten-

tional distortions in the outcomes. That is, if there is a religious prohibition of showing off good behavior such as charitable contributions, experimentation using a single-blind setting may lead to a misinterpretation of the results. This article increases our understanding of the extent experimental data different under different privacy regimes.

References

Alpizar, Francisco, Fredrik Carlsson and Olof Johansson-Stenman. 2008. "Anonymity, reciprocity, and conformity: Evidence from voluntary contributions to a national park in Costa Rica." *Journal of Public Economics* 92(5-6):1047–1060.

Andreoni, James. 1989. "Giving with impure altruism: Applications to charity and Ricardian equivalence." *Journal of Political Economy* 97(6):1447–1458.

Andreoni, James. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The Economic Journal* 100(401):464–477.

Andreoni, James. 1995. "Warm-glow versus cold-prickle: the effects of positive and negative framing on cooperation in experiments." *Quarterly Journal of Economics* 110(1):1–21.

Andreoni, James, Justin M Rao and Hannah Trachtman. 2017. "Avoiding the ask: A field experiment on altruism, empathy, and charitable giving." *Journal of Political Economy* 125(3):625–653.

Andreoni, James and Ragan Petrie. 2004. "Public goods experiments without confidentiality: a glimpse into fund-raising." *Journal of Public Economics* 88(7-8):1605–1623.

Arrow, Kenneth J. 1972. "Gifts and exchanges." *Philosophy & Public Affairs* pp. 343–362.

Benjamin, Daniel J, James J Choi and Geoffrey Fisher. 2016. "Religious identity and economic behavior." *Review of Economics and Statistics* 98(4):617–637.

Benjamin, Daniel J, James O Berger, Magnus Johannesson et al. 2018. "Redefine statistical significance." *Nature Human Behaviour* 2(1):6.

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the false discovery rate: a practical and powerful approach to multiple testing." *Journal of the Royal Statistical Society. Series B (Methodological)* pp. 289–300.

Bergstrom, Theodore, Lawrence Blume and Hal Varian. 1986. "On the private provision of public goods." *Journal of Public Economics* 29(1):25–49.

Bernheim, B Douglas. 1986. "On the voluntary and involuntary provision of public goods." *American Economic Review* pp. 789–793.

Cason, Timothy N and Vai-Lam Mui. 1997. "A laboratory study of group polarisation in the team dictator game." *The Economic Journal* 107(444):1465–1483.

Chuan, Amanda, Judd B Kessler and Katherine L Milkman. 2018. "Field study of charitable giving reveals that reciprocity decays over time." *Proceedings of the National Academy of Sciences* p. 201708293.

Croson, Rachel and Uri Gneezy. 2009. "Gender differences in preferences." *Journal of Economic Literature* 47(2):448–74.

Crumpler, Heidi and Philip J Grossman. 2008. "An experimental test of warm glow giving." *Journal of Public Economics* 92(5-6):1011–1021.

Dana, Jason, Roberto A Weber and Jason Xi Kuang. 2007. "Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness." *Economic Theory* 33(1):67–80.

Dufwenberg, Martin and Astri Muren. 2006. "Generosity, anonymity, gender." *Journal of Economic Behavior & Organization* 61(1):42–49.

Feddersen, Timothy, Sean Gailmard and Alvaro Sandroni. 2009. "Moral bias in large elections: theory and experimental evidence." *American Political Science Review* 103(2):175–192.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10(2):171–178.

Harbaugh, William T, Ulrich Mayr and Daniel R Burghart. 2007. "Neural responses to taxation and voluntary giving reveal motives for charitable donations." *Science* 316(5831):1622–1625.

Lambarraa, Fatima and Gerhard Riener. 2015. "On the norms of charitable giving in Islam: Two field experiments in Morocco." *Journal of Economic Behavior & Organization* 118:69–84.

List, John A, Robert P Berrens, Alok K Bohara and Joe Kerkvliet. 2004. "Examining the role of social isolation on stated preferences." *American Economic Review* 94(3):741–752.

Luhan, Wolfgang J, Martin G Kocher and Matthias Sutter. 2009. "Group polarization in the team dictator game reconsidered." *Experimental Economics* 12(1):26–41.

Null, Clair. 2011. "Warm glow, information, and inefficient charitable giving." *Journal of Public Economics* 95(5-6):455–465.

Rege, Mari and Kjetil Telle. 2004. "The impact of social approval and framing on cooperation in public good situations." *Journal of Public Economics* 88(7):1625–1644.

Roberts, Russell D. 1984. "A positive model of private charity and public transfers." *Journal of Political Economy* 92(1):136–148.

Shariff, Azim F, Aiyana K Willard, Teresa Andersen and Ara Norenzayan. 2016. "Religious priming: A meta-analysis with a focus on prosociality." *Personality and Social Psychology Review* 20(1):27–48.

Warr, Peter G. 1982. "Pareto optimal redistribution and private charity." *Journal of Public Economics* 19(1):131–138.

Supplemental Online Appendix

A. Notes on Field Implementation

Before conducting the experiments in the field, we assembled a focus group with nine local Muslim students. They were invited to join a free discussion on the choices of questions for the religious thinking exercise. The experimenters wanted to use the most neutral and relevant but not "extreme" or "misleading" questions in the experiment. It is not impossible that some religious questions could be sensitive to the subjects who participated in our study, especially when the experimenters are not Muslims.

The focus group participants were given 107 Islam-relevant religious questions. These Islam-relevant questions were found online. By using the Chinese search engine (www.baidu.com), it is not difficult to find question sets on Islamic daily life questions.

To double check the validity of the answers to these questions, the experimenters also explored a few most influential and popular Chinese Muslim online communities, such as www.ms15.com and www.muslimwww.com. The focus group participants were asked to answer these questions. They were asked to eliminate the obvious easiest and hardest questions, and select 30 questions they believed to be moderate and neutral. Then they were given the answers, and they were asked to vote to decide the final 20 questions that we used in the experiment. Translations of the religious thinking questions can be found below.

Religious Thinking Exercise

1. What is the meaning of "al-Salam'Alaykum" (Chinese pronunciation: An Se Lia Mu Er Lai Ku Mu) in Chinese?

   A. Hello

   B. Goodbye

2. Who is the ancestor of human beings?

   A. Prophet Adam

   B. Jesus Christ

3. Learning is the ＿＿＿ of Muslims. Fill-in-the-blank.

4. Is Quran only given to Muslims?

   A. Yes

   B. No

5. The person that Muslims should respect the most is:

   A. Father

   B. Mother

6. Which one is more important, Jum'a or Eid?

   A. Jum'a

   B. Eid

7. What is the meaning of "Tajweed" (Chinese pronunciation: Tai Zhi Wei De) in Chinese?

   A. The standard way of 'singing' Quran

   B. The Islamic World

8. Putting on a headscarf (or a veil) is＿＿＿?

   A. Sunnah

   B. Godly prescription, or "ana'l-haqq"

   C. Custom

9. What are the most famous/influential Islamic temples (al-Haram)?

   A. Al-Masjid al-Haram, Al-Masjid an-Nabawi, Al-Masjid al-Aqsa

   B. Al-Masjid al-Haram, Great Mosque of Xi'an, Al-Masjid al-Aqsa

   C. Al-Masjid al-Haram, Al-Masjid an-Nabawi, Great Mosque of Yinchuan

10. What are the most important traditional holidays in Islam? [or What are the most important days in Islam?]

    A. Eid al-Adha, Eid al-Fitr, and Jumah

    B. Eid al-Adha, Eid al-Fitr, and Spring Festival

    C. Eid al-Adha, Mid-Autumn Festival, and Jumah

11. When do boys have the responsibility/duty for ana'l-haqq?____ When do girls have the responsibility/duty for ana'l-haqq?____ Fill-in-the-blank.

12. Is lottery allowed in Islam?

    A. Yes

    B. No

13. How many chapters does the Quran consist of ? ____ Fill-in-the-blank.

14. What is the only holy book that has never been manipulated/tampered/garbled/misrepresented?

    A. Quran

    B. Bible

    C. Diamond Sutra

15. Although different sages may preach different laws, there is one thing that every sage preaches uniquely.

    A. Honor your parents

    B. Help brothers

    C. Tawhid, Worship the Lord alone

16. Is it right to perform Tatammum when you cannot find water?

    A. Yes

    B. No

17. To do worship, there are ___ inside ana'l-haqq and ___ outside ana'l-haqq. Fill-in-the-blank.

18. In the eyes of the God/Allah, what is the most valuable worship?

    A. On time worship

    B. Sincere worship

19. In what age are boys expected to do worship? ____ Fill-in-the-blank.

20. According to Islamic calendar, when should Muslims fast? ____ Fill-in-the-blank.

B. Translated Experimental Instructions for the Computerized Baseline/Muslim Treatment

Welcome to our experiment. During the following experiment, we require your complete and undivided attention and ask that you follow the instructions carefully. Please turn off your cell phones. For the duration of the experiment, do not take actions that could distract you or other participants. Peeking at other participants' decisions are not allowed during the session. And do not let others observe your decisions. If you have any questions during the experiment, please raise your hand. The experimenter will come to you privately and answer your question. If we think the question is of a general nature, we will announce the question and the answer to everyone. Please restrict your questions to clarifications about the directions only. If you break silence while the experiment is in progress, you will be asked to leave the experiment with no pay.

Procedure

You will be independently and randomly assigned to a committee with the other 2 participants. Each committee has 3 participants all of whom are Muslims. Your committee will be asked to collectively decide whether to give an amount of money to a needy Muslim undergraduate student in your university. As you can see, on each of your desks, there is a sheet of which you can find the needy person's basic information. We have 18 participants, so we will have 6 committees. Each committee is matched with one needy person. That is the reason why you can find 6 needy persons' information. As you can

see, we do not show their names, but they are real and they are selected from the database of the "needy-person" of your university. The basic information includes gender, ethnic identity, birthplace, etc.

Your committee will collectively decide whether to help a needy person. Your committee's decision is based on a simple majority rule. For example, please see a payoff matrix below

	Option A	Option B
Your Income	10	8
Needy Person's Income	0	6

At each time, you can vote for Option A or Option B. If you vote for Option A, it means you do not want to help the needy person at that level. If you vote for Option B, it means you want to help the needy person at that level. In your committee, if there are two or more votes for Option A, then Option A is your committee's decision. If there are two or more votes for Option B, then Option B is your committee's decision.

If your committee's decision is Option A, then your income will be 10 RMB, the needy person's income will be 0 RMB. If your committee's decision is Option B, according to the example above, your income will be 8 RMB, which means that each of you will give $(10 - 8 =)$ 2 RMB to the needy person, and the needy person will have $(2 \times 3 =)$ 6 RMB. The payoff table above is an example. In the experiment, you will see different payoff tables in the different periods.

Your committee's decision will be implemented for every committee members, which means even if your vote is different from your committee's decision, you need to follow the group decision. In the example above, if you vote for Option A, for example, but Option B is your committee's decision, Option B will be implemented. Your income will be 8 RMB, the needy person's income will be 6 RMB. If you vote for Option B, but Option A is your committee's decision, Option A will be implemented. Your income will be 10 RMB and the needy person's income will be 0 RMB.

The voting will be repeated for 60 periods. The payoff tables and the amount of money that will be transferred to the needy person can be different from period to period. Please pay attention to the instructions and read the payoff table carefully in every period.

Earnings

In every 20 periods, 1 period will be randomly selected as the period to be paid. Because we have 60 periods, at the end of the session, 3 periods will be randomly selected to be paid. Your committees' decisions in the selected 3 periods will determine your income and the needy person's income. Please note that the experimenters have selected the 3 periods and write them on board. These 3 numbers have been hidden behind the opaque papers. At the end of the session, these numbers will be revealed to you publicly.

Your accumulated income of these pre-selected periods will be used to calculate your cash payment. You will be paid in cash. Besides, you will get an additional 10 RMB as showing up payment. Whatever your or your committee's decision is, you will get the 10 RMB showing up payment.

Anonymity

All your decisions and your earnings will remain anonymous.

If you have any questions, please raise your hand. Otherwise, please proceed to answer the questions on the screen. The purpose of the questions is to make sure that you understand the different elements of the experiment. Any unclear question will be explained by the experimenter. Once you have answered all the questions, please press the "Continue" button to proceed. The computer will check your answers.