



Re-examining the role of family relationships in structuring perceived helping obligations, and their impact on moral evaluation[☆]

Ryan M. McManus^{*}, Jordyn E. Mason, Liane Young

Department of Psychology and Neuroscience, Boston College, United States

ARTICLE INFO

Keywords:

Obligation
Helping
Moral character
Moral values
Social interactions

ABSTRACT

Although recent research has highlighted that social relationships influence moral judgment, many questions remain. Across two pre-registered experiments (total $N = 1310$), we investigated one social relationship and its link to morality: kinship and its obligations. Experiment 1 varied genetic relatedness between helpers and beneficiaries (i.e., strangers, cousins, siblings), investigating differences in perceived obligations to help and downstream moral evaluations of helpers. Experiment 2 investigated whether these patterns varied via agents being estranged versus friendly, and whether relatedness impacts obligation judgments through other social interaction inferences (e.g., social closeness, frequency of prior help). Before helping occurred, agents were judged as having stronger obligations toward relatives than strangers, and closer relatives (i.e., siblings) than distant relatives (i.e., cousins). After helping occurred, agents who helped strangers were judged as more morally good than agents who helped relatives, but agents who helped strangers instead of relatives (or cousins instead of siblings) were judged as less morally good than agents who did the opposite. Perceived obligation differences shaped moral evaluation differences at the individual level only in contexts where agents helped one beneficiary over another. Importantly, social interaction inference differences were always more strongly correlated with obligation judgment differences than relatedness judgment differences were. Additionally, endorsement of family values and ingroup-loyalty correlated positively with obligations toward family, whereas endorsement of impartial beneficence correlated positively with obligations toward strangers. By broadening the theoretical and methodological scope of prior work, this research offers a richer characterization of some of the determinants and consequences of perceiving obligations to help.

1. Introduction

I know I should have been with you at such a difficult hour ... but you would not have liked me to leave them [the lepers] uncared for. I would have had my pleasure of your company only at the cost of hopes and aspirations of the poor, helpless, lepers ... Do you really think it is right to have me by your side in these circumstances?

- Excerpt from *Strangers Drowning* (MacFarquhar, 2015)

Morally speaking, many stories in Larissa MacFarquhar's book *Strangers Drowning* are simultaneously inspiring and confusing. In the excerpt above, readers are introduced to a man who devotes his life to living among and caring for sick strangers—lepers—rendering him a moral exemplar. However, the excerpt is taken from a letter to his wife—who at the time had fallen very ill herself and who, along with

their ill infant, had to travel away from her husband for treatment. These details, side by side, detract from the man's moral status and raise key questions about folk moral psychology.

In the current paper, we build on recent research that has begun to investigate questions that arise from situations in which the welfare of unknown strangers and close others are pitted against one another (Everett, Faber, Savulescu, & Crockett, 2018; Hughes, 2017; Marshall, Wynn, & Bloom, 2020; Marshall et al., 2021, *invited revision*; McManus, Kleiman-Weiner, & Young, 2020). Importantly, we go beyond prior work—in theory, methods, and measures—by conducting two experiments investigating whether and in what ways: (1) Beliefs about pro-social obligations to help depend on the degree of relatedness between the helper and the helped; (2) Differences in perceived obligation strength shape differences in moral evaluations of helpers; and (3) Moral values influence beliefs about obligations to help closely related,

[☆] This paper has been recommended for acceptance by Professor. Rachel Barkan.

^{*} Corresponding author.

E-mail address: mcmmanurd@bc.edu (R.M. McManus).

distantly related, and unrelated others.

2. Literature review

Recently, many empirical studies (e.g., Curry, Chesters, & van Lissa, 2019; Curry, Mullins, & Whitehouse, 2019; Everett et al., 2018; Hughes, 2017; Hughes, Creech, & Strosser, 2016; Kurzban, DeScioli, & Fein, 2012; Lee & Holyoak, 2020; Lieberman & Lobel, 2012; Marshall et al., 2020; McManus et al., 2020; Simpson, Laham, & Fiske, 2016; Sznycer, De Smet, Billingsley, & Lieberman, 2016; Tepe and Aydinli-Karakulak (2019); Uhlmann, Zhu, Pizarro, & Bloom, 2012; Waytz, Dungan, & Young, 2013; Weidman, Sowden, Berg, & Kross, 2020; Yudkin, Gantman, Hofmann, & Quoidbach, 2021) and conceptual analyses (e.g., Berry, Lewis, & Sowden, 2021, in press; Curry, 2016; Hester & Gray, 2020; Rai & Fiske, 2011; Schein, 2020; Tomasello, 2020) have focused on, and argued in favor of, the importance of understanding relationship ties in moral judgments and behavior. As the space of relationships is large and multidimensional, morality within this space will be incredibly complex. Here, we start mapping this terrain by zeroing in on a specific kind of relationship and its link to morality: kinship and the obligations it entails.

2.1. A sense of familial obligation

First, we review evidence that people report being more likely to help and protect their family members, compared to non-family. We argue that these first-person intentions point to a sense of obligation that is absent (or at least weaker) in the case of non-family.

Research abounds suggesting that people are more willing to help kin than non-kin. For example, in one study, people reported being more willing to help a negligent sibling than an acquaintance who was in danger through no fault of their own (Greitemeyer, Rudolph, & Weiner, 2003). Relatedly, people reported being most likely to help a sibling, next most likely to help a cousin, and least likely to help an acquaintance if they had time to help only one person (Burnstein, Crandall, & Kitayama, 1994). Similarly, people report being more willing to help full siblings than half- or step-siblings (Sznycer et al., 2016), and family members compared to friends and strangers (Passarelli & Buchanan, 2020). Finally, in a convincing in-lab experiment, Madsen et al. (2007) monetarily incentivized engagement in uncomfortable physical exercise (i.e., wall squats), paying more to those who held the position for longer. People held the uncomfortable position for longer when the beneficiary was a 50% genetic relative (e.g., parent or sibling) than when the beneficiary was a 25% (e.g., grandparent) or a 12.5% genetic relative (e.g., cousin).

People are also sensitive to these distinctions when contemplating decisions that would protect their kin. For example, people were more certain that they would report a stranger than a sibling for committing identical crimes (Lee & Holyoak, 2020), an effect that is enhanced when people imagine a distant other versus a close other committing severe crimes (Weidman et al., 2020). In other work, people reported being less likely to “blow the whistle” (Waytz et al., 2013) and more willing to pay for a transgressor’s crime (Linke, 2012) when hypothetical perpetrators were family members, compared to close friends, acquaintances, or strangers. People’s predictions about their own behavior in sacrificial moral dilemmas also reveals their consideration of whether or not the to-be-sacrificed others are family (Bleske-Rechek, Nelson, Baker, Remiker, & Brandt, 2010; Kurzban et al., 2012; Petrinovich, O’Neill, & Jorgensen, 1993).

Overall, the available evidence suggests that people experience a sense of familial obligation, and perhaps an especially strong obligation to help or protect their closely related family members (e.g., siblings and parents), compared to more distantly related family members (e.g., cousins) or non-family (e.g., strangers).

2.2. Perceptions of familial obligations and their role in moral evaluations

Although it seems clear that people themselves experience a sense of familial obligation, whether people believe *others* ought to adhere to these obligations, and whether these beliefs play a role in their moral evaluations of others, is an area of ongoing research.

For example, adolescents and young adults judged that ultimately not helping genetic relatives was more wrong than not helping non-genetic relatives (see Killen & Turiel, 1998). Relatedly, people judged not donating bone marrow to a critically in-need patient as more wrong when the in-need patient was the potential donor’s cousin compared to a stranger (Baron & Miller, 2000). Additionally, when considering stories of agents whose personal desires conflicted with requests to spend time with or support a close other, more people judged an agent’s fulfilling their personal desire as unacceptable when the requester was a family member versus a friend (Neff, Turiel, & Anshel, 2002). In the context of prosocial dilemmas, agents who chose to help (or simply endorsed helping) a larger number of strangers instead of a family member were judged as less moral than agents who did the opposite (Everett et al., 2018; Hughes, 2017). Similar effects emerge when people judge the moral acceptability of helping a larger number of socially distant others (e.g., people in another country) instead of a smaller number of socially close others (e.g., friends; Law, Campbell, & Gaesser, 2021). Although a violation of an obligation was offered as one potential mechanism for these effects, this hypothesis was not directly tested.

In more recent work, McManus et al. (2020) built on these findings by attempting to experimentally control for social distance, describing family members as otherwise like strangers (e.g., “having not seen or spoken to one another in years”), and using relatives that were distantly related (e.g., second cousin) to the target agent. In these experiments, non-dilemmas were also investigated (i.e., contexts in which it was clear that an agent’s helping someone did not come at the cost of neglecting someone else), with agents who helped a stranger being judged as more morally good than agents who helped a family member. However, in prosocial dilemmas contexts (like those used in Hughes, 2017; Everett et al., 2018; Law et al., 2021), agents who helped a stranger instead of a family member were judged as *less* morally good than agents who did the opposite. Importantly, McManus et al. (2020) measured perceptions of obligations’ being violated versus fulfilled, finding that helping strangers was judged as fulfilling an obligation less than helping family members, providing indirect support for the role of familial obligations in moral evaluations. Consistent with these findings, agents who helped strangers were judged as nicer than agents who helped friends (Marshall et al., 2020), but agents who failed to help strangers were judged as less immoral or less mean than agents who failed to help friends (Haidt & Baron, 1996; Marshall et al., 2020). Overall, the available evidence suggests that perceptions of familial obligations may in turn influence moral evaluations. However, the cited work leaves open the question of whether and exactly how perceived obligations directly shape moral evaluations.

2.3. Methodological concerns in McManus et al. (2020)

While McManus et al. (2020) provides the most comprehensive investigation of familial obligations and moral evaluations, we highlight two methodological problems in this work, and their consequences on inference. First, when participants made obligation judgments, they responded to an item that read “To what extent did X violate or fulfill an obligation they had?” with response options ranging from “completely violated” to “completely fulfilled.” Because of the semantic anchors used and how this item was worded, it is unclear how participants interpreted it. For example, what would it mean for one agent to “somewhat” fulfill an obligation, and another agent to “completely” fulfill an obligation? A more interpretable measure would assess the presence or strength of an obligation rather than its graded violation or fulfillment. Second, participants made obligation and moral character judgments

simultaneously only *after* the outcome of the scenario was known. However, prescriptive judgments (like obligation judgments), by their nature, are future-oriented (Malle, 2021). Because participants made a prescriptive judgment after the outcome was known, this work could not disentangle whether obligation judgments were inputs to moral character judgments, or moral character judgments were retroactively contaminating obligation judgments.

In our experiments, we adopt a pre-/post-outcome design (see Marshall et al., 2020; Marshall et al., 2021, *invited revision*) that allow us to measure prescriptive judgments *before* the outcome and moral character judgments *after* the outcome. Two important consequences follow from these methodological changes. First, with a new measure and a new pre-/post-outcome design, it is possible to answer the question of whether, on average, differences in perceived obligation strength correspond to differences in perceived moral character without compromising the hypothesized temporal link between these two judgment types (Malle, 2021), and therefore eliminating the possibility of moral character judgments contaminating obligation judgments. Second, this question can (and ought to) be answered at multiple levels of analysis. Beyond the statistical conclusion of mean differences for each judgment, it is possible to further probe their relationship with one another. That is, assuming that there is variability in by-relationship obligation judgments and by-relationship moral character judgments, is by-relationship discrimination in perceived obligation strength related to by-relationship discrimination in perceived moral character? For example, is it the case that there is a trend in which high discrimination in obligation judgments is associated with high but oppositely signed discrimination in moral character judgments? Importantly, this difference score correlation analysis can reveal results that are consistent or inconsistent with the mean difference analyses. Therefore, each of these tests is crucial for making claims about the relationship between obligation judgments and moral character judgments at multiple levels of analysis. Although these tests were possible in McManus et al. (2020), because of the measurement and design issues, making strong inferences from such tests would have been unwarranted.

3. Current research

In the current work, we go beyond prior research to explore the possibility that people will judge others as having an especially strong obligation to help their closest genetic relatives, followed by a weaker obligation to help more distant genetic relatives, and, last, an even weaker (or perhaps absent) obligation to help unrelated others. We focus our experiments on siblings, cousins, and strangers, as previous research suggests that people reliably distinguish among these categories when judging perceived kinship and when reporting their own helping/saving intentions (Bleske-Rechek et al., 2010; Burnstein et al., 1994)¹.

Experiment 1 is a high-powered, pre-registered experiment that investigates non-dilemma and prosocial dilemma contexts. Furthermore, in relation to prior work, Experiment 1 offers a richer characterization of the determinants and consequences of differential obligations by examining whether obligation judgments are influenced by self-reported moral values, and by directly testing whether and how obligation judgments influence moral character judgments. Experiment 2 is an even higher-powered, pre-registered replication that investigates whether describing the helping agents' relationships with their relatives as "stranger-like" versus "friend-like" affects judgments of obligation, moral character, and their relationship with one another. Experiment 2 also assesses whether relatedness impacts obligation judgments through

¹ Burnstein et al. (1994) tested family members against "acquaintances" rather than "strangers." Therefore, that Burnstein et al. used "acquaintances" and found that people discriminate between them and family members makes it especially likely that people will discriminate between the relationship categories studied here.

a suite of social interaction inferences (e.g., social/emotional closeness, frequency of past help, etc.). Importantly, all of our experiments target everyday, low-stakes helping behavior, rather than high-stakes (e.g., life or death) situations; this approach is a conservative test of our hypotheses given past work's showing enhanced discrimination among genetic relatives in high-stakes cases (Burnstein et al., 1994).

We outline the general logic behind our hypotheses here. If people are more likely to help related versus unrelated (or closely related versus distantly related) others in hypothetical (e.g., Burnstein et al., 1994) and real-world situations (e.g., Madsen et al., 2007), this suggests that people believe that they have a *stronger obligation* to help closely related others than they do to help more distantly related (or unrelated) others. Therefore, if people believe that they have stronger obligations to closely related others than they do to more distantly related (or unrelated) others, then they may also use this information when evaluating others, resulting in differences in third-person moral evaluations. The logic of this assumption is consistent with research showing that first-person moral beliefs shape third-person moral judgments (e.g., Niemi & Young, 2016). More specifically, in contexts where helpers do not have to choose between multiple potential beneficiaries, we predict that people will judge a helper who fulfills a stronger obligation as less morally good than a helper who fulfills a weaker (or non-existent) obligation. This hypothesis is broadly consistent with attribution theory (Kelley, 1967), as an obligation is a situational feature that makes it relatively more difficult for a third party judge to infer whether the helper has true prosocial motives. However, in contexts where helpers do (or must) choose between multiple beneficiaries, we predict that people will judge a helper who fulfills their stronger obligation as more morally good than a helper who fulfills their weaker (or non-existent) obligation. This hypothesis is broadly consistent with Relationship Regulation Theory (Rai & Fiske, 2011) and Morality-as-Cooperation (Curry, 2016). RRT suggests that communal sharing relationships (like those with family) carry with them inherent obligations that, if violated, will be judged negatively; similarly, MAC suggests that helping family is a universal moral good.

4. Open science

All materials, including experimental stimuli, data, and analysis code and output are available on our Open Science Framework (OSF) page at (https://osf.io/expmr/?view_only=6e80bfb029e64d02baaaf276a61d425). In these experiments, we report all measures, manipulations, and exclusions. All analyses were conducted in R version 4.0.4 (R Core Team, 2021); raw data were wrangled and plotted using the *tidyverse* package (Wickham et al., 2019); all descriptive statistics were computed using the *psych* package (Revelle, 2020); all standardized effect sizes were computed using *effsize* (Torchiano, 2020), *sjstats* (Ludecke, 2020), and *correlation* packages (Makowski, Ben-Shachar, Patil, & Ludecke, 2020); and correlation comparisons were conducted with the *cocor* package (Diedenhofen & Musch, 2015) using Steiger's method (Steiger, 1980). Experiment 1 was pre-registered at https://osf.io/6q4kr/?view_only=8033c65c9dfc41bcb528fe94409e4e8, and Experiment 2 was pre-registered at https://osf.io/uxysz/?view_only=cad9b09ce5d1494eb023a604d9f6feb3.

Importantly, because we did not pre-register outlier removal methods, all analyses are conducted on data which simply excludes duplicate responses and participants who failed any attention check. However, for all reported condition comparisons and tests of association, we conducted robustness checks and demonstrated that none of our results were the product of statistical outliers (see the supplemental online materials [SOM] on our OSF page).

5. Experiment 1

Experiment 1 was conducted to address limitations of and extend recent work on perceptions of familial obligations and their impact on

moral evaluations. First, and most importantly, we extended the theoretical scope of prior work by explicitly manipulating *how closely related* agents were to their potential (and eventual) beneficiaries, as opposed to simply manipulating *whether* agents were related to their beneficiaries. Second, in past work (McManus et al. (2020)), people simultaneously made judgments of the extent to which agents violated or fulfilled an obligation and how morally bad or good agents were, but only after helping had already occurred. Experiment 1 improved upon these methods in two ways by employing a more interpretable measure of our key concept of obligation and measuring obligation judgments before helping occurs. Third, unlike prior work, we probed the relationship—at various levels of analysis—between obligation judgments and moral character judgments. Fourth, we also investigated how individual differences in moral values influence perceptions of obligations toward differently related potential beneficiaries.

5.1. Method

5.1.1. Participants

All participants were U.S. residents recruited and compensated via Prolific² (Palan & Schitter, 2018). We decided a priori to collect data from 690 participants to obtain 600 analyzable responses. This sample size was chosen to yield 200 responses per between-participants condition (see *Statistical Power* for further justification). However, when applying our pre-registered exclusion criteria (excluding duplicate IP addresses and participants who failed any of three attention checks), we did not meet the 200 responses threshold for one of our between-participants conditions. Therefore, we collected data from an additional 10 participants; after reapplying our exclusion criteria, this yielded the desired sample size for each between-participants condition ($N_{Total} = 611$, 52.4% female, $M_{Age} = 31.52$ years).

5.1.2. Design

This experiment used a 2 (Relation: Distant vs Close) x 2 (Choice Context: No Choice vs Choice) x 3 (Relatedness Between Beneficiaries) mixed design, in which “Relation” and “Choice Context” were manipulated within-participants, whereas “Relatedness Between Beneficiaries” was manipulated between participants (see below). Specifically, participants were asked to make judgments of an agent who helped (1) a genetically distant other (e.g., cousin), (2) a genetically closer other (e.g., sibling), (3) a genetically distant other *instead of* a genetically closer other, and (4) a genetically closer other *instead of* a genetically distant other. Additionally, participants were randomly assigned to one of three conditions which varied in how related the target agents were to *both* beneficiaries. One group of participants read stories involving agents helping strangers and siblings; a second group of participants read stories involving agents helping strangers and cousins; and a third group of participants read stories involving agents helping cousins and siblings. Importantly, “No Choice” conditions did not mean that the agents had to help; rather, agents decided whether to help, but they did not have to make a choice about whom to help—only one possible beneficiary was present in the scenario. Similarly, “Choice” conditions did not mean that agents were forced to make a choice; instead, agents decided to help, but they chose to help one of two potential beneficiaries. Consistent with McManus et al. (2020), genetic relatives were always described as otherwise stranger-like.

We note that the intention behind this design was not to conduct hypothesis tests by comparing between-participants conditions, as we would lack adequate statistical power to detect the small expected effects for such comparisons. Rather, the between-participants conditions serve as internal replications of within-participants effects under slightly

varied conditions.

5.1.3. Procedure

After consenting, participants were randomly assigned to one of the three between-participants conditions. Before engaging in the experimental task, participants completed an ostensibly unrelated judgment task in which they rated the moral character of two agents who differed dramatically in their extreme (im)moral behavior. This task was included to decrease the likelihood of floor or ceiling effects in the experimental task, where agents engaged in relatively less extreme behavior. Following the non-experimental task, participants learned that they would be presented with four short stories and be asked to evaluate people in each story. The presentation of condition order was randomized across participants, and no story was repeated across conditions within the same participant (with all possible permutations of stories to conditions being evenly presented across participants [four stories to four conditions = 24 permutations in total]). Experiment 1 used a two-stage procedure adopted from developmental research on a similar topic (Marshall et al., 2020; Marshall et al., 2021, *invited revision*), with a pre-outcome judgment task followed by a post-outcome judgment task for each story (see Table 1 for an example scenario and its variants). After participants finished the experimental task, they filled out subscales from three separate moral values questionnaires. The order of questionnaires and the order of items within each questionnaire were randomized across participants.

5.1.4. Experimental Task Measures

In the pre-outcome segment, participants made judgments about how much of an obligation the agent had to help ($0 = none\ at\ all$ to $100 = a\ great\ deal$). In the post-outcome segment, when the agent’s helpful behavior was revealed, participants made moral character judgments of the agent ($0 = extremely\ bad$ to $50 = neither\ bad\ nor\ good$ to $100 = extremely\ good$). We adopted 0 to 100 scales to allow participants to make finer-grained distinctions than was possible in prior work³. We also collected pre-outcome ratings of how likely the agent was to help, and what proportion of the population would help if in the same situation. However, neither of these ratings was the focus of this paper.

Importantly, in contexts where agents did not have to consider whether to help one of two potential beneficiaries (i.e., “No Choice” conditions), participants made only one obligation judgment. However, in conditions where agents considered whether to help one of two potential beneficiaries (i.e., “Choice” conditions), participants made two obligations judgments (one judgment about the target agent’s obligation to help each potential beneficiary). When conducting analyses on obligation judgments in the latter conditions, ratings were averaged across the potential beneficiaries of the same relation to the target agent, which allowed us to conduct all pre-registered tests. For example, when participants read a story in which an agent ultimately helped a stranger instead of a cousin, and a separate story in which an agent ultimately helped a cousin instead of a stranger, participants’ two stranger obligation judgments were averaged into one stranger obligation judgment, whereas their two cousin obligation judgments were averaged into one cousin obligation judgment. This indexing did not apply to moral character judgments because participants judged the agent’s character only *after* they knew who the agent had helped.

5.1.5. Moral Values Measures

After participants completed the experimental task, they responded to subscales of three separate moral values questionnaires. Participants responded to the “family values” subscales of the Morality-as-

² Although we pre-registered data collection via Amazon’s Mechanical Turk (MTurk), before launching, our lab decided to temporarily switch to Prolific because of recent issues with poor-quality data on MTurk.

³ A pilot study (in non-dilemma contexts only) was conducted with obligation judgments as “yes/no” and as a graded Likert-scale, and moral character judgments as a graded Likert-scale. Results were qualitatively similar across measurement type (see SOM).

Table 1
Example story and its experimental variants.

	Distant	Close
No Choice	<p>Pre-Outcome: Bella was sitting on the subway when an injured woman on crutches boarded the train. Bella did not recognize this woman; she was a stranger. The train was already so full that many people were standing because all of the seats were taken. For the injured woman to sit down, someone needed to help by standing up and giving her their seat.</p> <p>Post-Outcome: Bella decided to stand up and offer her seat to the injured stranger.</p>	<p>Pre-Outcome: Bella was sitting on the subway when an injured woman on crutches boarded the train. Bella realized that she recognized this woman; she was Bella's sister whom Bella had not seen or spoken to in years. The train was already so full that many people were standing because all of the seats were taken. For the injured woman, Bella's sister, to sit down, someone needed to help by standing up and giving her their seat.</p> <p>Post-Outcome: Bella decided to get up and offer her seat to her injured sister.</p>
Choice	<p>Pre-Outcome: Bella was sitting on the subway when two injured women on crutches boarded the train at the same stop. Bella did not recognize one of the women; she was a stranger. However, Bella recognized the other woman as her sister whom she had not seen or spoken to in years. The train was already so full that many people were standing because all of the seats were taken. For either of the injured women to sit down, someone needed to help by standing up and giving them their seat.</p> <p>Post-Outcome: Bella decided to stand up and offer her seat to the injured stranger instead of her injured sister.</p>	<p>Pre-Outcome: Bella was sitting on the subway when two injured women on crutches boarded the train at the same stop. Bella did not recognize one of the women; she was a stranger. However, Bella recognized the other woman as her sister whom she had not seen or spoken to in years. The train was already so full that many people were standing because all of the seats were taken. For either of the injured women to sit down, someone needed to help by standing up and giving them their seat.</p> <p>Post-Outcome: Bella decided to stand up and offer her seat to her injured sister instead of the injured stranger.</p>

Note: Underlining signifies phrases that changed between “Distant” and “Close” conditions. For Experiment 1, “Distant” and “Close” changed meaning depending on random assignment to different between-participants conditions (i.e., “Distant” was not always Stranger, and “Close” was not always Sibling; see Method section). Importantly, participants never saw the same story across conditions; rather, stories were randomly assigned to each condition for each participant. See SOM for all stories and their manipulations across experiments (Importantly, not all scenarios were structured such that the agent had a resource that could only be used to help one potential beneficiary. For example, in one scenario, an agent simply decided to help one of two new neighbors move furniture into their new apartment, without it being clear that the agent only had limited time to help that day (see the “Moving Furniture” scenario in the SOM).

Cooperation Questionnaire (MAC-Q; Curry, Chesters, & Van Lissa, 2019), the “ingroup loyalty” subscales of the Moral Foundations Questionnaire (MFQ-30; Graham et al., 2011), and the “impartial beneficence” subscale of the Oxford Utilitarianism Scale (OUS; Kahane et al., 2018).

The subscales from the MAC-Q and the MFQ-30 both contain further subscales within them, one referring to moral *judgment* and the other referring to moral *relevance*. In this experiment, consistent with our pre-registration, we averaged judgment and relevance subscales together within each questionnaire, resulting in single scores for MAC-Q’s “family values” ($\alpha = 0.92$ for each between-participants dataset), and MFQ-30’s “ingroup loyalty” ($\alpha = 0.73\text{--}0.80$ for each between-participants dataset). All reported results are qualitatively similar when using individual judgment and relevance subscales (see SOM). The OUS’s “impartial beneficence” subscale ($\alpha = 0.74\text{--}0.78$ for each between-participants dataset) does not have constituent subscales.

5.1.6. *Statistical Power*

As reported in our pre-registration, we aimed for at least 200 participants per between-participants condition. This determination was based on an internal meta-analysis of the mean differences described in McManus et al. (2020). Each of the final sample sizes ($N_{Stranger/Sibling} = 203$; $N_{Stranger/Cousin} = 203$; $N_{Cousin/Sibling} = 205$) yielded at least 80% power to detect within-participant differences of $d_z = 0.20$, and correlations of $r = 0.20$, assuming two-tailed tests at an alpha level = 0.05 (Faul et al., 2007).

5.2. *Hypotheses*

Here, because we had many specific hypotheses (and we were more confident in some hypotheses than others), we communicate only our general hypotheses. However, interested readers can view specific hypotheses on our OSF pre-registration (https://osf.io/6q4kr/?view_only=8033c65c9dfc41bcb528fe94409e4e8):

- (1) Agents who can help a genetically distant other will be judged as having less of an obligation than agents who can help a genetically close other. Similarly, agents who can help either a distant other or a close other will be judged as having less of an obligation to help the distant other.
- (2) Agents who ultimately help a genetically distant other will be judged as more morally good than agents who help a genetically close other, but agents who help a more distant other instead of a closer other will be judged as less morally good than agents who help a closer other instead of a more distant other.
- (3) Obligation difference scores (i.e., “distant other” minus “close other” ratings) will be correlated with moral character difference scores; consistent with our mean difference predictions, the direction of these correlations will differ across our “Choice Context” factor.
- (4) Endorsement of family values and ingroup loyalty will be uncorrelated with obligation judgments to help strangers, but positively correlated with obligation judgments to help genetic relatives. Endorsement of these same values will also be correlated with obligation difference scores (i.e., higher endorsement = more discrimination). On the other hand, endorsement of impartial beneficence will be positively correlated with obligation judgments to help strangers. Endorsement of this same value will also be correlated with obligation difference scores (i.e., higher endorsement = less discrimination).

5.3. *Results*

Analyses will be reported for each measure within each dataset. As a reminder, by “dataset,” we mean an independent sample of participants who completed the fully within-participants design under slightly

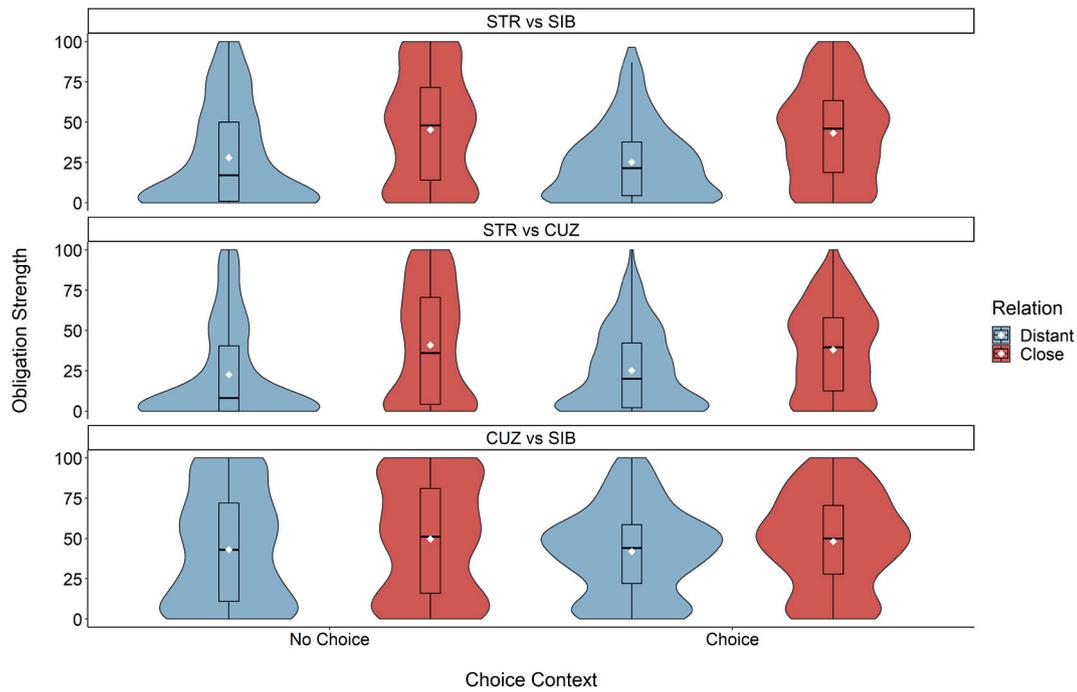


Fig. 1. Experiment 1: Violin plot of judgments of agent's strength of obligation to help as a function of how closely related the agent was to the potential beneficiary and whether another potential beneficiary was available to help instead (0 = none at all to 100 = a great deal). Different rows correspond to different datasets. In each violin, solid horizontal lines in boxplots represent medians, whereas diamonds represent means.

different conditions (i.e., varying levels of relatedness between beneficiaries).

5.3.1. Obligation Strength

In No Choice conditions, agents who could help a stranger were judged as less obligated to help than agents who could help a sibling, $t(202) = -6.16, p < .001, d_z = -0.43 [-0.57, -0.29], d_{av} = -0.54 [-0.73, -0.36]$; agents who could help a stranger were judged as less obligated to help than agents who could help a cousin, $t(202) = -6.28, p < .001, d_z = -0.44 [-0.58, -0.30], d_{av} = -0.57 [-0.76, -0.38]$; and agents who could help a cousin were judged as less obligated to help than agents who could help a sibling, $t(204) = -2.49, p = .014, d_z = -0.17 [-0.31, -0.04], d_{av} = -0.19 [-0.34, -0.04]$. In Choice conditions, agents were judged as less obligated to help a stranger than a sibling, $t(202) = -11.48, p < .001, d_z = -0.81 [-0.96, -0.65], d_{av} = -0.67 [-0.80, -0.55]$; agents were judged as less obligated to help a stranger than a cousin, $t(202) = -10.25, p < .001, d_z = -0.72 [-0.87, -0.56], d_{av} = -0.49 [-0.59, -0.39]$; and agents were judged as less obligated to help a cousin than a sibling, $t(204) = -7.13, p < .001, d_z = -0.50 [-0.64, -0.35], d_{av} = -0.22 [-0.28, -0.16]$. See Fig. 1 for judgments plotted by dataset and condition, and Table 2 for more detailed statistics.

5.3.2. Moral Character

When agents helped strangers and siblings, a 2×2 within-subjects ANOVA revealed an interaction pattern on moral character judgments, $F(1,202) = 44.26, p < .001, \eta_p^2 = 0.18$. This pattern replicated when agents helped strangers and cousins, $F(1, 202) = 13.14, p < .001, \eta_p^2 = 0.06$, as well as when agents helped cousins and siblings, $F(1, 204) = 21.49, p < .001, \eta_p^2 = 0.10$. Specifically, in No Choice conditions, agents who helped a stranger were judged as more morally good than agents who helped a sibling, $t(202) = 2.69, p = .008, d_z = 0.19 [0.05, 0.33], d_{av} = 0.19 [0.05, 0.32]$; agents who helped a stranger were judged as more morally good than agents who helped a cousin, $t(202) = 2.77, p = .006, d_z = 0.19 [0.06, 0.33], d_{av} = 0.18 [0.05, 0.30]$; however, agents who helped a cousin were judged no differently from agents who helped a

sibling, $t(204) = 0.35, p = .730, d_z = 0.02 [-0.11, 0.16], d_{av} = 0.02 [-0.10, 0.15]$. Although this last test was unable to directly support the null hypothesis, we note that the point estimates are closer to zero than they are to very small effects that some researchers may consider as theoretically meaningful (i.e., $|d_z/d_{av}| = 0.10$). In Choice conditions, agents who helped a stranger instead of a sibling were judged as less morally good than agents who did the opposite, $t(202) = -5.90, p < .001, d_z = -0.41 [-0.55, -0.28], d_{av} = -0.54 [-0.73, -0.35]$; agents who helped a stranger instead of a cousin were judged as less morally good than agents who did the opposite, $t(202) = -2.41, p = .017, d_z = -0.17 [-0.31, -0.03], d_{av} = -0.23 [-0.41, -0.04]$; and agents who helped a cousin instead of a sibling were judged as less morally good than agents who did the opposite, $t(204) = -5.12, p < .001, d_z = -0.36 [-0.50, -0.22], d_{av} = -0.43 [-0.60, -0.26]$. See Fig. 2 for judgments plotted by dataset and condition, and Table 2 for more detailed statistics.

5.3.3. Moral Character ~ Obligation strength

We note that for testing the relationship between obligation judgments and moral character judgments, we pre-registered hypotheses only about difference score correlations in No Choice conditions; therefore, we caution readers to take this into account when interpreting other results. However, we also note that replications across datasets should be considered as strong evidence for any reported effect, regardless of pre-registration.

In No Choice conditions, obligation difference scores were consistently uncorrelated with moral character difference scores, contrary to our pre-registered hypotheses. However, in Choice conditions, obligation difference scores were consistently positively correlated with moral character difference scores. These relationships held across all datasets. See Fig. 3 for difference score relationships plotted by dataset, and Table 3 for difference score statistics.

5.3.4. Obligation Strength ~ Moral Values

Across No Choice and Choice conditions, endorsement of family values and ingroup loyalty were consistently more strongly positively

Table 2
Experiment 1 judgments by relation and choice context (split by dataset).

	Distant	Close	<i>p</i>	<i>r</i>	<i>d_z</i>	<i>d_{av}</i>
STR vs SIB						
<i>No Choice</i>						
Oblig	28.02 (30.19)	45.34 (33.41)	< .001	.21	-0.43 [-0.57, -0.29]	-0.54 [-0.73, -0.36]
Moral	83.85 (14.53)	81.07 (15.46)	.008	.52	0.19 [0.05, 0.33]	0.19 [0.05, 0.32]
<i>Choice</i>						
Oblig	25.27 (22.96)	43.29 (28.87)	< .001	.65	-0.81 [-0.96, -0.65]	-0.67 [-0.80, -0.55]
Moral	58.92 (19.48)	68.79 (16.89)	< .001	.15	-0.41 [-0.55, -0.28]	-0.54 [-0.73, -0.35]
STR vs CUZ						
<i>No Choice</i>						
Oblig	22.62 (29.18)	40.87 (34.51)	< .001	.16	-0.44 [-0.58, -0.30]	-0.57 [-0.76, -0.38]
Moral	83.25 (15.56)	80.45 (16.13)	.006	.59	0.19 [0.06, 0.33]	0.18 [0.05, 0.30]
<i>Choice</i>						
Oblig	25.19 (24.22)	38.04 (27.40)	< .001	.77	-0.72 [-0.87, -0.56]	-0.49 [-0.59, -0.39]
Moral	63.90 (19.45)	68.08 (17.71)	.017	.11	-0.17 [-0.31, -0.03]	-0.23 [-0.41, -0.04]
CUZ vs SIB						
<i>No Choice</i>						
Oblig	43.17 (34.07)	49.69 (34.84)	.014	.41	-0.17 [-0.31, -0.03]	-0.19 [-0.34, -0.04]
Moral	81.11 (15.48)	80.76 (16.34)	.730	.59	0.02 [-0.11, 0.16]	0.02 [-0.10, 0.15]
<i>Choice</i>						
Oblig	41.90 (26.81)	48.14 (28.86)	.003	.90	-0.50 [-0.64, -0.35]	-0.22 [-0.28, -0.16]
Moral	60.08 (17.71)	67.53 (17.08)	< .001	.28	-0.35 [-0.50, -0.22]	-0.43 [-0.60, -0.26]

Note. STR = Stranger; CUZ = Cousin; SIB = Sibling. Different rows correspond to different datasets. Means and (SDs) are reported in the first two columns. Correlation coefficients, *r*, are reported so *d_z* and *d_{av}* effect sizes can be calculated directly from descriptive statistics (see Lakens, 2013). Brackets underneath effect sizes denote 95% confidence intervals.

correlated with obligations toward genetically close (versus genetically distant) others. Endorsement of family values also tended to be more strongly correlated with obligations toward genetic relatives than was ingroup loyalty. Moreover, endorsement of family values tended to be more strongly correlated with obligation difference scores than endorsement of ingroup loyalty was. That is, the more participants endorsed family values, the more they distinguished between obligation judgments to help genetically closer and genetically more distant others. Within the dataset that compared obligation judgments to members of the same family (i.e., CUZ vs SIB), we note that although not all family values/obligation correlations were statistically different from ingroup loyalty/obligation correlations, this was the case for 67% of the family values/obligation correlations, suggesting the unique predictive power of family values relative to ingroup loyalty. We conducted these comparisons only within the CUZ vs SIB dataset because the influence of family values and ingroup loyalty on obligation judgments are

potentially confounded in the other datasets due to strangers likely being categorized as outgroup members. On the other hand, across No Choice and Choice conditions, endorsement of impartial beneficence was consistently positively correlated with obligation judgments to help strangers, and sometimes positively correlated with obligation judgments to help genetic relatives. However, endorsement of impartial beneficence was not consistently correlated with obligation difference scores. That is, participants who scored high versus low in impartial beneficence were no more likely to distinguish between obligations toward genetically closer and genetically more distant others. See Table 4 to compare correlations across datasets and moral values questionnaires, and see SOM for relationships plotted by dataset.

6. Experiment 2

Experiment 2 was conducted to replicate, address limitations of, and further extend Experiment 1. First, by describing agents as “stranger-like” in Experiment 1, we intended to experimentally control for other dimensions along which differently related agents might differ from one another. However, participants may have still made different inferences across differently related agents. For example, participants may have inferred that a sibling, compared to a cousin, has more frequently helped the target agent in the past, which in turn could have accounted for differences in perceptions of obligations⁴. To test this possibility, in Experiment 2, we systematically varied whether agents were described as “stranger-like” versus “friend-like” (e.g., agents grew up in the same household, always got along, still regularly help one another when in need, and continue to behave as if they are best friends). We therefore omitted the “stranger” category and focused only on cousin and sibling relationships. Second, we directly measured perceptions of relatedness, as well as inferences of past and future interactions between helpers and beneficiaries, to explore whether objective relatedness impacts obligation judgments via perceived relatedness versus inferences about past and future social interactions.

6.1. Method

6.1.1. Participants

All participants were U.S. residents recruited and compensated via CloudResearch (Chandler, Rosenzweig, Moss, Robinson, & Litman, 2019). We decided a priori to collect data from 735 participants to obtain 660 analyzable responses. This sample size was chosen to yield 330 responses per between-participants condition (see Statistical Power for further justification). After excluding duplicate responses and participants who failed an attention check, this yielded the desired sample size for each between-participants condition ($N_{Total} = 699$, 53.5% female, $M_{Age} = 40.51$ years).

6.1.2. Design

This experiment used a 2 (Relation: Distant vs Close) x 2 (Choice Context: No Choice vs Choice) x 2 (Nature of Relationship Between Agents) mixed design, in which “Relation” and “Choice Context” were manipulated within-participants, whereas “Nature of Relationship Between Agents” was manipulated between participants (see below). Specifically, participants were asked to make judgments of an agent who helped (1) a cousin, (2) a sibling, (3) a cousin *instead of* a sibling, and (4) a sibling *instead of* a cousin. Additionally, participants were randomly assigned to one of two conditions which varied how the relationship between agents was described. One group of participants read stories involving agents being described as “stranger-like” (as in Experiment 1); another group of participants read the same stories but instead learned that agents were “friend-like.” See Table 5 for an example scenario’s “friend-like” descriptions and its variants across Relation and Choice

⁴ We thank the editor and two anonymous reviewers for raising this concern.

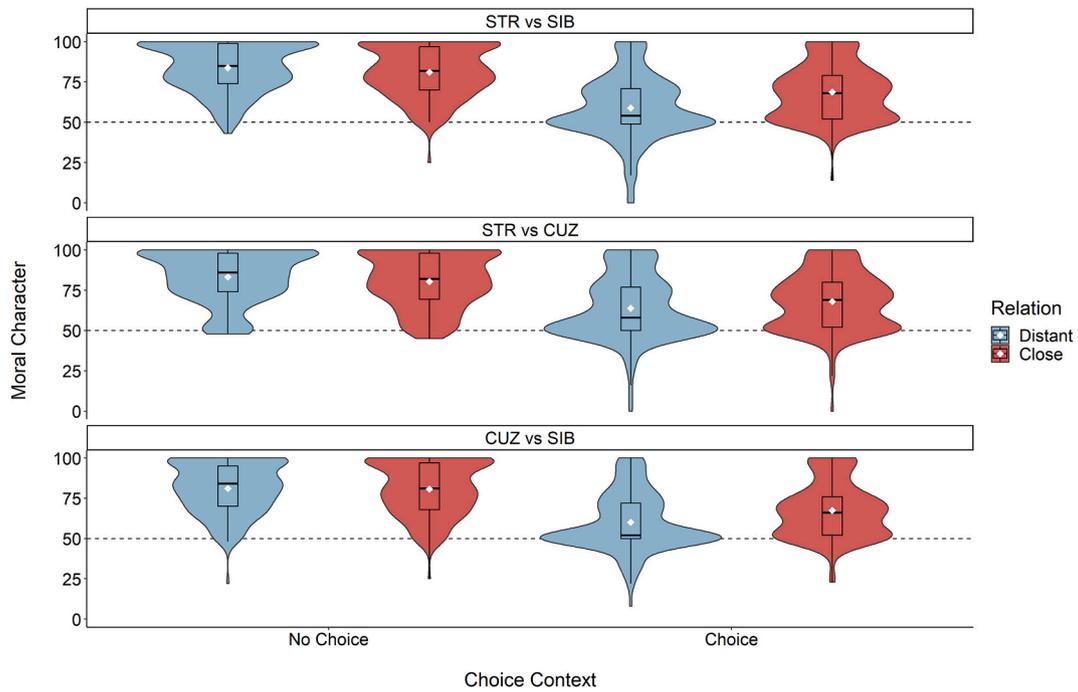


Fig. 2. Experiment 1: Violin plot of judgments of agent’s moral character as a function of how closely related the agent was to their beneficiary and whether another potential beneficiary was available to help instead (0 = extremely bad to 50 = neither bad nor good to 100 = extremely good). Different rows correspond to different datasets. Dashed horizontal lines depicts the midpoint of the scale. In each violin, solid horizontal lines in boxplots represent medians, whereas diamonds represent means.

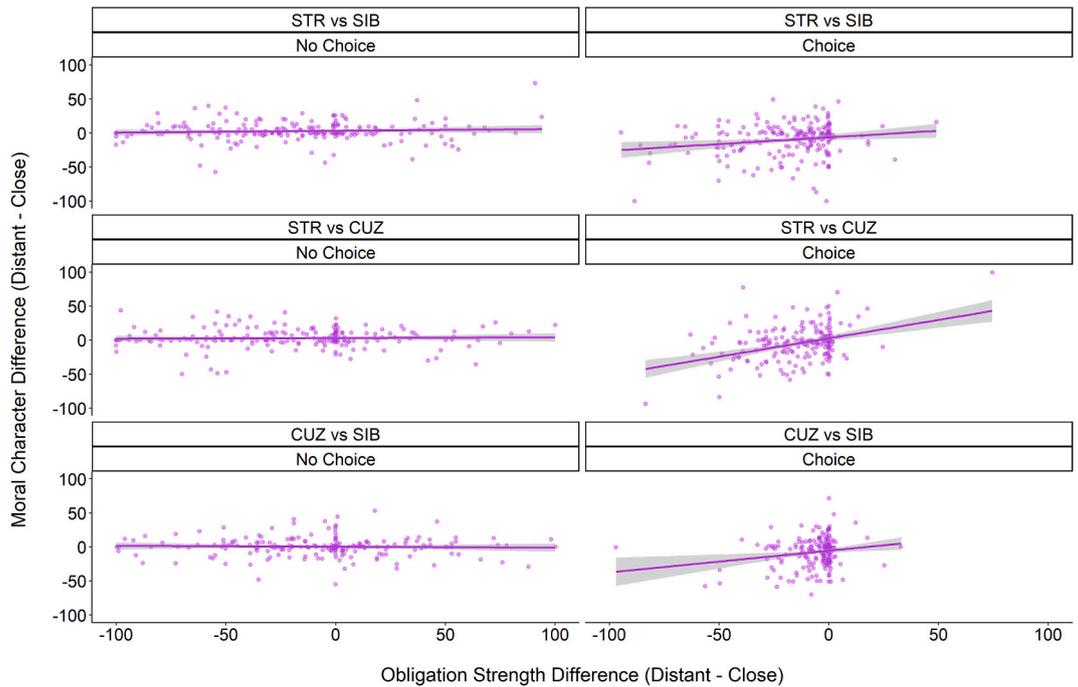


Fig. 3. Experiment 1: Scatterplots of moral character difference scores as a function of obligation difference scores. Differences for each variable were calculated by subtracting “close other” judgments from “distant other” judgments. Different rows correspond to different datasets; different columns correspond to different levels of Choice Context factor (left = No Choice, right = Choice). Solid shaded lines show OLS regression best-fits.

Table 3

Experiment 1 correlations between obligation judgments and moral character judgments by choice context (split by dataset).

Difference score correlations		
	<i>r</i>	
STR vs SIB		
No Choice	.07	
Choice	.18	**
STR vs CUZ		
No Choice	.03	
Choice	.39	***
CUZ vs SIB		
No Choice	-.03	
Choice	.19	**

Note: STR = Stranger; CUZ = Cousin; SIB = Sibling. Different row headers (e.g., STR vs SIB and STR vs CUZ) correspond to different datasets. Difference score is calculated by subtracting obligation judgments to help genetically closer others from obligation judgments to help more genetically distant others (i.e., Distant – Close). *** $p \leq .001$; ** $p \leq .010$; * $p \leq .050$; † $p \leq .100$.

Context factors.

6.1.3. Procedure

Experiment 2’s procedure was identical to Experiment 1 except for two minor additions. First, before participants started the experimental task, they were given a definition of genetic relatedness along with an example of how different two people were in terms of genetic relatedness⁵; participants were told to keep this definition in mind when answering the new relatedness measure (see below). Second, after participants finished the experimental task, they filled out all subscales (rather than only a specific set of subscales) from three separate moral values questionnaires.

6.1.4. Experimental Task Measures

Participants made the same obligation and moral character judgments as in Experiment 1. However, we also collected pre-outcome ratings of how closely genetically related agents were (0% related to 100% related), how socially/emotionally close the agents were (0 = not at all to 100 = very much so), perceived frequency of past help and future help (i.e., how often the potential beneficiary had helped/will help the potential helper), and perceived frequency of past and future interactions generally (0 = not at all to 100 = very much).

As in Experiment 1, in contexts where agents did not have to consider whether to help one of two potential beneficiaries (i.e., “No Choice” conditions), participants made only one pre-outcome judgment (e.g., one obligation judgment), as there was only one potential beneficiary. However, in conditions where agents considered whether to help one of two potential beneficiaries (i.e., “Choice” conditions), participants made two pre-outcome judgments (e.g., one judgment about the target agent’s obligation to help each potential beneficiary). When conducting analyses on pre-outcome judgments in the latter conditions, we applied the same indexing as in Experiment 1 so that there was only one judgment for each pre-outcome measure for potential beneficiaries of the same

⁵ This question and example were included due to another pre-registered experiment that was conducted to investigate the genetic relatedness hypothesis by systematically varying relationship categories (e.g., half-cousin vs cousin). As explained in more detail in this experiment’s our RMarkdown file (see “BadE Analysis Script.nb.html” on our OSF page), participants did not seem to recognize some of these relationship categories, nor did they understand the meaning of genetic relatedness in this context, leading to inconsistent results across different relationship categories.

Table 4

Experiment 1 correlations between moral values and obligation judgments by relation and choice context (split by dataset).

	MAC Family Values		MFT Ingroup Loyalty		OUS Impartial Beneficence	
	<i>r</i>		<i>r</i>		<i>r</i>	
STR vs SIB						
No Choice						
Distant	.23	***	.29	***	.15	*
Close	.35	***	.32	***	.08	
Difference	-.12	†	-.05		.05	
Choice						
Distant	.25	***	.22	***	.31	***
Close	.43	***	.36	***	.21	**
Difference	-.30	***	-.24	***	.04	
STR vs CUZ						
No Choice						
Distant	.06		.04		.17	*
Close	.22	**	.16	*	.24	***
Difference	-.14	†	-.10		-.08	
Choice						
Distant	.19	**	.22	***	.42	***
Close	.36	***	.37	***	.46	***
Difference	-.30	***	-.27	***	-.12	†
CUZ vs SIB						
No Choice						
Distant	.26	***	.17	*	.23	***
Close	.48	***	.25	***	.38	***
Difference	-.21	**	-.07		-.14	*
Choice						
Distant	.32	***	.16	*	.36	***
Close	.40	***	.24	***	.34	***
Difference	-.23	***	-.20	**	-.00	

Note: STR = Stranger; CUZ = Cousin; SIB = Sibling. Different row headers (e.g., STR vs SIB and STR vs CUZ) correspond to different datasets. “Difference” is calculated by subtracting obligation judgments to help genetically closer others from obligation judgments to help more genetically distant others (i.e., Distant – Close). Plus signs (+) indicate which family values/obligation correlations are statistically different from the same condition’s ingroup loyalty/obligation correlations. As described in the results section, these comparisons were conducted only within the CUZ vs SIB because the influence of family values and ingroup loyalty on obligation judgments are potentially confounded in the other datasets. *** $p \leq .001$; ** $p \leq .010$; * $p \leq .050$; † $p \leq .100$.

relation.

6.1.5. Moral Values Measures

After participants completed the experimental task, they responded to all subscales of the Morality-as-Cooperation Questionnaire (MAC-Q; Curry, Chesters, & Van Lissa, 2019), the Moral Foundations Questionnaire (MFQ-30; Graham et al., 2011), and the Oxford Utilitarianism Scale (OUS; Kahane et al., 2018). However, as in Experiment 1, we were primarily interested in the “family values” subscale ($\alpha = 0.89\text{--}0.90$ for each between-participants dataset), the “ingroup loyalty” subscale ($\alpha = 0.74\text{--}0.79$ for each between-participants dataset), and the “impartial beneficence” subscale ($\alpha = 0.72\text{--}0.78$ for each between-participants dataset). Because we did not find any differences between judgment and relevance subscales of the MAC-Q or MFQ-30 in Experiment 1, we conducted analyses only on their composites in Experiment 2. See our OSF page for analyses of all questionnaires’ other subscales.

6.1.6. Statistical Power

Each of the final sample sizes ($N_{\text{Stranger-Like}} = 354$; $N_{\text{Friend-Like}} = 345$) yielded at least 95% power to detect within-participant differences of d_z

Table 5
Example story's opening descriptions for only "Friend-Like" stimuli.

	Distant	Close
No Choice	Bella and Nichole (Bella's cousin) are like best friends. They grew up in the same household and are only a few years apart in age. As adults, they still talk on the phone every couple days and always help each other when possible. They even happen to live in the same city which allows them to get together every weekend.	Bella and Suzie (Bella's sister) are like best friends. They grew up in the same household and are only a few years apart in age. As adults, they still talk on the phone every couple days and always help each other when possible. They even happen to live in the same city which allows them to get together every weekend.
Choice	Bella, Nichole (Bella's cousin), and Suzie (Bella's sister) are like best friends. They all grew up in the same household and are only a few years apart in age. As adults, they still talk on the phone every couple days and always help each other when possible. They even happen to live in the same city which allows them to get together every weekend.	Bella, Nichole (Bella's cousin), and Suzie (Bella's sister) are like best friends. They all grew up in the same household and are only a few years apart in age. As adults, they still talk on the phone every couple days and always help each other when possible. They even happen to live in the same city which allows them to get together every weekend.

Note: We include only the opening sentences of the story's variants here, as this was the primary difference between "Stranger-Like" and "Friend-Like" stimuli. Therefore, these opening descriptions were not used in the "Stranger-Like" stimuli. Importantly, although all stories within the "Friend-Like" stimuli communicated similar friend-like histories, each story differed in their exact details. See SOM for all full stories and their manipulations across experiments.

= 0.20, and correlations of $r = 0.20$, assuming two-tailed tests at an alpha level = 0.05.

6.2. Hypotheses

Here, as in Experiment 1, because we had many specific hypotheses, we communicate only our general hypotheses. Interested readers can view specific hypotheses on our OSF pre-registration (https://osf.io/uxysz/?view_only=cad9b09ce5d1494eb023a604d9f6feb3).

- (1) Agents who can help a genetically distant other will be judged as having less of an obligation than agents who can help a genetically close other. Similarly, agents who can help either a distant other or a close other will be judged as having less of an obligation to help the distant other.
- (2) Agents who help a more genetically distant other instead of a genetically closer other will be judged as less morally good than agents who help a genetically closer other instead of a more genetically distant other.
- (3) Obligation difference scores (i.e., "distant other" minus "close other" ratings) will be positively correlated with moral character difference scores only in our "Choice" conditions.
- (4) Endorsement of family values and ingroup loyalty will be positively correlated with obligation judgments to help genetic relatives. Endorsement of these same values will also be correlated with obligation difference scores (i.e., higher endorsement = more discrimination).

6.3. Manipulation Check

6.3.1. Relatedness

To ensure that participants were indeed distinguishing between cousins and siblings in terms of *how closely related* each relative was to the target agent, as a manipulation check, we conducted paired-samples *t*-tests within each level of the "Choice Context" factor within each dataset.

As expected, In No Choice conditions, agents were judged as less closely related to stranger-like cousins than stranger-like siblings, $t(353) = -34.66, p < .001, d_z = -1.84 [-2.01, -1.67], d_{av} = -2.00 [-2.20, -1.80]$; similarly, agents were judged as less closely related to friend-like cousins than friend-like siblings, $t(344) = -33.58, p < .001, d_z = -1.81 [-1.98, -1.64], d_{av} = -1.92 [-2.11, -1.73]$. Likewise, in Choice conditions, agents were judged as less closely related to stranger-like cousins than stranger-like siblings, $t(353) = -35.61, p < .001, d_z = -1.89 [-2.07, -1.72], d_{av} = -1.98 [-2.17, -1.79]$; and agents were judged as less closely related to friend-like cousins than friend-like siblings, $t(344) = -33.77, p < .001, d_z = -1.82 [-1.99, -1.65], d_{av} = -1.87 [-2.05, -1.69]$. See SOM for plots and more detailed statistics.

6.4. Primary Results

Analyses will be reported for each primary measure (i.e., obligation and moral character) within each dataset. As a reminder, by "dataset," we mean an independent sample of participants who completed the fully within-participants design under slightly different conditions (i.e., with agents being described as "stranger-like" versus "friend-like").

6.4.1. Obligation Strength

In No Choice conditions, agents who could help a stranger-like cousin were judged as less obligated to help than agents who could help a stranger-like sibling, $t(353) = -4.07, p < .001, d_z = -0.22 [-0.32, -0.11], d_{av} = -0.24 [-0.36, -0.12]$; and agents who could help a friend-like cousin were judged as less obligated to help than agents who could help a friend-like sibling, $t(344) = -2.88, p = .004, d_z = -0.16 [-0.26, -0.05], d_{av} = -0.15 [-0.25, -0.05]$. Similarly, in Choice conditions, agents were judged as less obligated to help a stranger-like cousin than a

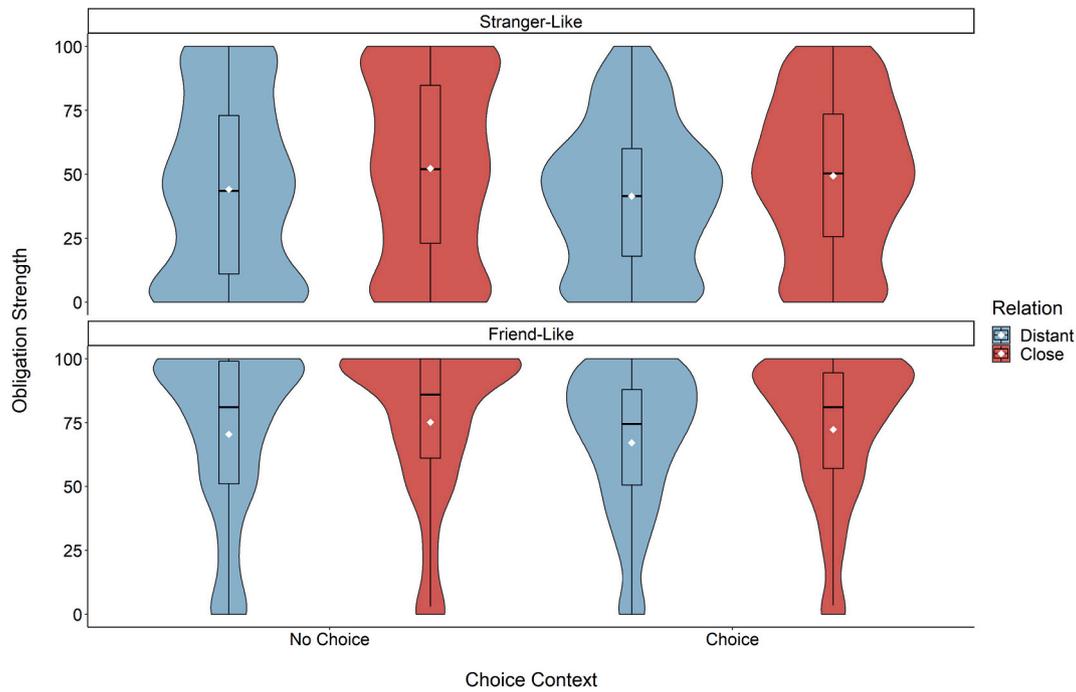


Fig. 4. Experiment 2: Violin plot of judgments of agent’s strength of obligation to help as a function of how closely related the agent was to the potential beneficiary and whether another potential beneficiary was available to help instead (0 = none at all to 100 = a great deal). Distant = CUZ; Close = SIB. Different rows correspond to different datasets. In each violin, solid horizontal lines in boxplots represent medians, whereas diamonds represent means.

stranger-like sibling, $t(353) = -12.34, p < .001, d_z = -0.66 [-0.77, -0.54], d_{av} = -0.27 [-0.31, -0.23]$; and agents were judged as less obligated to help a friend-like cousin than a friend-like sibling, $t(344) = -10.58, p < .001, d_z = -0.57 [-0.68, -0.46], d_{av} = -0.19 [-0.22, -0.15]$. All results are consistent with those found in Experiment 1. Interestingly, obligation judgments were overall lower for stranger-like relatives than friend-like relatives. See Fig. 4 for judgments plotted by dataset and condition, and Table 6 for more detailed statistics⁶.

6.4.2. Moral Character

When agents helped stranger-like cousins and siblings, a 2×2 within-subjects ANOVA revealed an interaction pattern on moral character judgments, $F(1, 353) = 64.77, p < .001, \eta_p^2 = 0.16$. This pattern replicated when agents helped friend-like cousins and siblings, $F(1, 344) = 17.74, p < .001, \eta_p^2 = 0.05$. Specifically, in No Choice conditions, agents who helped a stranger-like cousin were judged as no more morally good than agents who helped a stranger-like sibling, $t(353) = 0.46, p = .649, d_z = 0.02 [-0.08, 0.13], d_{av} = 0.02 [-0.06, 0.10]$; similarly, agents who helped a friend-like cousin were judged as no more morally good than agents who helped a friend-like sibling, $t(344) = 0.64, p = .521, d_z = 0.03 [-0.07, 0.14], d_{av} = 0.03 [-0.06, 0.11]$. Although these tests were unable to directly support the null hypothesis, we note that the point estimates are similar to those from Experiment 1 for the same simple effect (i.e., closer to zero than they are to very small effects that some researchers may consider as theoretically meaningful). However, consistent with Experiment 1, in Choice conditions, agents who helped a stranger-like cousin instead of a stranger-like sibling were judged as less morally good than agents who did the opposite, $t(353) = -8.98, p < .001, d_z = -0.48 [-0.59, -0.37], d_{av} = -0.56 [-0.69, -0.43]$; similarly, agents who helped a friend-like cousin instead of a friend-like sibling were judged as less morally good than agents who did

Table 6
Experiment 2 judgments by relation and choice context (split by dataset).

	Distant	Close	<i>p</i>	<i>r</i>	<i>d_z</i>	<i>d_{av}</i>
Stranger-Like						
<i>No Choice</i>						
Oblig	44.14 (33.94)	52.33 (34.73)	< .001	.39	-0.22 [-0.32, -0.11]	-0.24 [-0.36, -0.12]
Moral	83.63 (16.26)	83.32 (16.72)	.649	.71	0.02 [-0.08, 0.13]	0.02 [-0.06, 0.10]
<i>Choice</i>						
Oblig	41.45 (27.93)	49.42 (30.09)	< .001	.92	-0.66 [-0.77, -0.54]	-0.27 [-0.31, -0.23]
Moral	58.33 (19.54)	69.31 (19.77)	< .001	.32	-0.48 [-0.59, -0.37]	-0.56 [-0.69, -0.43]
Friend-Like						
<i>No Choice</i>						
Oblig	70.51 (31.68)	75.16 (29.76)	.004	.53	-0.16 [-0.26, -0.05]	-0.15 [-0.25, -0.05]
Moral	86.32 (15.56)	85.89 (16.23)	.521	.69	0.03 [-0.07, 0.14]	0.03 [-0.06, 0.11]
<i>Choice</i>						
Oblig	67.19 (27.46)	72.36 (27.62)	< .001	.95	-0.57 [-0.68, -0.46]	-0.19 [-0.22, -0.15]
Moral	62.35 (22.32)	68.51 (20.94)	< .001	.32	-0.24 [-0.35, -0.14]	-0.28 [-0.41, -0.16]

Note. Distant = CUZ; Close = SIB. Different rows correspond to different datasets. Means and (SDs) are reported in the first two columns. Correlation coefficients, *r*, are reported so *d_z* and *d_{av}* effect sizes can be calculated directly from descriptive statistics (see Lakens, 2013). Brackets underneath effect sizes denote 95% confidence intervals.

⁶ Although we pre-registered tests using dataset as a between-subjects factor, we do not report those analyses because our intention was to run those analyses only if we detected effects in one dataset but not the other. However, interested readers can perform these analyses using our data on OSF.

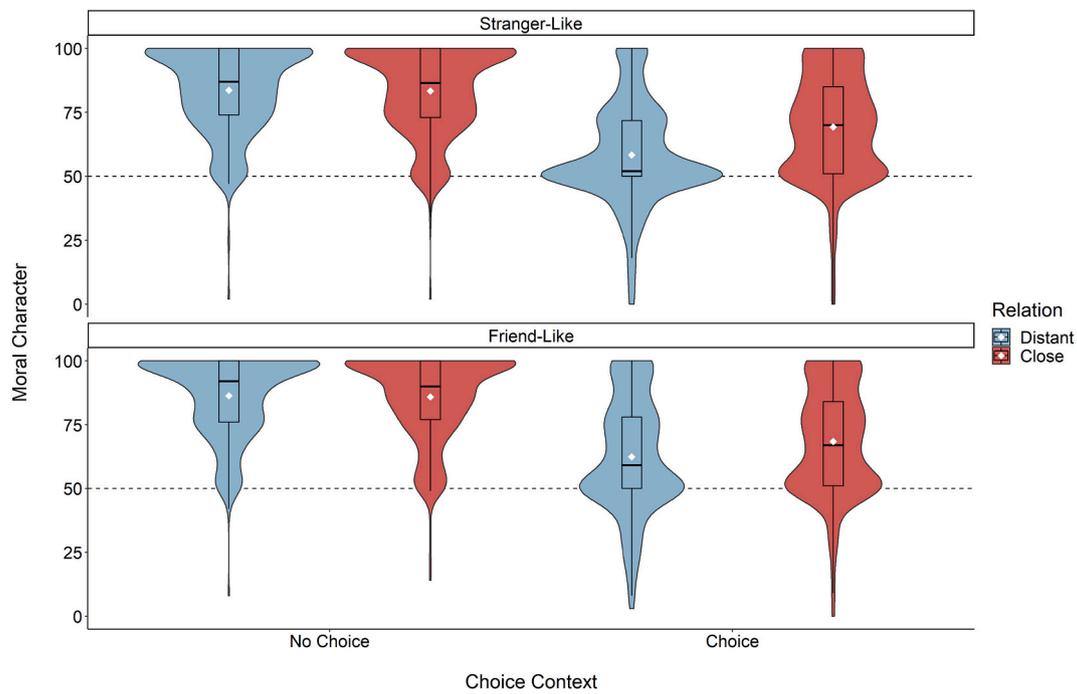


Fig. 5. Experiment 2: Violin plot of judgments of agent's moral character as a function of how closely related the agent was to their beneficiary and whether another potential beneficiary was available to help instead ($0 = \text{extremely bad}$ to $50 = \text{neither bad nor good}$ to $100 = \text{extremely good}$). Distant = CUZ; Close = SIB. Different rows correspond to different datasets. Dashed horizontal lines depicts the midpoint of the scale. In each violin, solid horizontal lines in boxplots represent medians, whereas diamonds represent means.

the opposite, $t(344) = -4.53, p < .001, d_z = -0.24 [-0.35, -0.14], d_{av} = -0.28 [-0.41, -0.16]$. See Fig. 5 for judgments plotted by dataset and condition, and Table 6 for more detailed statistics.

6.4.3. Moral Character ~ Obligation strength

In No Choice conditions, obligation difference scores were uncorrelated with moral character difference scores. However, in Choice conditions, obligation difference scores were positively correlated with moral character difference scores. These relationships held across datasets and are consistent with results from Experiment 1. See Fig. 6 for difference score relationships plotted by dataset, and Table 7 for difference score statistics.

6.4.4. Obligation Strength ~ Moral values

Across No Choice and Choice conditions, endorsement of family values and ingroup loyalty were consistently more strongly positively correlated with obligations toward siblings (versus cousins). Endorsement of family values also tended to be more strongly correlated with obligations toward relatives generally than was ingroup loyalty. Moreover, endorsement of family values tended to be more strongly correlated with obligation difference scores than endorsement of ingroup loyalty was. Unlike Experiment 1, though, these relationships with difference scores were only evident in the Choice conditions. We note that although not all family values/obligation correlations were statistically different from ingroup loyalty/obligation correlations, this was the case for 83% of the family values/obligation correlations; the only correlations that were not statistically different were those with obligation differences in the No Choice conditions. Across No Choice and Choice conditions, endorsement of impartial beneficence was weakly (and inconsistently) positively correlated with obligations toward relatives generally. Like Experiment 1, endorsement of impartial beneficence was uncorrelated with obligation differences. See Table 8 to compare correlations across datasets and moral values questionnaires, and see SOM for these relationships plotted by dataset.

6.5. Relatedness vs. Social Interactions Results

Here, we report whether and how perceived relatedness and social interaction inferences vary across our relatedness manipulation. First, we test whether perceived relatedness and social interaction inferences indeed differ as a function of our "Relation" manipulation. Second, we test whether and how perceived relatedness is associated with obligation judgments. Third, we test whether and how social interaction inferences are associated with obligation judgments. Last, to better understand which inferences are more strongly or more consistently associated with obligation differences, we compare relatedness/obligation associations to social interaction/obligation associations.

6.5.1. Relatedness

As noted in the Primary Results section (as a manipulation check), across contexts and datasets, participants consistently judged cousins as less related than siblings. See SOM for plots and statistics of mean differences across contexts and datasets.

6.5.2. Social Interactions

Consistent with relatedness judgments, across contexts and datasets, participants consistently judged cousins, compared to siblings, as less socially/emotionally close, as less frequently helping the potential helper in the past (and as less frequently helping the potential helper in the future), and as less frequently interacting with the potential helper in the past (and less frequently helping the potential in the future). Consistent with obligation judgments, all social interactions judgments were lower for stranger-like relatives than friend-like relatives. See SOM for plots and statistics of mean differences across contexts and datasets for each judgment.

6.5.3. Obligation ~ Relatedness

Across contexts and datasets, relatedness difference scores were not consistently correlated with obligation difference scores. See Table 9 for difference score correlation statistics.

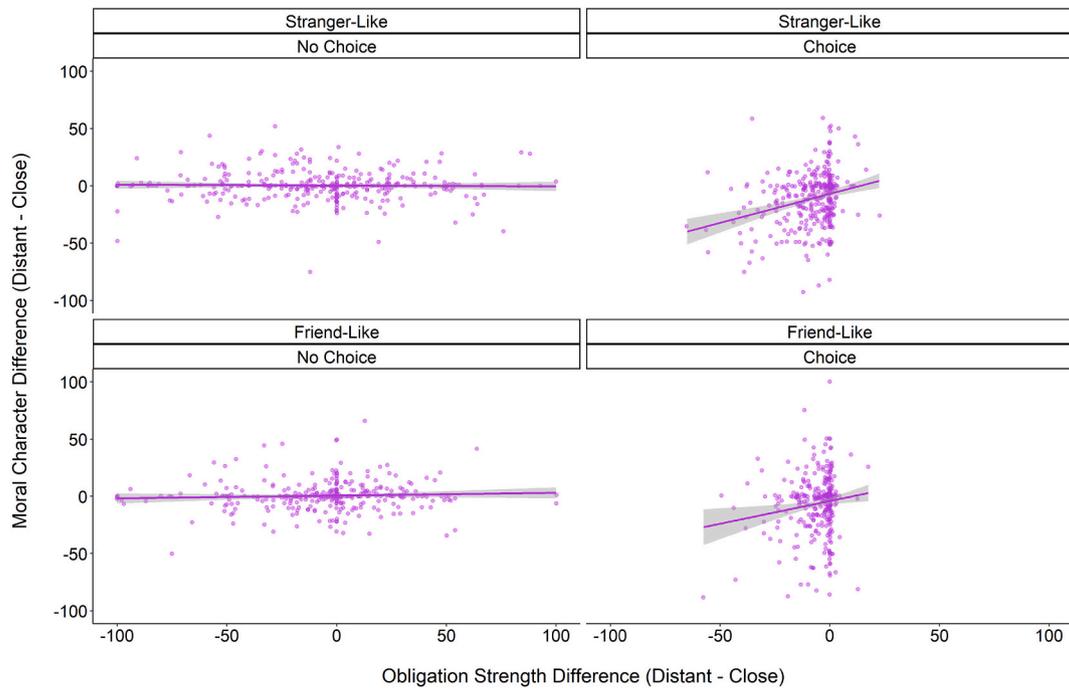


Fig. 6. Experiment 2: Scatterplots of moral character difference scores as a function of obligation difference scores. Differences for each variable were calculated by subtracting “close other” judgments from “distant other” judgments. Different rows correspond to different datasets; different columns correspond to different levels of Choice Context factor (left = No Choice, right = Choice). Solid shaded lines show OLS regression best-fits.

6.5.4. *Obligation ~ Social Interactions*

Across contexts and datasets, social interaction difference scores were consistently positively correlated with obligation difference scores. The magnitudes of these relationships were consistently stronger than the magnitudes of the relationships between relatedness differences and obligation differences. We note that although not all social interaction/obligation correlations were statistically different from relatedness/obligation correlations, this was the case for 80% of the social interaction/obligation correlations. However, we do not interpret the small percentage of non-differences as evidence in favor of the null; instead, in conjunction with the majority of comparisons being statistically different, we interpret the consistent numerical differences in magnitude to suggest that we were not statistically well-powered to detect the smallest differences (as we did not specifically power these tests in the design of this experiment). See Table 9 to compare relatedness difference scores’ relationships with obligation difference scores to social interaction difference scores’ relationships with obligation difference scores.

Table 7
Experiment 2 Correlations Between Obligation Judgments and Moral Character Judgments by Choice Context (Split by Dataset).

	Difference score correlations	
	<i>r</i>	
Stranger-Like		
No Choice	-.02	
Choice	.27	***
Friend-Like		
No Choice	.06	
Choice	.14	**

Note: Different row headers correspond to different datasets. Difference score is calculated by subtracting obligation judgments to help genetically close others from obligation judgments to help genetically distant others. *** $p \leq .001$; ** $p \leq .010$; * $p \leq .050$; † $p \leq .100$.

7. General Discussion

Ten years ago, Bloom (2011) identified a gap in our understanding and investigation of moral psychology:

Table 8
Experiment 2 correlations between moral values and obligation judgments by relation and choice context (split by dataset).

	MAC Family Values		MFQ Ingroup Loyalty		OUS Impartial Beneficence	
	<i>r</i>		<i>r</i>		<i>r</i>	
Stranger-Like						
<i>No Choice</i>						
Distant	.31	***	+	.21	***	.13 *
Close	.33	***	+	.18	***	.13 *
Difference	-.03			.02		.00
<i>Choice</i>						
Distant	.37	***	+	.27	***	.19 ***
Close	.43	***	+	.28	***	.16 **
Difference	-.20	***	+	-.07		.05
Friend-Like						
<i>No Choice</i>						
Distant	.25	***	+	.14	*	.12 *
Close	.32	***	+	.19	***	.06
Difference	-.06			-.04		.06
<i>Choice</i>						
Distant	.29	***	+	.15	**	.12 *
Close	.34	***	+	.17	**	.11 *
Difference	-.17	***	+	-.06		.02

Note: Distant = CUZ; Close = SIB. Different row headers correspond to different datasets. “Difference” is calculated by subtracting obligation judgments to help genetically close others from obligation judgments to help genetically distant others. Plus signs (+) indicate which family values/obligation correlations are statistically different from the same condition’s ingroup loyalty/obligation correlations. *** $p \leq .001$; ** $p \leq .010$; * $p \leq .050$; † $p \leq .100$.

Table 9

Experiment 2 correlations between additional pre-outcome measures (relatedness and social interactions) and obligation judgments by relation and choice context (split by dataset).

	Difference score correlations		
	<i>r</i>		
Stranger-Like			
<i>No Choice</i>			
Relatedness	.10	†	
Closeness	.22	***	
Prior Help	.21	***	
Future Help	.36	***	+
Prior Interactions	.16	**	
Future Interactions	.19	***	
<i>Choice</i>			
Relatedness	.10	†	
Closeness	.37	***	+
Prior Help	.42	***	+
Future Help	.51	***	+
Prior Interactions	.34	***	+
Future Interactions	.45	***	+
Friend-Like			
<i>No Choice</i>			
Relatedness	.02		
Closeness	.25	***	+
Prior Help	.26	***	+
Future Help	.35	***	+
Prior Interactions	.26	***	+
Future Interactions	.19	***	+
<i>Choice</i>			
Relatedness	.13	*	
Closeness	.60	***	+
Prior Help	.56	***	+
Future Help	.60	***	+
Prior Interactions	.45	***	+
Future Interactions	.54	***	+

Note: Different row headers correspond to different datasets. "Difference Score" is calculated by subtracting obligation judgments to help genetically close others from obligation judgments to help genetically distant others. Plus signs (+) indicate which social interaction/obligation correlations are statistically different from the same condition's relatedness/obligation correlations. *** $p \leq .001$; ** $p \leq .010$; * $p \leq .050$; † $p \leq .100$.

The problem is that most research in this field, including my own, focuses almost entirely on how people make sense of, judge, and respond to interactions of unrelated strangers. We have little to say about how people think of interactions that occur between parent and child, brother and sister, and closely related individuals. We also often ignore moral judgments and moral feelings that concern spouses, close friends, colleagues, allies, and compatriots.

Bloom goes on to argue that these are precisely the interactions that matter most, and, in turn, that the field's failure to explore them leads to the development of theories that do not capture our everyday moral psychology.

Similarly, much recent conceptual and empirical work has pointed out that, either implicitly or explicitly, the field of moral psychology has often operated as if social relationships are not important for third-person moral judgment (e.g., Earp, McLoughlin, Monrad, Clark, & Crockett, 2021, preprint; Everett et al., 2018; Hester & Gray, 2020; Law et al., 2021; Marshall et al., 2020; Marshall et al., 2021, invited revision; Schein, 2020). Indeed, many prominent theories do not make predictions about how social relationships will impact moral judgment (e.g., Gray, Young, & Waytz, 2012; Schein & Gray, 2016, 2018; Young et al., 2010; Young, Cushman, Hauser, & Saxe, 2007). Some prominent theories, however, stress the importance of social relationships to different degrees (e.g., Graham et al., 2011; Rai & Fiske, 2011).

Specifically, Relationship Regulation Theory (Rai & Fiske, 2011), which partly draws on relationship psychology (e.g., Clark & Mills, 1979), suggests that communal sharing relationships, like those with family, are characterized by providing preferential aid to those within the group. Therefore, RRT predicts that failure to uphold this relationship obligation will be judged negatively. Similarly, Morality-as-Cooperation (Curry, 2016) predicts that helping family is considered a universal moral good. The current findings support these predictions, but add greater nuance.

Contributing to the recent resurgence in research on everyday moral psychology, our findings suggest that people view relatedness between potential helpers and beneficiaries to be a key determinant in perceived strength of obligations to help. Agents were judged as having stronger obligations to help close genetic relatives (i.e., siblings) than more distant genetic relatives (i.e., cousins) and non-relatives (i.e., strangers), and agents were also judged as having stronger obligations to help distant genetic relatives than non-relatives. Importantly, however, results of our second experiment suggest that relatedness likely exerts its influence on obligation judgments through differential inferences about social interactions (e.g., how frequently each potential beneficiary has helped the potential helper in the past). Moreover, we found consistent evidence that moral values serve as additional determinants of perceived obligation strength. Specifically, endorsement of family values (Curry, Chesters, & Van Lissa, 2019; Curry, Mullins, & Whitehouse, 2019) and ingroup loyalty (Graham et al., 2011) were consistently and more strongly positively associated with obligation judgments to help family members than to help strangers, whereas endorsement of impartial beneficence (Kahane et al., 2018) was consistently positively associated with obligation judgments to help strangers. Replicating prior research (Marshall et al., 2021, invited revision; McManus et al., 2020), we also found that perceived obligations appeared to inform moral judgments. On the one hand, people judged agents who helped a stranger as more morally good than agents who helped a family member. On the other hand, people judged agents who helped a stranger instead of a family member as less morally good than agents who helped a family member instead of a stranger. In extending our understanding of these patterns, consistent correlations were found between obligation judgment differences and moral character judgment differences only in contexts where agents can make a choice of whom to help, revealing when and how perceived obligations can impact moral judgment.

7.1. Determinants of perceived obligation strength

Unexpectedly, across different choice contexts, differences in obligation strength were smallest when comparing judgments to help cousins to judgments to help siblings, contrary to what might be expected based on differences in the "relatedness coefficient" (Hamilton, 1964). As a result, when people make comparative third-person judgments, perhaps specific categorizations are most psychologically accessible, with strangers being thought of as "not family," and cousins and siblings being thought of as "family." In turn, people may not distinguish between obligations to help members of the same category (e.g., cousins and siblings) as much as they distinguish between obligations to help members of different categories (e.g., cousins and strangers). This interpretation finds support in recent research where adults distinguished in their third-person judgments between obligations to help strangers and family members (outgroup-ingroup) to a greater extent than between obligations to help friends and family members (ingroup-ingroup; Marshall et al., 2021, invited revision). Another possibility is that people simply encode social agents as being related or not, and, if related, then other factors are more important in determining obligation strength.

However, in our second experiment, when comparing judgments about members of the same family, relatedness judgment differences were always less strongly correlated with obligation judgment

differences than other social interaction judgment differences were. Therefore, a divergent interpretation is that relatedness simply serves as a cue for the frequency of past and future positive social interactions between others. That is, people may use relationship-based categories (e.g., cousin/sibling) to infer, for example, how often each currently in-need agent has helped a current potential helper in the past. In turn, people may then use these social interaction inferences to determine how much of an obligation the potential helper has to help each in-need agent in the present (see Table 9 for supporting evidence). This is consistent with research showing that co-residence duration is a better predictor of real and intended kin altruism than are genetic relatedness beliefs (Lieberman, Tooby, & Cosmides, 2007), as well as recent work that documents higher moral acceptability judgments of helping socially close (e.g., friends) over socially distant others (e.g., people in another country; Law et al., 2021). However, we caution readers that, in our experiments, we did not manipulate social interaction information within the same set of participants. We made inferences from correlational data that were collected simultaneously with obligation judgments during the pre-outcome task; therefore, further experimental research is needed to determine whether social interaction inferences indeed precede and mediate obligation judgment differences (see Pirlott & MacKinnon, 2016), and if so, which specific social interaction inferences are most important. Additionally, by manipulating (between participants) the nature of the relationship between agents as “stranger-like” versus “friend-like,” we discovered that people may believe that obligations to help are much weaker when agents are estranged as opposed to friendly. Together, these results suggest that not only do people’s obligation judgments track with relationship category, but they may also track with relationship quality.

We also documented important individual differences that influence judgments of obligation strength (see Tables 4 and 8). Across choice contexts, endorsement of family values (Curry, Chesters, & Van Lissa, 2019; Curry, Mullins, & Whitehouse, 2019) and ingroup loyalty (Graham et al., 2011) were more strongly positively associated with obligation judgments to help family members than non-family members, and more strongly positively associated with obligation judgments to help close family members than distant family members. Moreover, endorsement of family values tended to be more strongly associated with obligation judgments, and differences in obligation judgments, than endorsement of ingroup loyalty was. This suggests that endorsement of family values is a better predictor of obligation judgments that involve family members than endorsement of ingroup loyalty is, perhaps because the ingroup loyalty foundation does not treat family as a special category. Therefore, we consider our results as additional validation of the moral relevance of family values, and its uniqueness relative to similar constructs, as posited by the recent theory of “Morality-as-Cooperation” (see Curry, 2016; Curry, Chesters, & Van Lissa, 2019; Curry, Mullins, & Whitehouse, 2019).

In contrast, impartial beneficence (Kahane et al., 2018) was sometimes positively associated with obligation judgments to help family members, but, importantly, always positively associated with obligation judgments to help strangers. Contrary to our expectations, however, impartial beneficence was not consistently associated with obligation differences. That is, endorsement of impartial beneficence did not close the gap in judgments about obligations toward distantly related (or unrelated) others and closely related others. This was surprising because impartial beneficence is described as the “radical view that we should treat the well-being of *all* sentient beings *equally*—and that we are therefore *required* to give as much moral weight to distant strangers as to our closest relatives” (Kahane et al., 2018). However, we consider two solutions to this issue. First, familial obligations may belong to a special class of obligations that are relatively immune to individual differences in expansiveness of moral concern (see Waytz, Iyer, Young, Haidt, & Graham, 2019 for a measure of expansiveness of moral concern). That is, people who believe that they (and others) have obligations to humanity in general may believe that they (and others) have *even stronger*

obligations to their family members. This could explain why endorsement of impartial beneficence was not related to differences in obligation judgments. Although, to our knowledge, no research to date has tested the relationship between impartial beneficence and expansiveness of moral concern, impartial beneficence has been shown to be associated with other similar constructs, such as identification with all of humanity (Kahane et al., 2018; McFarland, Webb, & Brown, 2012). Second, the fact that none of our scenarios involved helping through the distribution of a finite but divisible resource may have produced these non-associations. That is, endorsement of impartial beneficence may close the gap in obligation judgments in contexts where agents can allocate a limited resource either equally or unequally (see Shaw, Choshen-Hillel, & Caruso, 2018). This is consistent with recent research documenting a negative relationship between endorsement of impartial beneficence and judgments of the morality of parochial empathy (Fowler, Law, & Gaesser, 2021).

7.2. Consequences of perceiving differential obligations

In accordance with prior research on the impact of obligations on moral evaluations (Everett et al., 2018; Hughes, 2017; Marshall et al., 2020; Marshall et al., 2021, *in prep*; McManus et al., 2020), our results further support the notion that people make moral evaluations based on others’ fulfilling or violating their relationship obligations. In general, we found that agents who helped unrelated others were judged as more morally good than agents who helped related others, but agents who helped distantly related (or unrelated) others instead of more closely related others were judged as less morally good than agents who did the opposite.

In contexts where there was only one potential beneficiary, however, one comparison did not follow these patterns. Across experiments, agents who helped a cousin were, on average, judged as no more morally good than agents who helped a sibling. To better understand these patterns, we investigated whether obligation judgments influenced moral character judgments at the individual level, rather than just on average (see Figs. 3 and 6, and Tables 3 and 7). Our data suggest that, in contexts where agents do not have to consider a choice about whom to help, differences in obligation strength are unassociated with differences in moral character. That is, there was not an association whereby the more people discriminated between distantly related (or unrelated) others and more closely related others in their obligation judgments, the more they discriminated (in the opposite direction) in their moral character judgments. This individual-level non-association also held when using repeated-measures correlations which ignore experimental condition information and analyze within-person variability as opposed to between-person variability (Bakdash & Marusich, 2017; see our OSF page).

These results suggest that the link between differential obligation judgments and differential moral evaluations may be more complicated than we and others had previously assumed (e.g., Marshall et al., 2020; Marshall et al., 2021, *invited revision*; McManus et al., 2020). Even when there were average differences in both obligation judgments and moral evaluations (e.g., STR vs SIB in Experiment 1), these individual-level relationships were still non-existent, suggesting that perhaps attribution theory (Kelley, 1967) is the wrong lens through which to frame these effects. We therefore interpret our data as supporting one of at least a few possibilities. Most people who made differential obligation judgments in this context did not also make differential moral evaluations, or, most people who made differential obligation judgments did not tend to agree on how to differentiate in their moral evaluations. However, it is also possible that these non-associations were an artefact of our experimental designs, as judgments in this context were made across two entirely different scenarios rather than within a single scenario (see Marshall et al., 2021, *invited revision* for a different design). Regardless, these findings highlight the need for more work exploring the underlying mechanisms for differences in moral evaluations in this

context. Further, the implications of our results echo recent calls to avoid using sets of aggregate estimates to draw conclusions about individual-level psychology (e.g., Fisher, Medaglia, & Jeronimus, 2018; Grice et al., 2020).

On the other hand, in contexts where agents had to consider a choice about whom to help, differences in obligation strength were consistently positively associated with differences in moral character. Specifically, the more people discriminated between distantly related (or unrelated) others and closely related others in their obligation judgments (with stronger obligations to help closer others), the more they discriminated in their moral character judgments (with more positive moral character judgments for agents who helped closer others). These individual-level associations also held when using repeated-measures correlations (see our OSF page). Together, the whole of our results suggest that obligations may be especially salient in contexts where choices about whom to help can or must occur, and in turn, this is when perceived obligation strength will be especially likely to structure subsequent moral evaluations.

7.3. Limitations and future directions

The current research has several important limitations. First, although we report our effects as being about perceived obligations (i.e., what people think others *ought* to do), these effects could be interpreted as being about more general expectations (i.e., what people think another [or most others] *would* do). However, there are empirical and conceptual reasons to be skeptical of this account. Extant research suggests that general expectations and obligations are separable, and they may exert distinct influences on moral evaluations. For example, adults believe, from a third party perspective, that friends are almost as likely to help one another as family members are, but, critically, they believe that family members are much more obligated to help one another than are friends, resulting in judging unhelpful family members as much meaner than unhelpful friends (Marshall et al., 2021, *invited revision*). We also speculate that obligations and expectations can be further disentangled experimentally. Specifically, we predict that manipulating an agent's prior behavior would affect judgments of how likely an agent would be to offer their help, but, importantly, this would not affect judgments of whether the agent is obligated to help. Ongoing research in our lab is investigating this possibility.

Second, we evaluated only U.S. adults' judgments of obligations, and so we are unable to make claims about their developmental trajectory in the U.S., or about their trajectory and endpoint outside of the U.S. However, an emerging body of research has started documenting these trajectories. Specifically, young children in the U.S. seem to *not* discriminate between family and non-family when making third-person obligation judgments, whereas older children and adults do make these distinctions (Dahl, Gross, & Siefert, 2020; Marshall et al., 2021, *invited revision*). These patterns hold across various cultures tested thus far (Marshall et al., 2021, *invited revision*), suggesting that young children may start out with more universal (versus parochial) conceptions of obligations. In contrast to these recent findings, U.S. infants seem to have a concept of ingroup obligation, expecting bystanders of transgressions to later withhold help from the transgressor if the transgression was against the bystander's ingroup, but not if the transgression was against the bystander's outgroup (Ting, He, & Baillargeon, 2019). This suggests that there may be important differences between obligations to help and obligations to punish. There may also be important differences between family versus non-family categories and the more general ingroup versus outgroup categories, as other research also suggests that infants (see Ting, Buyukozker Dawkins, Stavans, & Baillargeon, 2021, *forthcoming*, for a review) and children may recognize that agents have special obligations to ingroup (but not outgroup) members (Chalik & Dunham, 2020; Rhodes, 2012; Rhodes & Chalik, 2013; Weller & Lagattuta, 2014). Mapping the developmental trajectories and endpoints of obligations, within and across cultures, as well as within and

across group categories, is an exciting area for future research.

Third, our design focused on people's third-person judgments of unknown agents, which were based on severely impoverished information (i.e., an agent's single behavior). However, people will generally have stronger positive priors about their own family members and friends versus acquaintances or strangers (see Berg, Kitayama, & Kross, 2021, for neural evidence consistent with this idea). Therefore, when close others fail to fulfill their obligations or behave in prior-inconsistent ways, people may more easily generate alternative, positive explanations than for strangers or acquaintances, leading to potentially rational differences in moral evaluation (see Kim, Park, & Young, 2020). Supporting this claim, when people are given strong (versus weak) evidence of a novel agent's character, their trustworthiness judgments are more extreme (Kim, Mende-Siedlecki, Anzellotti, & Young, 2020). Further, when given false information about real-world others having behaved unethically, people judged close others as more moral than mere acquaintances (Forbes & Stellar, 2021). A related point is that people may make different judgments when they themselves are the recipients of help or harm, perhaps with important social consequences. Consistent with this idea, when people witnessed both friends and strangers stealing from them, even though people updated their impressions of their friends less so than strangers (Park, Fareri, Delgado, & Young, 2021), the degree to which people negatively updated their impression of their friends (relative to strangers) was associated with having fewer real-world friends (Park & Young, 2020). Similarly, when imagining a dispute with an acquaintance, people judged that their close friend's taking the side of the acquaintance, or their staying neutral, would be damaging to their friendship (Shaw, DeScioli, Barakzai, & Kurzban, 2017).

Last, recent calls have been made for researchers to communicate constraints on the generalizability of their findings (see Simons, Shoda, & Lindsay, 2017; Yarkoni, 2020). Another issue is that we cannot be certain that our experimental effects will generalize to new stimuli or new measures, as we only used a small set of stimuli and measures to test our main hypotheses. However, the fact that other researchers—who have used different stimuli and different measures—have found converging effects (Dahl et al., 2020; Everett et al., 2018; Hughes, 2017; Law et al., 2021; Marshall et al., 2020; Marshall et al., 2021, *invited revision*) suggests that our experimental effects may indeed generalize. Moreover, because we documented an effect of “choice context” on the direction of moral character judgments, as well as on the individual-level associations between obligation and moral character judgments, an important open question is what other kinds of relationships or contexts create obligations, and if (and how) these are integrated into moral evaluations and moral behavior (see Haidt & Baron, 1996; McManus et al., 2020, for demonstrations of how some non-kinship obligations are associated with moral evaluations, and Berry et al., 2021, *in press*; Earp et al., 2021, *preprint*, for similar ongoing conceptual and empirical work). Finally, another generalizability question concerns how varying the morally relevant behavior would affect relationship-based obligations and downstream moral evaluations. For example, prior research has shown that people believe lying to one's spouse, compared to a close friend or employee, reflects more poorly on the liar's character (Hughes et al., 2016). Therefore, it will be important for future research to broaden the scope of investigation of obligations beyond the domain of helping behavior.

8. Conclusion

By broadening the theoretical and methodological scope of prior work, the current research sheds light on when (i.e., in what contexts) and how (i.e., in what ways) relationship-based obligations may structure moral evaluation. This work suggests that relationship information influences perceived obligation strength, which, in turn, directly influences third-person moral evaluations in contexts where agents can make a choice of whom to help. Moreover, these data suggest that

personal moral values contribute to beliefs about prosocial obligations, with endorsement of family values, ingroup loyalty, and impartial beneficence being associated with judgments of obligation strength. Although untested thus far, these findings may shed light on how helping occurs in the real world, as beliefs about prosocial obligations may determine who helps whom. These findings add to a growing body of literature that suggests continuing to study the interplay between context and social relationships is important for understanding everyday moral cognition, and perhaps everyday moral behavior.

Author contributions

R.M. developed the experimental concept. R.M. and L.Y. developed the experimental designs. Stimulus construction and modification, and literature review, were performed by R.M. and J.M. Data collection and analyses were performed by R.M. R.M. interpreted analyses under the supervision of L.Y. R.M. drafted the manuscript, and J.M. and L.Y. provided critical revisions. All authors approved the final version of the manuscript for submission.

Author note

All data, analysis code, and supplemental materials for all experiments have been made publicly available via the Open Science Framework and can be accessed at https://osf.io/expmr/?view_only=6e80bfb029e64d02baafe276a61d425. The design and analysis plan for Experiment 1 was pre-registered at https://osf.io/6q4kr/?view_only=8033c65c9dfc41bcbc528fe94409e4e8, and Experiment 2 was pre-registered at https://osf.io/uxysz/?view_only=cad9b09ce5d1494eb023a604d9f6feb3.

Declaration of Competing Interest

The authors declare that there were no conflicts of interest with regard to the authorship or the submission of this article.

Acknowledgements

We thank the John Templeton Foundation (award number 61061) and the National Science Foundation (award number 1627157) for monetary support. We also thank members of the Morality Lab at Boston College, members of the Gaesser Lab at SUNY Albany, and Kurt Gray, who provided thoughtful discussion and helpful feedback at various stages of this project. We are especially thankful to Kayla Carew and Dan Ulanovsky, who also helped with stimulus construction and review of the literature in the early stages of this project. Similarly, we are especially thankful to Isaac Handley-Miner, Kevin Jiang, and Minjae Kim, who each reconducted portions of our analyses and then checked our reported statistics against their own analyses.

References

- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology, 8*, 456.
- Baron, J., & Miller, J. G. (2000). Limiting the scope of moral obligations to help: A cross-cultural investigation. *Journal of Cross-Cultural Psychology, 31*(6), 705–727.
- Berg, M. K., Kitayama, S., & Kross, E. (2021). How relationships bias moral reasoning: Neural and self-report evidence. *Journal of Experimental Social Psychology, 95*, 104156.
- Berry, Z., Lewis, N. A., & Sowden, W. J. (2021). The double-edged sword of loyalty. In *Current directions in psychological science* (in press).
- Bleske-Rechek, A., Nelson, L. A., Baker, J. P., Remiker, M. W., & Brandt, S. J. (2010). Evolution and the trolley problem: People save five over one unless the one is young, genetically related, or a romantic partner. *Journal of Social, Evolutionary, and Cultural Psychology, 4*(3), 115–127.
- Bloom, P. (2011). Family, community, trolley problems, and the crisis in moral psychology. *The Yale Review, 99*(2), 26–43.
- Burnstein, E., Crandall, C., & Kitayama, S. (1994). Some neo-Darwinian decision rules for altruism: Weighing cues for inclusive fitness as a function of the biological importance of the decision. *Journal of Personality and Social Psychology, 67*(5), 773–789.
- Chalik, L., & Dunham, Y. (2020). Beliefs about moral obligation structure children's social category-based expectations. *Child Development, 91*, e108–e119.
- Chandler, J., Rosenzweig, C., Moss, A. J., Robinson, J., & Litman, L. (2019). Online panels in social science research: Expanding sampling methods beyond mechanical Turk. *Behavior Research Methods, 51*(5), 2022–2038.
- Clark, M., & Mills, J. (1979). Interpersonal attraction in exchange and communal relationships. *Journal of Personality and Social Psychology, 37*, 12–24.
- Curry, O. S. (2016). Morality as cooperation: A problem-centred approach. In T. K. Shackelford, & R. D. Hansen (Eds.), *Evolutionary psychology: The evolution of morality* (pp. 27–51). Springer International Publishing.
- Curry, O. S., Chesters, M. J., & Van Lissa, C. J. (2019). Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *Journal of Research in Personality, 78*, 106–124.
- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology, 60*(1), 47–69.
- Dahl, A., Gross, R. L., & Siefert, C. (2020). Young children's judgments and reasoning about prosocial acts: Impermissible, suberogatory, obligatory, or supererogatory? *Cognitive Development, 55*, 100908.
- Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS One, 10*(4), Article e0121945.
- Earp, B. D., McLoughlin, K. L., Monrad, J. T., Clark, M. S., & Crockett, M. J. (2021). How social relationships shape moral judgment (preprint) <https://psyarxiv.com/e7cgq/>.
- Everett, J. A. C., Faber, N. S., Savulescu, J., & Crockett, M. J. (2018). The costs of being consequentialist: Social inference from instrumental harm and impartial beneficence. *Journal of Experimental Social Psychology, 79*, 200–216.
- Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191.
- Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106–E6115.
- Forbes, R. C., & Stellar, J. (2021). When the ones we love misbehave: Exploring moral processes within intimate bonds. (in press).
- Fowler, Z., Law, K. F., & Gaesser, B. (2021). Against empathy bias: The moral value of equitable empathy. *Psychological Science, 32*(5), 766–779.
- Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366–385.
- Gray, K., Young, L., & Waytz, A. (2012). Mind perception is the essence of morality. *Psychological Inquiry, 23*(2), 101–124.
- Greitemeyer, T., Rudolph, U., & Weiner, B. (2003). Whom would you rather help: An acquaintance not responsible for her plight or a responsible sibling? *The Journal of Social Psychology, 143*(3), 331–340.
- Grice, J. W., Medellin, E., Jones, I., Horvath, S., McDaniel, H., O'Lansen, C., & Baker, M. (2020). Persons as effect sizes. *Advances in Methods and Practices in Psychological Science, 3*(4), 443–455.
- Haidt, J., & Baron, J. (1996). Social roles and the moral judgment of acts and omissions. *European Journal of Social Psychology, 26*, 201–218.
- Hamilton, W. D. (1964). The genetical evolution of social behavior. *Journal of Theoretical Biology, 7*(1), 1–16.
- Hester, N., & Gray, K. (2020). The moral psychology of raceless genderless strangers. *Perspectives on Psychological Science, 15*(2), 216–230.
- Hughes, J. S. (2017). In a moral dilemma, choose the one you love: Impartial actors are seen as less moral than partial ones. *British Journal of Social Psychology, 56*, 561–577.
- Hughes, J. S., Creech, J., & Strosser, G. (2016). *Attributions about morally unreliable characters: Relationship closeness affects moral judgments*.
- Kahane, G., Everett, J. A. C., Earp, B. D., Caviola, L., Faber, N. S., Crockett, M. J., & Savulescu, J. (2018). Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychological Review, 125*(2), 131–164.
- Kelley, H. H. (1967). Attribution theory in social psychology. *Nebraska Symposium on Motivation, 15*, 192–238.
- Killen, M., & Turiel, E. (1998). Adolescents' and young adults' evaluations of helping and sacrificing for others. *Journal of Research on Adolescence, 8*(3), 355–375.
- Kim, M., Mende-Siedlecki, P., Anzelotti, S., & Young, L. (2020). Theory of mind following the violating of strong and weak prior beliefs. *Cerebral Cortex, 31*(2), 884–898.
- Kim, M., Park, B., & Young, L. (2020). The psychology of motivated versus rational impression updating. *Trends in Cognitive Sciences, 24*(2), 101–111.
- Kurzban, R., DeScioli, P., & Fein, D. (2012). Hamilton vs. Kant: Pitting adaptations for altruism against adaptations for moral judgment. *Evolution and Human Behavior, 33*, 323–333.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863.
- Law, K. F., Campbell, D., & Gaesser, B. (2021). Biased benevolence: The perceived morality of effective altruism across social distance. *Personality and Social Psychology Bulletin, https://doi.org/10.1177/01461672211002773*.
- Lee, J., & Holyoak, K. J. (2020). "But he's my brother": The impact of family obligation on moral judgments and decisions. *Memory and Cognition, 48*(1), 158–170.
- Lieberman, D., & Lobel, T. (2012). Kinship on the kibbutz: Coresidence duration predicts altruism, personal sexual aversions and moral attitudes among communally reared peers. *Evolution and Human Behavior, 33*, 26–34.
- Lieberman, D., Tooby, J., & Cosmides, L. (2007). The architecture of human kin detection. *Nature, 445*, 727–731.
- Linke, L. H. (2012). Social closeness and decision making: Moral, attributive and emotional reactions to third party transgressions. *Current Psychology, 31*, 291–312.

- Ludecke, D. (2020). *sjstats: Statistical functions for regression models*.
- MacFarquhar, L. (2015). *Strangers drowning: Grappling with impossible idealism, drastic choices, and the overpowering urge to help*. New York, New York: Penguin Press.
- Madsen, E. A., Tunney, R. J., Fieldman, G., Plotkin, H. C., Dunbar, R. I. M., Richardson, J., & McFarland, D. (2007). Kinship and altruism: A cross-cultural experimental study. *British Journal of Psychology*, *98*, 339–359.
- Makowski, D., Ben-Shachar, M. S., Patil, I., & Ludecke, D. (2020). Methods and algorithms for correlation analysis in R. *Journal of Open Source Software*, *5*(51), 2306.
- Malle, B. (2021). Moral judgments. *Annual Review of Psychology*, *72*, 293–318.
- Marshall, J., Gollwitzer, A., Mermin-Bunnell, K., Shinomiya, M., Relebsdorf, J., & Bloom, P. (2021). *How development and culture shape intuitions about prosocial obligations* (invited revision).
- Marshall, J., Wynn, K., & Bloom, P. (2020). Do children and adults take social relationship into account when evaluating people's actions? *Child Development*, *91*(5), e1082–e1100.
- McFarland, S., Webb, M., & Brown, D. (2012). All humanity is my ingroup: A measure and studies of identification with all humanity. *Journal of Personality and Social Psychology*, *103*(5), 830–853.
- McManus, R. M., Kleiman-Weiner, M., & Young, L. (2020). What we owe to family: The impact of special obligations on moral judgment. *Psychological Science*, *31*(3), 227–242.
- Neff, K. D., Turiel, E., & Anshel, D. (2002). Reasoning about interpersonal responsibility when making judgments about scenarios depicting close personal relationships. *Psychological Reports*, *90*, 723–742.
- Niemi, L., & Young, L. (2016). When and why we see victims as responsible: the impact of ideology on attitudes toward victims. *Personality and Social Psychology Bulletin*, *42*(9), 1227–1242.
- Palan, S., & Schitter, C. (2018). Prolific. Ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, *17*, 22–27.
- Park, B., Fareri, D., Delgado, M., & Young, L. (2020). The role of right temporo-parietal junction in processing social prediction error across relationship contexts. *Social Cognitive and Affective Neuroscience*, nsaa072.
- Park, B., & Young, L. (2020). An association between biased impression updating and relationship facilitation: A behavioral and fMRI investigation. *Journal of Experimental Social Psychology*, *87*.
- Passarelli, T. O., & Buchanan, T. W. (2020). How do stress and social closeness impact prosocial behavior? *Experimental Psychology*, *67*(2), 123–131.
- Petrinovich, L., O'Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology*, *64*(3), 467–478.
- Pirlott, A. G., & MacKinnon, D. P. (2016). Design approaches to experimental mediation. *Journal of Experimental Social Psychology*, *66*, 29–38.
- R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review*, *188*, 57–75.
- Revelle, W. (2020). *psych: Procedures for personality and psychological research*. Evanston, Illinois, USA: Northwestern University.
- Rhodes, M. (2012). Naïve theories of social groups. *Child Development*, *83*(6), 1900–1916.
- Rhodes, M., & Chalik, L. (2013). Social categories as markers of intrinsic interpersonal obligations. *Psychological Science*, *24*(6), 999–1006.
- Schein, C. (2020). The importance of context in moral judgments. *Perspectives on Psychological Science*, *15*(2), 207–215.
- Schein, C., & Gray, K. (2016). Moralization and harmification: The dyadic loop explains how the innocuous becomes harmful and wrong. *Psychological Inquiry*, *27*, 62–65.
- Schein, C., & Gray, K. (2018). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review*, *22*(1), 32–70.
- Shaw, A., Choshen-Hillel, S., & Caruso, E. M. (2018). Being biased against friends to appear unbiased. *Journal of Experimental Social Psychology*, *78*, 104–115.
- Shaw, A., DeScioli, P., Barakzai, A., & Kurzban, R. (2017). Whoever is not with me is against me: The costs of neutrality among friends. *Journal of Experimental Social Psychology*, *71*, 96–104.
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, *12*(6), 1123–1128.
- Simpson, A., Laham, S. M., & Fiske, A. P. (2016). Wrongness in different relationships: Relational context effects on moral judgment. *The Journal of Social Psychology*, *156*(6), 594–609.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, *87*, 245–251.
- Szycer, D., De Smet, D., Billingsley, J., & Lieberman, D. (2016). Coresidence duration and cues of maternal investment regulate sibling altruism across cultures. *Journal of Personality and Social Psychology*, *111*(2), 159–177.
- Tepe, B., & Aydinli-Karakulak, A. (2019). Beyond harmfulness and impurity: Moral wrongness as a violation of relational motivations. *Journal of Personality and Social Psychology*, *117*(2), 310–337.
- Ting, F., Buyukozzer Dawkins, M., Stavans, M., & Baillargeon, R. (2021). Principles and concepts in early moral cognition. In J. Decety (Ed.), *The social brain: A developmental perspective*. Cambridge, MA: MIT Press (forthcoming).
- Ting, F., He, Z., & Baillargeon, R. (2019). Toddlers and infants expect individuals to refrain from helping an ingroup victim's aggressor. *Proceedings of the National Academy of Sciences*, *116*(13), 6025–6034.
- Tomasello, M. (2020). The moral psychology of obligation. *Behavioral and Brain Sciences*, *43*, Article E56.
- Torchiano, M. (2020). *effsize: Efficient effect size computation*.
- Uhlmann, E. L., Zhu, L. L., Pizarro, D. A., & Bloom, P. (2012). Blood is thicker: Moral spillover effects based on kinship. *Cognition*, *124*(2), 239–243.
- Waytz, A., Dungan, J., & Young, L. (2013). The whistleblower's dilemma and the fairness-loyalty tradeoff. *Journal of Experimental Social Psychology*, *49*, 1027–1033.
- Waytz, A., Iyer, R., Young, L., Haidt, J., & Graham, J. (2019). Ideological differences in the expanse of the moral circle. *Nature Communications*, *10*, 4389.
- Weidman, A. C., Sowden, W. J., Berg, M. K., & Kross, E. (2020). Punish or protect? How close relationships shape responses to moral violations. *Personality and Social Psychology Bulletin*, *46*(5), 693–708.
- Weller, D., & Lagattuta, K. (2014). Children's judgments about prosocial decisions and emotions: Gender of the helper and recipient matters. *Child Development*, *85*(5), 2011–2028.
- Wickham, H., et al. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686.
- Yarkoni, T. (2020). The generalizability crisis. *Behavioral and Brain Sciences*. <https://doi.org/10.1017/S0140525X20001685>.
- Young, L., Camprodon, J. A., Hauser, M., Pasual-Leone, A., & Saxe, R. (2010). Disruption of the right temporo-parietal junction with transcranial magnetic stimulation reduces the role of beliefs in moral judgments. *Proceedings of the National Academy of Sciences*, *107*(15), 6753–6758.
- Young, L., Cushman, F., Hauser, M., & Saxe, R. (2007). The neural basis of the interaction between theory of mind and moral judgment. *Proceedings of the National Academy of Sciences*, *104*(20), 8235–8240.
- Yudkin, D. A., Gantman, A., Hofmann, W., & Quoidbach, J. (2021). Binding moral values gain importance in the presence of close others. *Nature Communications*, *12*, 2718.