UNIVERSITY of CALIFORNIA PRESS | Collabra

ORIGINAL RESEARCH REPORT

Moral Framing and Charitable Donation: Integrating Exploratory Social Media Analyses and Confirmatory Experimentation

Joe Hoover[*], Kate Johnson[*], Reihane Boghrati[†], Jesse Graham[*] and Morteza Dehghani[‡]

Do appeals to moral values promote charitable donation during natural disasters? Using Distributed Dictionary Representation, we analyze tweets posted during Hurricane Sandy to explore associations between moral values and charitable donation sentiment. We then derive hypotheses from the observed associations and test these hypotheses across a series of preregistered experiments that investigate the effects of moral framing on perceived donation motivation (Studies 2 & 3), hypothetical donation (Study 4), and real donation behavior (Study 5). Overall, we find consistent positive associations between moral care and loyalty framing with donation sentiment and donation motivation. However, in contrast with people's perceptions, we also find that moral frames may not actually have reliable effects on charitable donation, as measured by hypothetical indications of donation and real donation behavior. Overall, this work demonstrates that theoretically constrained, exploratory social media analyses can be used to generate viable hypotheses, but also that such approaches should be paired with rigorous controlled experiments.

The 900-mile-wide Hurricane Sandy hit the Atlantic Coast of the United States on October 29th, 2012 with record-breaking rainfall and 80-mile-per-hour winds. Within 5 days after landfall, more than 20 million tweets related to the storm were posted to Twitter (Guskin & Hitlin 2012). While the content of those tweets ranged widely, references to charitable donation were common. For example, the Red Cross estimated that over 2 million tweets were posted about their charitable efforts during and after Hurricane Sandy (Virtual Social Media Working Group and DHS First Responders Group, 2013).

This data affords an opportunity to investigate psychological phenomena relevant to charitable donation in a real-world setting, because digital traces of psychological processes are embedded in social media language (Boyd et al., 2015; Brady, Wills, Jost, Tucker, & Van Bavel, 2017; Park et al., 2015; Pennebaker, 2011; Sagi & Dehghani, 2014; Tausczik & Pennebaker, 2010). In this work, we gain insight into the representation of

charitable donation in public discourse by modeling the language used to frame charitable donation in the wake of Hurricane Sandy. We then use these insights to generate hypotheses regarding psychological factors that promote charitable donation and relevant constructs.

Developing a better understanding of charitable donation is doubly important. First, it is both a central example of prosociality and a widely acknowledge behavioral puzzle. Giving money to a stranger is a strange behavior indeed, and, as reviewed below, charitable donation is often multiply determined by a wide range of constructs that span multiple levels of analysis. However, despite this complexity, charitable donation can be operationalized and measured with relative ease and it thus offers a valuable opportunity to gain insight into more general prosocial processes. Second, beyond the contributions to basic science that charitable donation research can make, it has the potential to confer substantial benefits to society. Charitable donations constitute an essential component of intra- and international social improvement efforts. A better understanding of the factors that influence charitable donation can thus help organizations develop more efficacious campaigns.

We believe our approach to studying charitable donation — which combines computational analysis of naturalistic data and controlled confirmatory experimentation — is valuable for two primary reasons. First, social media has

[*] Department of Psychology, University of Southern California, Los Angeles, CA 90089, US

[†] Department of Computer Science, University of Southern California, Los Angeles, CA 90089, US

[‡] Department of Psychology and Computer Science, University of Southern California, Los Angeles, CA 90089, US

Corresponding author: Morteza Dehghani (mdehghan@usc.edu)

been identified as a powerful mechanism for collecting donations (Dietz, Druart, & Edge Research, 2015). For example, Dietz et al. (2015) found that $\frac{1}{3}$ of online donations are made via peer-to-peer social media mechanisms, such as peer-to-peer solicitation (Castillo, Petrie, & Wardell, 2014; Meer, 2011), which involves a person soliciting donations from their peers for an organization or cause. Social media has thus become an important ecosystem in which charitable donation dynamics with high-stakes play out. Despite the fact that social media is a relatively new phenomenon, understanding charitable donation processes in social media contexts is necessary for understanding charitable donation in the 21st century.

Second, the psychological information contained in social media offers an opportunity to conduct large-scale exploratory studies using data generated by a natural mechanism. Exploration is an essential component of the scientific process (Box, 1976; Tukey, 1977); but, it can be costly, particularly for resource intensive studies like those that require human participants, and this cost is further compounded by publishing biases toward confirmatory research. The psychological information contained in social media data enables relatively inexpensive exploratory analysis of large scale, naturally generated data.

However, while we believe that social media offers a high-value opportunity for exploratory hypothesis generation, taking advantage of this opportunity can be difficult. The methods employed must be powerful enough to discern the potentially small signals amid the considerable noise that characterizes social media data. However, if the goal is hypothesis generation, analyses must also be theoretically constrained, such that they target specific constructs. While it is now well-established that psychologically meaningful patterns can be detected in social media language, often these patterns are identified via theory-agnostic, data driven methods (See Dehghani et al., 2016 for discussion). While such bottom-up approaches has provided valuable evidence for the presence of psychological information in social media, they do not generally enable the derivation of hypotheses that target specific constructs.

In the current research, we use a two-stage approach that we believe maximizes the value of exploratory social media analysis. This approach pairs a theoretically constrained exploratory social media study with subsequent confirmatory experimental studies. To generate exploratory hypotheses, we estimate a set of hierarchical linear models using measurements obtained via a recently developed Natural Language Processing algorithm, Distributed Dictionary Representation (DDR; Garten et al., 2017), that harnesses the power of data-driven language modeling but also offers the precision of theory-driven measurement specificity. We then programmatically test these hypotheses with a series of preregistered, confirmatory experiments.

Promotive Factors of Charitable Donation
Previous research has identified a diverse range of factors associated with charitable donation. At a high level, most of these factors can be categorized into one of three taxonomies of effects: individual differences, situation characteristics, and solicitation framing. We see these taxonomies as corresponding to the three major components involved in a charitable donation: the person donating and the person (s)/entity they are donating to; the specific characteristics of the donation situation; and the donation solicitation itself. While this is by no means a perfect hierarchy or mapping, we believe it is useful for organizing a field of research that spans many distinct sub-fields, levels of analysis, and constructs.

Individual Differences Effects
In addition to the distinctions drawn above, the effects of individual differences on charitable donation can be further divided into functional groups: other-oriented, group-oriented, or values-oriented. Though, these boundaries are porous and we see these distinctions, again, more as a useful schema for conceptual organization than a rigid theoretical taxonomy.

Other-oriented factors. Canonical examples of other-oriented factors are empathy and perspective taking, which have been shown to increase charitable donation (Ashar et al., 2016; Penner, Dovidio, Piliavin, & Schroeder, 2005; Tusche, Böckler, Kanske, Trautwein, & Singer, 2016; Yamamoto, Yoo, & Matsui, 2015). When people feel empathy, tenderness (Ashar et al., 2016), or an awareness of need (Bekkers & Wiepking, 2011) toward victims, they are more likely to make a charitable contribution. Further, engaging in perspective taking or feeling empathy is thought to induce personal distress that further promotes charitable donation (Ashar et al., 2016), which, in such cases, is generally interpreted as a mechanism for emotional conflict reduction (Levine & Crowther, 2008; Piferi, Jobe, & Jones, 2006).

Group-oriented factors. In addition to other-oriented factors, charitable donation is also moderated by group-oriented factors. When people consider victims to be members of their own group, they are more likely to make charitable donations (Levine & Crowther, 2008; Piferi et al., 2006). However, these group biases appear to be moderated by complex social dynamics (Branscombe, Spears, Ellemers, & Doosje, 2002) and previous work has shown that they can be mitigated by interventions. For example, Freeman, Aquino, and McFerran (2009) found that an experimentally induced increase in feelings of moral elevation was associated with decreased in-group donation biases. Stronger valuations of morals associated with care and fairness are also associated with decreased in-group biases (Nilsson, Erlandsson, & Västfjäll, 2016).

Values-oriented factors. Finally, other research has focused on individual differences in values associated with charitable donation. As Bekkers and Wiepking (2011) note, a range of survey-based studies have identified associations between charitable donation and individual values related to prosociality (Bekkers, 2006; Van Lange, Bekkers, Schuyt, & Vugt, 2007), altruism (Bekkers & Schuyt, 2008; Farmer & Fedor, 2001), moral care (Schervish & Havens, 2002; Wilhelm & Bekkers, 2010), and social order (Todd & Lawson, 1999). Bekkers and Wiepking (2011) further note that people who feel a sense of obligation or

responsibility to society tend to show a greater tendency to make charitable donations (Amato, 1985; Reed & Selbee, 2002; Schuyt, Bekkers, & Smit, 2010).

In other work, Nilsson et al. (2016) used Moral Foundations Theory (MFT; Graham et al., 2012) to investigate associations between moral values and donation. They found that individual-oriented moral values (i.e. values associated with care and fairness) were associated with charitable behavior independent of target group affiliations. In contrast, they observed that while group-oriented moral values were positively associated with donation to in-group members, they were *negatively* associated with donation to out-group members.

Situation Characteristics Effects

Beyond individual differences, researchers also have identified situational factors — such as, reputational concerns, potential costs and benefits, awareness, and instrumental value (Bekkers & Wiepking, 2011) — that influence charitable donation. While these factors most likely interact with individual differences, the constitute distinct effects with potentially substantial consequences. For example, Ashar et al. (2016) found that both the blamelessness of victims and the instrumental value of donating both mediated the effect of compassion on donation. Donation is also influenced by other situational factors such as social signaling (Piferi et al., 2006) and desired reciprocity (Levine & Crowther 2008). Even simple factors like disaster awareness have important effects on donation (Martin 2013).

Framing Effects

While there are numerous individual and situational factors that influence charitable donation, most instances of donation include a third component: the solicitation itself (Bekkers & Wiepking, 2011; Bryant, Jeon-Slaughter, Kang, & Tax, 2003). Characteristics of solicitations affect charitable donation above and beyond individual-differences and situational effects. Most of the research on charitable donation and framing effects has focused on constructs from behavioral economics and decision-making.

For instance, when solicitations are framed as economic transactions rather than acts of charity, they tend to be more effective (Holmes, Miller, & Lerner, 2002; Zlatev & Miller, 2016) and framing donation opportunities as exceptional, rather than ordinary, tends to increases donations (Sussman, Sharma, & Alter, 2015). Charitable donations are also counter-intuitively moderated by the number of victims emphasized in solicitations, such that frames that focus on one victim are more effective than frames that focus on many victims (Slovic, 2010; Small, Loewenstein, & Slovic, 2007; Västfjäll, Slovic, Mayorga, & Peters, 2014). Still other factors such as image valence, temporal framing (Chang & Lee, 2009), message abstractness (Das, Kerkhof, & Kuiper, 2008), and self- vs. other-beneficial frames (Fielding, Knowles, & Robertson, 2017) have also been identified as moderators of charitable donation.

Current Work

Charitable donation is moderated by individual-differences with different functional emphases, situational effects, and framing effects. While these distinctions are not necessarily explicitly drawn in the literature, most research has been focused on mechanisms that fall under one of these super-ordinate categories. However, we suspect that charitable donation is likely influenced by a more complex network of factors; that is, the super-ordinate categories of factors that we note, which have largely been studied in isolation, most likely interact in the real-world. Accordingly, we propose that research on charitable donation that draws from multiple factor categories may lead to a more nuanced understanding of charitable donation.

Starting from this position, we investigate the function of moral values in solicitation framing. While values have been noted as important charitable donation factors (Amato, 1985; Bekkers, 2006; Bekkers & Schuyt, 2008; Farmer & Fedor, 2001; Nilsson et al., 2016; Reed & Selbee, 2002; Schervish & Havens, 2002; Schuyt et al., 2010; Todd & Lawson, 1999; Van Lange et al., 2007; Wilhelm & Bekkers, 2010), they have primarily been treated as a source of individual difference in the charitable donation literature. However, given the sensitivity of charitable donation behavior to solicitation framing, we hypothesize that the efficacy of solicitations is moderated by the moral values that they evoke.

To operationalize moral values, we rely on the framework established by Moral Foundations Theory (MFT; Graham et al., 2012). MFT proposes that moral values can be decomposed into five distinct foundations constituted by a bi-polar continuum between so-called virtues and vices. These five foundations are: care/harm, fairness/cheating, loyalty/betrayal, authority/subversion, and purity/degradation. While a relatively large body of research has relied on MFT (*Publication Search|moralfoundations. org*, n.d.), *the* degree to which it is psychologically valid is still debated. Most prominently, critics of MFT propose that moral values are best understood as variations of harm concerns (Gray & Keeney, 2015a, 2015b), rather than emergent products of distinct classes of concerns (Graham 2015). Despite these open questions, we operationalize moral values using MFT for several reasons. First, MFT offers the most diverse and well-established pluralistic model of moral values. Even critics of MFT acknowledge that moral values are pluralistic and propose that MFT reflects different "flavors" of values (Schein & Gray, 2017). Regardless of whether these different flavors emerge from a single omnibus concern about harm or multiple foundations, we are interested in their differential effects on solicitation frames. Accordingly, it is necessary that our operationalization of moral values reflects their pluralism.

Our second reason for privileging MFT is that it is the only taxonomy of moral values (1) for which term dictionaries have been developed (Graham et al., 2011) and (2) that has been repeatedly employed in studies of natural language (Dehghani et al., 2016; Graham et al., 2011; Sagi & Dehghani, 2014). Because exploratory natural language modeling is a major component of the current

work, selecting an operationalization of moral values with a precedent in such research is important. This precedent increases our baseline certainty in the validity of our measurement models and thus helps reduce potential sources of error.

To investigate the effects of morally evocative solicitation frames, we begin with a large-scale, naturalistic exploration of the moral language used to frame messages relevant to charitable donation. Rather than starting with a set of a priori hypotheses, this approach allows us to observe the linguistic dynamics related to charitable donation in a natural setting. We then use the observations drawn in this exploratory phase to derive specific hypothesis about the association between expressions of moral values and charitable donation solicitations. Specifically, we conduct five studies that investigate the relationship between moral framing and multiple constructs relevant to charitable donation: donation sentiment (Study 1), perceived donation motivation (studies 2 & 3), hypothetical donation (studies 4 & 5), and real donation (Study 5).

Beyond focusing on an under-explored domain, this research demonstrates how cutting-edge, data-driven Natural Language Processing (NLP) methods can be used with theoretical constraints and integrated into an experimental research paradigm. In Study 1, we estimate the semantic association between charitable donation sentiment and moral values using DDR (Garten et al., 2017), an NLP framework that uses distributed representations (Le & Mikolov, 2014; Mikolov, Yih, & Zweig, 2013) learned by a neural network to measure the presence of latent semantic constructs in short texts. Specifically, we rely on DDR to model the association between expressions of moral values and language associated with charitable donation in a corpus of tweets posted during and after Hurricane Sandy. While other recent research has used similar approaches to conduct confirmatory hypothesis tests (Dehghani et al., 2016; Sagi & Dehghani, 2014), here we use this as an exploratory framework for investigating the association between moral values and three donation relevant phenomena that we then study across 4 controlled experiments:

- *Perceived donation motivation.* The perceived motivational power of a donation solicitation (Studies 2 & 3).
- *Hypothetical donation.* The hypothetical monetary donation amount reported for a donation solicitation (Studies 4 and 5).
- *Real donation.* The real monetary donation amount given in response to a solicitation (Study 5).

We believe this approach, which uses large scale analysis of observational data to develop hypotheses that are then tested using controlled experiments, is a valuable tool for psychological research. We show how the traces left in social media can be modeled in a way that informs hypothesis generation. Importantly, we also show that hypotheses generated using this approach not only correspond to previous research, but also can be replicated using conventional experimental paradigms.

Participant and Population Considerations
Over-reliance on student subject pools has led to well-known biases across much of social psychological research. Such biases pose particularly serious problems for research on charitable donation. The majority of charitable donation behavior is *not* enacted by undergraduate students and it is thus not at all clear what can be learned about charitable donation via studies conducted with student subject pools. To avoid these issues, none of the studies reported in this work were conducted with student populations.

However, more importantly than simply avoiding student subject pools, the reported studies rely on data generated by a diverse set of mechanisms. Our first study relies on Twitter data posted during and after Hurricane Sandy. Importantly, Twitter users are *not* randomly sampled and Twitter data introduces its own set of problematic biases (Ruths & Pfeffer, 2014). For example, Twitter accounts can be registered for bots and organizations with disproportionately high rates of tweeting; the demographic distributions of Twitter users are not representative; and Twitter itself employs undisclosed algorithms designed to promote and influence user engagement. Such factors complicate the relationship between sample findings and population dynamics and they raise potential issues for generalization. To account for the first issue, we focus only on unique tweets which limits the influence of high-output accounts; however, it is much more difficult to address issues of representativeness and ecological manipulation. Nonetheless, this data is still drawn from a population that is more diverse than standard student populations and, despite the noted limitations, the data we analyze is particularly valuable because it was generated by a relevant real-world mechanism. That is, our first study intends to model the moral frames people employ when discussing charitable donation and, to that end, we use data from real public discourse to model the target construct.

Further, studies 2–5 rely on online data collection platforms that offer additional inclusiveness. Specifically, studies 2 and 5 were conducted using Amazon Mechanical Turk (MTurk), which offers a more diverse and representative subject pool, compared to standard student subject pools (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011). Studies 3 and 4 were conducted using www.yourmorals.org, an online platform where users can participate in surveys relevant to moral psychology (Graham et al., 2011). By relying on data generated by three distinct mechanisms (Twitter, MTurk, and yourmorals. org), our goal was to ensure greater sample diversity and inclusiveness, compared to student samples, as well as maximize generalizability and external validity.

Power and Methodological Precision Considerations
The reported sequence of studies was carefully designed to maximize power and precision. Our initial hypotheses were generated from a high-powered, ecologically valid, exploratory analysis of data naturally generated by a mechanism that is directly relevant to our phenomena of interest. To maximize statistical precision and reduce overall model variance, the parameter estimates used

to derive hypotheses in this study were generated by aggregating across 500 hierarchical linear models. However, prior to this analysis, rigorous method validation was conducted to ensure the validity of our estimates of moral sentiment (See Method Validation in Supplemental Material).

Further, prior to the confirmatory testing conducted across the next five studies, we developed donation solicitation items via a multi-staged design process (See Study 2 Method section). Each confirmatory experiment was also preceded by a power-analysis to determine sample size. The power analysis for each study was based on parameter estimates observed in the prior studies.

To further ensure the validity of our confirmatory process, each confirmatory study (studies 2–5) was preregistered via the Open Science Foundation and links to these pre-registrations are provided under each study method section. Finally, across each study, we aimed to replicate some set of the results observed in the prior studies. For example, Study 3, in addition to reporting new tests, replicates the tests reported in Study 2; and Study 5 conceptually replicates the tests reported in Study 4, in addition to reporting new tests.

Overall, in this work we sought to maximize power by including data drawn from diverse sources, relying on diverse methodologies, collecting samples large enough for sufficient power, and conducting nested replications of effects. We would like to note again that all the behavioral studies in this paper were pre-registered prior to data collection, and all measures, manipulations, and exclusions in the studies are disclosed.

Study 1

During and after Hurricane Sandy, more than 20 million tweets related to the storm were posted to Twitter (Guskin & Hitlin, 2012) and many of these mentioned opportunities for charitable donation (Virtual Social Media Working Group and DHS First Responders Group, 2013). Accordingly, this natural phenomenon offers a valuable opportunity to explore how charitable donation is framed in daily social discourse and, more specifically, (1) whether references to charitable donation (a construct we refer to as donation sentiment) occur within moral values frames and (2) whether distinct classes of moral values frames are differentially associated with donation sentiment. To model the association between donation sentiment and specific moral frames, we use DDR to estimate the presence of specific moral values in a corpus of tweets associated with Hurricane Sandy.

We operationalize moral values using MFT; however, as each of the five MFT foundations is constituted by two distinct subordinate domains (e.g. Care vs. Harm) we treat the MFT framework as a model of 10 moral values, rather than conceptualizing each foundation as an atomic unit. The advantage of this approach is that any effects that hold for an entire foundation (i.e. virtue and vice) can still be identified in the 10 value model. In contrast, if we were to rely on the standard 5 value model, it would be impossible to detect a positive association between, say,

care and donation sentiment in the presence of a negative association between harm and donation sentiment.

Method

The data used in this study consists of 7,222,763 tweets posted between 10/16/2012 and 11/05/2012 and containing Hurricane Sandy (HS) hashtags ('#sandy,' '#HurricaneSandy').[1] For this analysis, we focus only on *original* and unique tweets, thus excluding retweets. These tweets were then preprocessed using standard methods (Vijayarani, Ilamathi, & Nithya, 2015). Specifically, punctuation and emoticons were stripped from each Tweet and URLs and user references (indicated on Twitter via the '@' symbol followed by a username) were replaced with the strings 'URL' and '@user,' respectively. Tweets were also labeled by location, time of post, and tweet/retweet status. Location was determined using a combination of geo-tag binning and string matching, where place names in the user Location field were matched to actual geographical place names. To ensure a minimal level of affiliation among the authors of the tweets in the corpus, we excluded tweets that were not posted in the United States from our analysis, which also necessitated excluding tweets for which location could not be determined (N tweets that met geographic inclusion criterion = 1,068,301). Finally, we also excluded tweets that were posted before Hurricane Sandy made landfall in the United States (N tweets that met both geographic and temporal inclusion criteria = 913,987). These exclusions reduced the original 7,222,763 tweets to 913,987, an admittedly large reduction. While there are certainly reasonable arguments that could be made against dropping so much data, one of the largest barriers to external validity in social media analysis is sample noise and bias (Ruths & Pfeffer, 2014), which arises from myriad factors including non-human accounts that post at disproportionately high rates (e.g. spam bots and organizational accounts). To mitigate these sources of bias, we focus only on unique and original tweets, which helps minimize the influence that non-human accounts have on the population of tweets by excluding tweet duplicates. Further, we exclude accounts that we cannot ensure meet our inclusion criterion (US based accounts), which also helps eliminate junk accounts from the sample. Of course, these decisions introduced new biases. For example, tweets posted by US-located users who did not report location information were excluded and we cannot determine whether and to what to degree any observed effects would be homogeneous within this population. Nonetheless, we believe that the potential biases introduced by our exclusion criteria are less severe than those present in unfiltered twitter data. However, the unknown risks of potential unidentified sample biases underscore the import of supplementing social media analyses with controlled experimentation, which is the approach applied in the current work.

After pre-processing and data selection, we then calculated the loading of each of the remaining tweets on each of the 10 moral foundations dimensions. As Garten et al. (2017) explain, DDR (1) uses distributed representations of words to generate distributed representations of latent

semantic constructs and (2) compares the representation of these latent constructs to representations of a given text (e.g. a tweet) in order estimate the loading of the text on the construct. As in Garten et al. (2017), the distributed representations we use to implement DDR were generated by the Word2Vec algorithm (Mikolov et al., 2013), which is a shallow neural network that can be efficiently trained on massive linguistic corpora.[2]

In essence, DDR provides the same kind of theoretical specificity characteristic of other dictionary-based methods (Pennebaker, 2011; Tausczik & Pennebaker, 2010); however, it also offers multiple advantages. Foremost is the fact that only a small set of words (e.g. 4, see **Table 1** for the seed words used in this study) are necessary for representing a semantic construct, which drastically reduces the cost of dictionary generation. Further, because the loading of a text on a latent construct does not depend on the presence of specific words, as in conventional dictionary-based approaches, but rather on proximity in high-dimensional space, DDR is robust to lexical gaps. Because Tweets are constrained to being very short by Twitter's platform, DDR's reliance on semantic proximity, as opposed to lexical matching, is particularly valuable because it relaxes the requirement that a text contain specific words from a given dimension in order to be counted as expressing that dimension.

We then use DDR to calculate the moral loadings for the entire Twitter corpus (N = 913,987). To operationalize donation sentiment, we used a straightforward approach that involved automatically labeling tweets as containing donation sentiment if they contained any of seven hashtags known to be associated with donation solicitations during Hurricane Sandy (#donate, #sandyaid, #sandy_aid, #redcross, #red_cross, #howtohelp, #90999). Importantly, however, to ensure independence between moral loadings and these hashtags, we excluded these hashtags when calculating the moral loadings. While previous research has demonstrated that DDR estimates of semantic loadings correspond well with human ratings (Garten et al., 2017), we also evaluated the validity of our DDR estimates in a pilot study by comparing them to human annotations of moral content in a subset of the

Hurricane Sandy data. The results of this study converged with Garten et al. (2017), indicating that DDR estimates, though noisy, reliably track human judgments of moral content (See Method Validation in Supplemental Materials for more details).

Finally, in order to model the association between donation sentiment and moral sentiment, a hierarchical linear model with variable intercepts was fit using Maximum Likelihood estimated. In this model, moral loading was the dependent variable. Fixed effects were estimated for factors representing donation sentiment (Donation = 0/No Donation = 1), moral foundation (11 levels; 10 moral foundation dimensions and non-moral), and their interaction. Finally, tweet ID was used as the second-level group variable. Permitting the tweet-level intercepts to vary is necessary, because each tweet has a value for each moral domain, which means that this is functionally a repeated measures design. Further, to in investigate potential differences in moral loading as a function of moral domain, pairwise comparisons with Tukey's multiple comparisons adjustment method were conducted. To reduce model variance, 500 separate models were estimated on subsets of the data that were randomly sampled without replacement. However, due to an imbalance between tweets containing non-donation (N = 888,023) and donation sentiment (N = 25,964), each sample consisted of the full 25,964 donation tweets and a randomly selected set of non-donation tweets. Thus, each model estimated compared our best idea of the estimated means of moral sentiment among tweets containing donation sentiment to the estimated means of moral sentiment among a random sample of non-donation tweets. To evaluate the results of these models, parameters were aggregated across models.

Results

Aggregated results from the random intercepts GLMs revealed that substantial variance in moral loading occurs at the tweet level ($\overline{\mu}$ ICC = 0.57, $SD = 0.001$). This indicates that the loadings of a given tweet on the 10 moral domains are considerably correlated, which suggests that people tend to evoke multiple moral domains simultaneously. Further, from **Figure 1** it is apparent that tweets containing donation sentiment exhibit far greater variance in mean levels of moral sentiment, compared to those that do not contain donation sentiment. More importantly, however, these results indicate that, among donation tweets, compared to other moral domains care ($\overline{\mu} = 0.51$, SE = 0.01, $d^3 = 0.36$, 95%CI [0.49, 0.52], p < 0.001)[3] and loyalty ($\overline{\mu} = 0.32$, SE = 0.01, d = 0.23, 95%CI [0.31, 0.33], p < 0.001), contain relatively high estimates of moral sentiment (See **Table 2** for examples of care and loyalty tweets).

Indeed, the only other positive mean loadings are for fairness ($\overline{\mu} = 0.19$, SE = 0.01, d = 0.14, 95%CI [0.17, 0.20], p < 0.001), cheating ($\overline{\mu} = 0.15$, SE = 0.01, d = 0.11, 95%CI [0.13, 0.16], p < 0.001), and harm ($\overline{\mu} = 0.13$, SE = 0.01, d = 0.09, 95%CI [0.12, 0.14], p < 0.001); of these, the foundation with the highest mean loading, fairness, has an effect that is substantially lower than those of care and loyalty, $\overline{\mu}_{diff} = 0.32$, $SD_{pooled} = 1.39$, $d = 0.23$ and $\overline{\mu}_{diff} = 0.13$, $SD_{pooled} = 1.39$, $d = 0.10$.[4]

Table 1: DDR moral seed words.

| Moral Domain	Seed Words
Care	kindness, compassion, nurture, empathy
Harm	suffer, cruel, hurt, harm
Fairness	fairness, equality, justice, rights
Cheating	cheat, fraud, unfair, injustice
Loyalty	loyal, solidarity, patriot, fidelity
Betrayal	betray, treason, disloyal, traitor
Authority	authority, obey, respect, tradition
Subversion	subversion, disobey, disrespect, chaos
Purity	purity, sanctity, sacred, wholesome
Degradation	impurity, depravity, degradation, unnatural
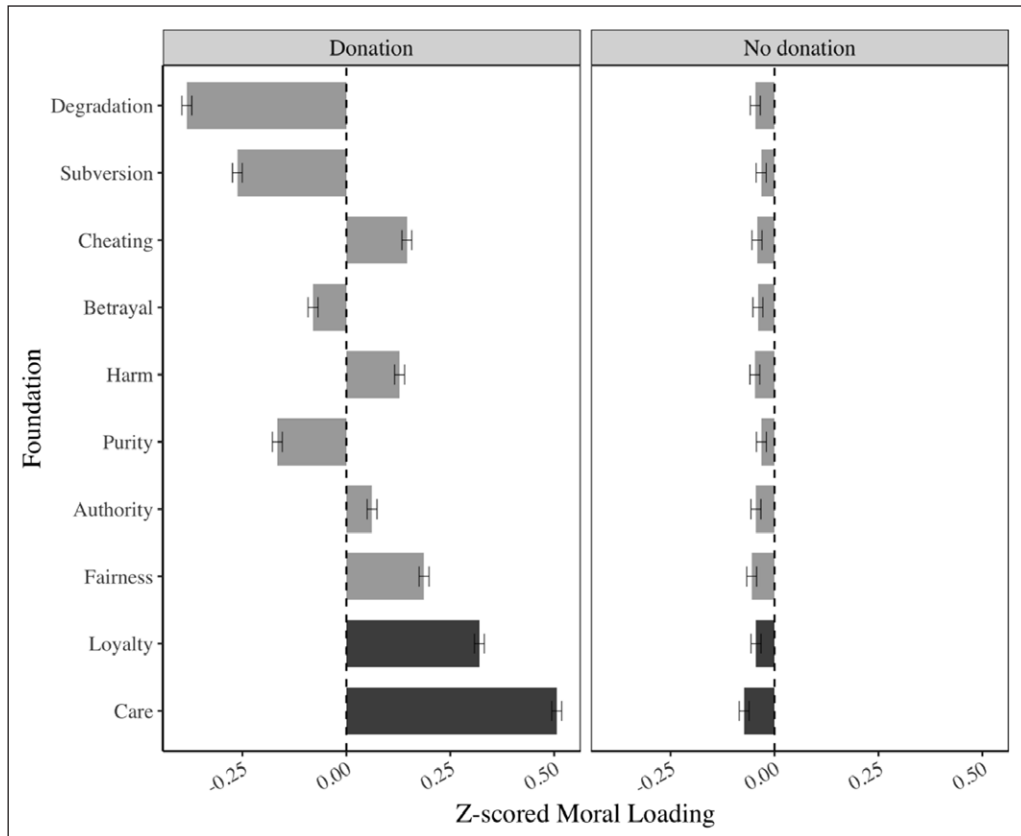
Figure 1: Moral loadings of donation sentiment. Error bars indicate 95%*CI*. The top and bottom panels depict loadings for tweets that do and do not contain donation sentiment, respectively.

Table 2: Study 1 Examples of donation Tweets with high estimated care and loyalty loadings.

Domain	Tweet Text	Standardized Loading
Care	sandyhelp show your compassion and donate today[†]	5.25
Care	Really inspiring stories of healing and humanity Time to donate SandyHelp	5.10
Care	Do something selfless donate SandyHelp	4.43
Loyalty	I just donated to help our fellow citizens SandyHelp at URL Show humanity your compassion Let's be there for others	4.39
Loyalty	Love my fellow brothers and sisters in New Jeersey [sic] And fellow Americans standing strong as a nation Sandy please donate to local shelters	4.78
Loyalty	Text REDCROSS to 90999 & give 10 to support our friends family and fellow countrymen Sandy Together we can make a difference	4.61

[†] As stated, indicators of donation sentiment such as 'sandyhelp' were removed from the tweets prior to analysis. However, such phrases are included in these examples for readability.

Notably, this analysis also revealed negative effects for subversion ($\bar{\mu} = -0.26$, SE $= 0.01$, d $= -0.19$, 95%CI [-0.27, -0.25], p < 0.001), degradation $\bar{\mu} = -0.38$, SE $= 0.01$, d $= -0.28$, 95%CI [-0.4, -0.37], p < 0.001, betrayal ($\bar{\mu} = -0.08$, SE $= 0.01$, d $= -0.06$, 95%CI [-0.09, -0.07], p < 0.001), and purity ($\bar{\mu} = -0.17$, SE $= 0.01$, d $= -0.12$, 95%CI [-0.18, -0.15], p < 0.001), such that tweets containing donation sentiment had loadings on these domains that were below the average. Thus, for example, Tweets containing donation sentiment emphatically *did not* use subversion or degradation frames.

Discussion

These results indicate that the moral frames people used in tweets about donation were highly heterogeneous, such that donation sentiment appears to be most strongly positively associated with care and loyalty and most strongly negatively associated with subversion and degradation. Both the observed negative and positive effects fit with previous literature. For example, cues of injustice are known to automatically draw attention (Baumert, Gollwitzer, Staubach, & Schmitt, 2011; Hafer, 2000) and elicit disgust (Haidt, 2003) and Baumert, Thomas, and Schmitt (2012) propose that observers' perceptions of injustice can diminish altruistic behavior. Further, Zagefka, Noor, Brown, de Moura, and Hopthrow (2011) found that study participants were less likely to help victims whom they perceived as implicitly responsible for their situation. While it is not necessarily the case that tweets containing subversion and degradation frames necessarily focus on perceptions of injustice or blameworthiness, makes sense that tweets containing

these frames do not tend to also contain donation sentiment. That is, in light of this research, when people are tweeting donation sentiment, it makes sense that such sentiment is not framed with moral concerns about subversion, degradation, and betrayal. That said, this framework does not offer an explanation for why tweets containing donation sentiment also had lower loadings on purity. Indeed, given the association between purity and sub-concepts like sanctity and wholesomeness, one might expect that donation tweets would be positively associated with purity frames.

Regarding the observed positive associations, the effects of care and loyalty echo previous research that finds that empathy and group affiliation are associated with increased donation (Levine, Cassidy, & Jentzsch, 2010). While care and loyalty do not necessarily map directly onto empathy and group affiliation, it would make sense if moral values associated with care and loyalty play a role in empathic and affiliative processes. Given this indirect precedent in the literature and our focus on promotive charitable donation factors, in the subsequent studies we chose to further investigate the potential association between moral care and loyalty and charitable donation. While there are doubtless multiple, overlapping motivations for tweeting about donation, we make the simplifying assumption that a major motivation of tweeting about donation is to motivate others to donate. From this view, one possible implication of the effects observed in this study is that people believe, on some level, that care and loyalty frames are effective donation motivators. Accordingly, in the next study, we conduct an experimental test of this hypothesis.

Study 2
In Study 1, we found that people tended to use language associated with moral care and loyalty when making charitable donation relevant posts during Hurricane Sandy. In the current study, we extend this finding by testing experimentally the hypothesis that donation solicitations containing care or loyalty rhetoric are perceived as stronger motivators of charitable donation. To maintain a clear division between the exploratory analyses in Study 1 and the current study, this study pre-registered via the Open Science Framework (https://osf.io/wvb3e/?view_only=b8f4b0b8f7bb441a8074a2901bc2956d).

In addition to testing our primary hypothesis, that care and loyalty rhetoric enhances the perceived efficacy of donation solicitations, we also preregistered and investigated a similar effect on diffusion motivation. That is, we tested the hypothesis that care and loyalty rhetoric enhances the perceived likelihood of voluntarily disseminating a donation solicitation. Further, in order to better understand the nature of the effects of donation solicitation, we explored whether the hypothesized effects of care and loyalty framing on donation motivation and diffusion motivation would remain robust after controlling for the degree to which the donation solicitations "made sense" to participants, a measure intended to approximate fluency (Reber & Schwarz, 1999; Reber, Winkielman, &

Schwarz, 1998). However, because the primary focus of the current research is the association between moral framing and charitable donation, we report the results of the diffusion motivation and "sense" control models in the Study 2 section of Supplemental Materials.

Method
Participants (N = 372, 51% Female, Mean Age = 36.38, SD Age = 12.31) were recruited from MTurk and paid $0.20 for their participation. This study employed a between-subjects design in which participants were randomly assigned to one of two conditions. In one condition, participants rated three non-moral tweets on three dimensions: the degree to which the tweets would motivate them to donate, the degree to which they would be motivated to retweet the tweets, and the degree to which the tweets "made sense". In the other condition, participants rated both three care and three loyalty tweets on the same dimensions.

The tweet stimuli that participants rated were selected from a pool of fake donation solicitation tweets written by the researchers to reflect different moral concerns. Initially, a large pool (N = 100; 10 per moral domain) of solicitation tweets was generated. These tweets were than coded by expert annotators and tweets that were mis-coded were dropped based on the assumption that they failed to adequately express the target domain. These tweets were then pre-rated by an independent sample of MTurk workers (N = 157) for vividness, arousal, and valence. For each moral domain, tweets with moderate scores on all three dimensions were selected for use in subsequent research. Specifically, depending on their assignment, participants responded to both three tweets expressing either care (e.g. 'Show your compassion for the people affected by #HurricaneSandy', please donate now. #Sandy #SandyDonate) and three tweets expressing loyalty (e.g. The people affected by #HurricaneSandy need your loyalty and support, please donate now. #SandyDonate'), or just three tweets expressing no moral values (e.g. '#SandyDonate is freaking awsome! Donate now! #Sandy #HurricaneSandy').

Prior to data collection, we conducted a power analysis in order to determine the requisite sample size to detect a moderately small effect (d = 0.30) with 80% power (342) and planned to collect an additional 30 participants to account for attrition and attention-check failures. Data was collected in two waves. The first wave (N = 210) was collected and used to estimate item reliabilities using Coefficient alpha (Cronbach, 1951). Given sufficient alphas, we proceeded to collect the remaining participants. In both waves, a three-item attention check was administered that required participants to report specific answers to these items. This study was approved by the USC IRB panel (ID UP-15-00380).

Results
In the first wave of data (N = 210), 37 participants failed the manipulation check, yielding an N of 173. Cronbach's Alpha for each condition (*Care* = 0.91, *Loyalty* = 0.91, *Non-moral* = 0.70) was greater than the

planned cut-off (0.60), thus the remainder of the dataset was collected. Due to a minor data collection error, 386 participants were collected, rather than the planned 372; however, after excluding participants who failed the manipulation check in both waves (Wave 1 = 37; Wave 2 = 36), *N* was reduced to 313. As planned, items were then averaged, creating a donation motivation index for each condition.

Congruent with our hypotheses, one-tailed t-tests[5] indicated that participants reported higher donation motivation in the care (*M* = 4.16, *SD* = 1.5) and loyalty (*M* = 4.28, *SD* = 1.58) conditions compared to the control condition (*M* = 3.55, *SD* = 1.40), *Cohen's d*[6] = 0.42, 95%*CI* [0.20, 0.65], *t*(308) = 3.75, 95%*CI* = [0.29, 0.94], *p* < 0.001 and *Cohen's d* = 0.49, 95%*CI* [0.26, 0.72], *t*(308) = 4.33, 95%*CI* = [0.39, 1.06], *p* < 0.001 (See **Figure 2**).

Discussion

The results of Study 2 correspond well to the observations made in Study 1, which suggests that the hypotheses derived from Study 1 are at least minimally stable. Specifically, both care and loyalty frames are perceived as stronger donation motivators compared to a non-moral control. Although the finding that care and loyalty rhetoric are perceived as stronger motivators of charitable donation than a non-moral control is perhaps unsurprising, this is nonetheless an important validation. That said, a more important question is whether the perceived donation motivation of care and loyalty exceeds that of other moral frames. Accordingly, we address this question in the next study.

Study 3

In Study 2, we found that donation solicitations containing care and loyalty frames were seen as stronger motivators of donation, compared to a non-moral control condition, and that they also were seen as stronger motivators of diffusion. In the current study, we further investigate these associations by comparing the effects of care and loyalty framing to the effects of fairness, cheating, and harm frames, the domains with the next highest positive effects in Study 1. While the previous study provided evidence for the effects of care and loyalty rhetoric on donation motivation, a replication of the differences in magnitude between care and loyalty and other moral concerns would provide much stronger evidence for the hypotheses that care and loyalty frames are particularly relevant to charitable donation, compared to other moral domains.

Method

Participants (N = 1116, % Female = 36) were recruited via www.yourmorals.org, an online platform where users can participate in surveys relevant to moral psychology (Graham et al., 2011). Participants were assigned to one of six conditions (care, loyalty, fairness, cheating, harm, and non-moral; *n* = 186 per condition) and, as in Study 2, they were asked to rate the degree to which three donation solicitations would motivate them to donate. The items for the care, loyalty, and non-moral conditions were identical to those used in Study 2 and the items for the fairness (e.g. 'Make sure the people affected by Hurricane Sandy get the assistance they deserve, #SandyDonate'), cheating (e.g.
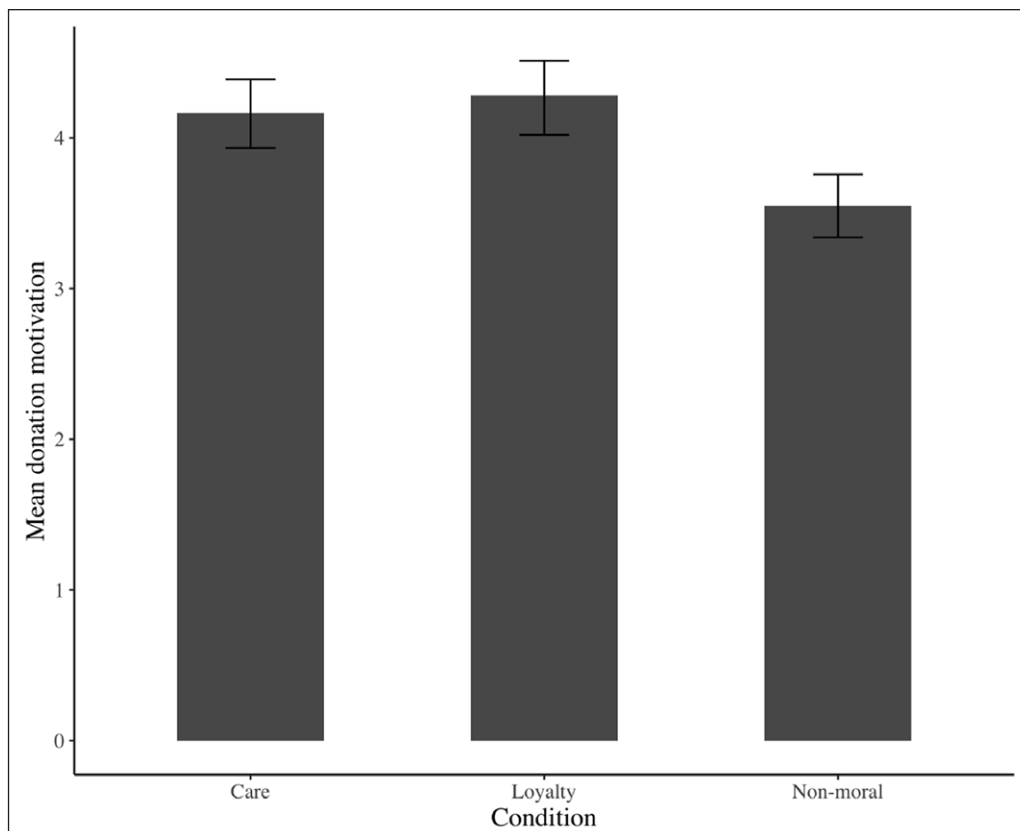

Figure 2: Mean donation motivation across conditions. Error bars represent bootstrapped 95%*CI*s.

'The people affected by Hurricane Sandy are being taken advantage of, please #donate now.') and harm (e.g. 'The people affected by Hurricane Sandy are suffering, please #donate now.') conditions were generated using the same processes described in Study 2.

Prior to conducting the planned hypothesis tests, Cronbach's Alpha was evaluated for each domain with the preregistered minimum inclusion threshold set to 0.60. A donation motivation index was then generated by averaging the individual item scores within each condition. To test the hypotheses that care and loyalty frames have stronger effects on donation motivation than fairness, cheating, harm, and non-moral frames, we conducted an analysis of variance with the donation motivation index as the dependent variable and condition as the independent variable and then generated planned comparisons comparing the effects of the care and loyalty conditions to each of the other conditions.

As in Study 2, we also tested the same effects on retweet motivation and we investigated the role of sense in these effects (See section Study 3 in Supplemental Materials for the results of these analyses). This study was preregistered via the Open Science Framework (https://osf.io/mpxrt/?view_only=fbbb005e70c046e59e3396ea501562b7) and approved by the USC IRB (ID UP-07-00393).

Results

As in Study 2, the care (α = 0.90), loyalty (α = 0.76), and control (α = 0.71) items demonstrated acceptable reliability, as did fairness (α = 0.74), cheating (α = 0.68) and harm (α = 0.79). Due to missing data, 81 participants were dropped from the analysis, yielding N = 1035. An ANOVA indicated substantial between-conditions variance in donation motivation, F (6,1029) = 988.8, $p < 0.001$. More importantly, planned-contrasts comparing the effects of

care and loyalty to all other conditions revealed a nearly perfect replication of the patterns identified in Study 1 (See **Figure 3**). However, one exception was observed: the effect of the harm condition was not significantly different from the effects of the care and loyalty conditions, (d = –0.03 and –0.29, respectively, ps = [0.10, 0.31]).

Discussion

Study 3 provides further experimental evidence that framing a donation solicitation with care or loyalty rhetoric increases perceived donation motivation compared to a non-moral frame. Importantly, this effect is congruent with the effects observed in Study 1 and it directly replicates the effects found in Study 2. Further, Study 3 provides partial evidence that care and loyalty frames have a stronger effect on donation motivation than other moral frames, as indicated by Study 1. However, the absence of reliable differences between the the care and loyalty and harm conditions contradicts this interpretation; while harm had a relatively small association with donation sentiment in Study 1, Study 3 indicates that harm frames are perceived as equivalently motivating, compared to care and loyalty frames. While this contradicts our expectations, it may also be important to acknowledge that MFT includes care and harm in the same superordinate category. Accordingly, it would not be entirely surprising if participants only weakly distinguished between care and harm frames, given their conceptual overlap.

Overall, however, this study provides additional evidence that care and loyalty frames are perceived as particularly strong motivators of donation. While the effects of these frames were not significantly different from that of harm, their positive difference from the effects of fairness and cheating suggest that the effects we have observed cannot be cleanly reduced to effects of moral frames in general, which
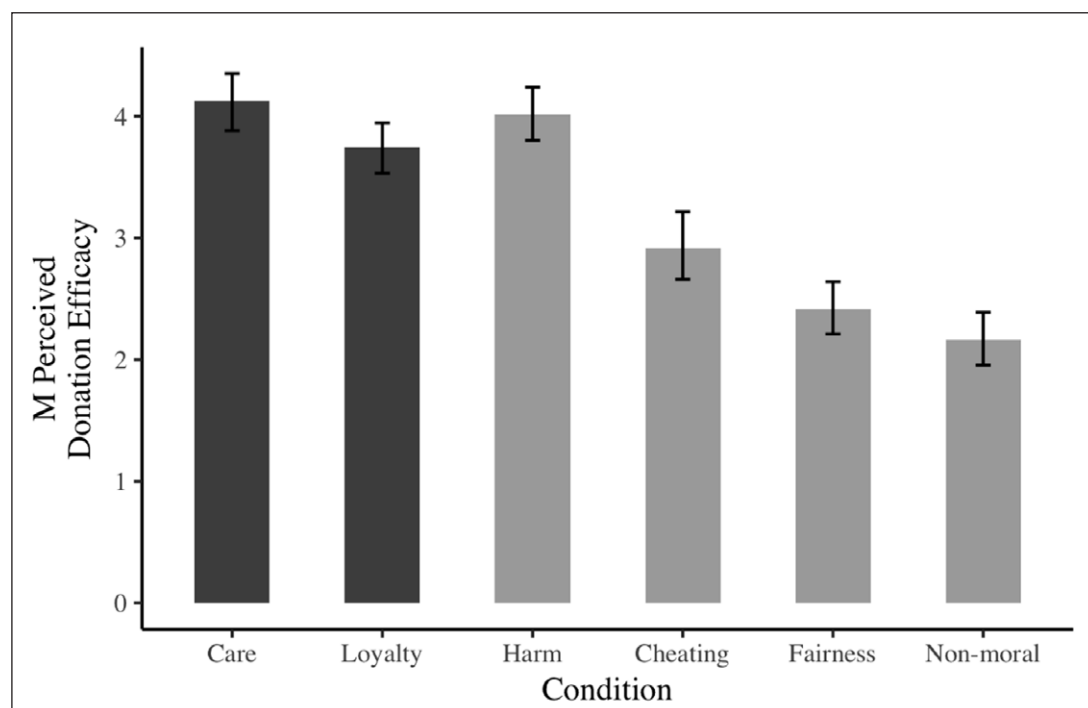

Figure 3: Mean donation motivation across moral and non-moral conditions. Error bars represent 95%*CI*.

is congruent with the hypotheses derived from Study 1. However, while Studies 1–3 indicate that care and loyalty frames are seen as relatively strong motivators of donation, none of these studies address whether these frames actually influence donation relevant behavior. That is, it could easily be the case that while people *believe* that these frames motivate donation, in reality, they have no such motivational power. Accordingly, in the next studies we investigate whether care and loyalty frames influence hypothetical (Studies 4 and 5) and real (Study 5) donation behavior.

Study 4

In the previous studies, we found compelling evidence that people associate charitable donation with care and loyalty (Study 1) and that solicitation frames that contain care and loyalty rhetoric are perceived as particularly strong motivators of charitable donation (Studies 2 & 3). In Study 4, we shift focus from perceived donation motivation to hypothetical donation amount. Whereas donation motivation targets participants beliefs about what kinds of frames motivate donation behavior, donation amount focuses on whether different frames motivate people to make larger or smaller donations. Further, with the aim of better understanding the mechanism driving the effects of loyalty on donation relevant constructs, here we also investigate whether loyalty rhetoric has an effect above and beyond that of care rhetoric. This is an important test, because it could be that framing a donation solicitation using both care and loyalty rhetoric could increase donation amount above and beyond using either frame in isolation. However, it could also be the case that the combined frame has no additional effect on donation amount. Finally, with the aim of better understanding the mechanism underlying the effect of loyalty framing on donation amount, we investigate the mediation of this effect by self-group overlap. Specifically, we test the hypothesis that loyalty framing is associated with increased hypothetical donation amount and also that this association is mediated by self-group overlap, a measure of identification with a given group.

Method

As in Study 3, participants (N = 930, % female = 0.42) were recruited from YourMorals.org. Sample size was determined via power analysis for an effect size of $d = 0.30$ with 80% power with an additional 30 participants collected to account for attrition, missing data, and attention check failure. As in Studies 2 and 3, participants were randomly assigned to one condition which manipulated moral framing. Specifically, the conditions used either care, loyalty, care + loyalty, cheating, or control frames. Cheating was chosen as the moral comparison case because it had a relatively strong effect in Study 1 and had the strongest effect, excluding care, loyalty, and harm, in Study 3. An alternative would have been to select harm as the comparison case, due to its effect size in Study 3. However, given the above noted possibility that care and harm frames may not be strongly differentiated in the current experimental paradigms, we used cheating instead.

The stimuli for the care + loyalty condition were created by combining the language used for the individual care and loyalty stimuli (e.g. 'Show compassion for the people affected by #HurricaneSandy, they need your loyalty and support. Please donate now. #SandyDonate'). The other stimuli were identical to those used in Studies 2 and 3. In the experiment, participants were asked to read all three tweets. On the next page, they were then asked to indicate 'how close or identified with the victims of Hurricane Sandy they feel' using a dynamic version of the conventional Inclusion of the Other in Self paradigm (Aron, Aron, & Smollan, 1992; Gómez et al., 2011). Participants were then asked to report how many dollars they thought they would donate using a continuous sliding scale with visible real values that ranged from 0 to 100.

Accordingly, three stages of analysis were planned and preregistered for this study. First, an ANOVA was estimated to determine whether there is significant between-condition variance; second, planned contrasts were implemented in order to determine (1) whether care, loyalty, and care + loyalty frames have stronger effects on hypothetical donation amount, compared to the other conditions and (2) whether the care + loyalty condition had stronger effects than the individual care and loyalty condition. Finally, we planned to estimate bootstrapped mediation model to determine whether the effect of loyalty framing on hypothetical donation amount is mediated by feelings of self-group overlap.

This study was preregistered via the Open Science Framework (https://osf.io/9cnk2/?view_only=baf7799b ca134531b14c4995dca7e25b) and approved by the USC IRB (ID UP-07-00393).

Results

Contrary to our expectations, an ANOVA in which the dependent variable was hypothetical donation amount and the independent variable was condition indicated only marginally significant ($p = 0.06$) between-condition variance. Further, examination of the condition means revealed no meaningful pattern (See **Table 3**; for box plots of each condition see **Figure 2** in Supplemental Materials). Accordingly, we did not proceed with the planned analyses.

Discussion

Contrary to our expectations, the current study found no evidence that moral framing a ects hypothetical donation amount. This is somewhat surprising, given that the previous studies provided consistent evidence that

Table 3: Study 4 condition means and standard deviations.

Donation condition	*M* Donation	SD
care	27.55	31.80
loyalty	31.09	34.29
care + loyalty	34.10	35.85
cheating	25.45	31.79
control	33.06	35.76

perceived donation motivation is associated with care and loyalty frames. While these null effects could have been caused by a measurement instrument (e.g. some arbitrary consequence of the sliding scale indicator), we believe it is more likely that perceived donation motivation and hypothetical donation amount are substantively different constructs. That is, it may simply be the case that people *believe* that care and loyalty frames motivate donation but that this belief has no direct connection to how much a person will donate in a hypothetical situation. Clearly, these results suggest that the effects of care and loyalty frames on charitable donation relevant constructs are not uni-dimensional. Further, these results highlight the fact that folk belief that a donation solicitation will be effective does not necessarily indicate that it will be.

Study 5

Together, Studies 1–3 indicate that people tend to frame donation sentiment with care and loyalty rhetoric and that they believe that donation solicitations framed with care and loyalty rhetoric are stronger motivations of donation, compared to both non-moral frames and frames that evoke other classes of moral concerns. In contrast, Study 4 suggests that perceived donation motivation may not be a reliable proxy for donation amount. However, because Study 4 focuses only on hypothetical donation amount, whether care or loyalty frames affect *actual* donation is unknown. Accordingly, in the current study, we offered participants an opportunity to make a real donation to the Greater New Orleans Foundation Tornado Relief Fund, which supports victims of the record-breaking series of tornadoes that hit the New Orleans area in February, 2017.

Method

Participants (N = 235,[7] % female = 0.54, Mean age = 33.70, SD age = 11.31) were recruited via Amazon Mechanical Turk and paid $0.50 for their participation. Participants were randomly assigned to one of three conditions (care, control, and loyalty). As in Young, Chakro, and Tom (2012), participants were told that our lab was considering offering future participants the opportunity to make a charitable donation at the end of studies. Participants were asked to read a donation solicitation and indicate how much of a hypothetical $20.00 bonus they would donate using a text entry box. All participants read solicitations with identical introductions:

> In February, 2017, New Orleans, Louisiana, was struck by seven tornadoes, one of which was the most powerful on record. The storm destroyed hundreds of homes, caused millions of dollars of damage, and the areas that were hit the hardest were among those that were most severely impacted by Hurricane Katrina in 2005.

To manipulate the moral frame, the donation solicitations also contained an additional component that varied between conditions. The components for care, loyalty, and control were, respectively:

> The people in these areas are still recovering from the February storm and they need your help and compassion. If you care about their well-being, please help the vulnerable by donating to the Greater New Orleans Foundation Tornado Relief Fund.
>
> The people in these areas are still recovering from the February storm and they need your help and compassion. If you care about your fellow Americans, please stand in solidarity with them by donating to the Greater New Orleans Foundation Tornado Relief Fund.
>
> The people in these areas are still recovering from the February storm. Please consider donating to the Greater New Orleans Foundation Tornado Relief Fund.

After completing the hypothetical donation task, participants were asked if they wanted to complete an additional task in order to earn a $1.00 bonus. If participants consented to completing the additional task, they then were asked to complete a filler task involving labeling the number of primary entities in two photographs, one containing multiple cats and another containing multiple vases. This filler task was assigned in order to make the bonus seem more legitimate and thus reduce the likelihood that participants might guess that we were interested explicitly in their donation behavior. Finally, after participants completed the bonus task, they were told that they could donate any portion ranging from 0 to 100% of their bonus to the same charitable cause they read about early in the study. Participants were then asked to indicate how much of their bonus in cents they wanted to donate using text entry. All data was collected during early April, 2017, within 90 days of the New Orleans tornadoes.

To test the hypotheses that care and loyalty frames increase hypothetical donation amount and real donation amount, two ANOVAs with planned comparisons were estimated, one with hypothetical donation amount and the other with real donation amount as the dependent variable and condition as the independent variable.

This study was preregistered via the Open Science Framework (https://osf.io/4z9v2/?view_only=fc1c2788e 11d4ad098d0326af6798e93) and approve by the USC IRB (ID UP-17-00078).

Results

Of 235 participants, 6 did not complete the hypothetical donation task and were dropped from further analysis. Further, across the care, loyalty, and control conditions 7 participants (2, 3, and 2, respectively) opted not to complete the bonus (i.e. experimental) portion of the study and an additional 10 participants did not complete the real donation portion of the survey, yielding an N of 218. An ANOVA with hypothetical donation as the dependent variable indicated significant between-condition variance $F(3, 226) = 139.2$, $p < 0.001$; however, planned comparisons revealed that neither the care ($M = 6.31$, $SD = 3.58$) nor the loyalty ($M = 5.97$, $SD = 2.90$)

conditions induced hypothetical donations that were significantly higher than those reported by participants in the control condition (M = 5.50, SD = 2.76), *diff* = 0.80, SE = 0.50, t = 1.61, p = 0.19 and *diff* = 0.47, SE = 0.50, t = 0.95, p = 0.53, respectively (See **Table 4** for condition means and standard deviations). Further, given this data's moderate violations of ANOVA's parametric assumptions of normality and equal variance, we conducted an additional ANOVA using robust procedures (Mair & Wilcox, 2016) that relax these assumptions. Specifically, we conducted a robust one-way ANOVA, which relied on 20% trimmed means, rather than sample means, and relaxed the assumption of homogeneous variances across groups. In contrast to the non-robust ANOVA, the robust ANOVA indicated no strong evidence for variance across groups, F (2, 87.59) = 1.40, p = 0.25. Accordingly, the current study replicated the null effects of care and loyalty framing on hypothetical donation observed in Study 4.

Further, while pre-registered non-robust analyses with *real* donation as the dependent variable indicated

a significant effect of care framing, relative to control, a robust ANOVA yielded no evidence of between-condition variation. More specifically, our pre-registered analyses identified significant between-condition variation F (3, 215) = 50.48, p < 0.001 and subsequent planned comparisons indicated that participants assigned to the care condition (*M = 37.49, SD = 40.10*) donated significantly more money than those assigned to the control condition (M = 23.92, SD = 33.06), *diff* = 13.57, SE = 5.83, 95%CI = [0.57, 26.67], t = 2.33, p = 0.04. However, there was no observed statistical difference in donation amounts between participants assigned to the loyalty (M = 24.56, SD = 31.75) and control conditions, *diff* = 0.64, SE = 5.82, 95%CI = [−12.31, 13.60], t = .111, p = 0.99. See **Figure 4**. Again, due to moderate violations of parametric assumptions, we also conducted robust a robust ANOVA. Under this model, there was no significant evidence of between-condition variation, F (2, 82.76) = 1.61, p = 0.20.

Discussion

In the current study, we tested the hypotheses that care and loyalty framing increase both hypothetical and actual donation behavior. As in Study 4, moral framing effects on hypothetical donation could not be distinguished from zero. Further, robust ANOVA estimates provided no reliable evidence for an effect on actual donation. This results indicate a possible disjunction between people's *perceptions* of framing efficacy and *actual* framing efficacy. While people reported thinking that care and loyalty frames would increase donation behavior (Studies 2 and 3), we find no evidence for any such effects when people are actually exposed to these frames.
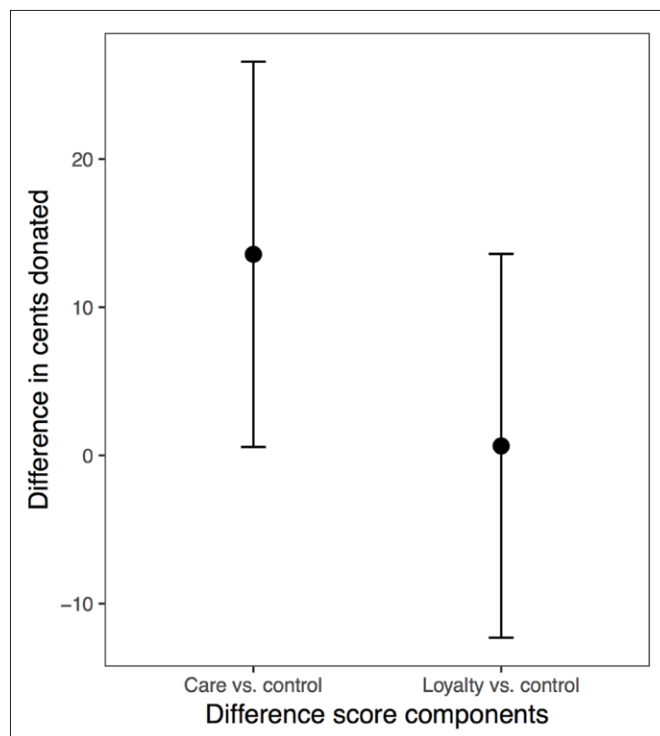
Table 4: Study 5 Hypothetical and Real Donations.

Condition	Hypothetical Donation	Real Donation
Care	6.31 (3.58)	37.49 (40.10)
Loyalty	5.97 (2.90)	24.56 (31.75)
Control	5.50 (2.76)	23.92 (33.06)

Hypothetical donations ranged from $0.00–$20.00 and real donations ranged from $0.00–$1.00. Standard deviations are shown in parentheses.

Figure 4: Differences in cents donated. Points represent point estimates of the difference in sense estimated between moral condition (X axis) and the control condition. Error bars represent 95%CI.

General Discussion

Across five studies, we investigated the effects of moral framing on constructs relevant to charitable donation. To motivate this exploration, we conducted theoretically constrained exploratory analysis of a corpus of tweets posted during Hurricane Sandy. By applying NLP to naturally generated data, we were able to derive hypotheses from observations of real-world dynamics. Specifically, these analyses revealed that tweets containing donation sentiment were much more strongly associated with the moral domains of care and loyalty, compared to other moral domains. These observations raised two questions: one, is this association between care and loyalty frames and charitable donation sentiment reliable; and two, what are the psychological implications of this association?

Across a sequence four pre-registered experiments, we then sought to determine the extent to which these real-world, observational effects correspond to three psychological phenomena associated with charitable donation: perceived donation motivation, hypothetical donation behavior, and real donation behavior.

In studies 2 and 3, we observed a pattern of moral framing effects on donation motivation that mostly fit with the effects found in Study 1, such that solicitations that contained care or loyalty frames were seen as more likely to motivate donation than non-moral controls and solicitations containing fairness and cheating frames. However, in contrast to Study 1, there was no difference between the effects of care, loyalty, and harm frames on perceived donation motivation. One explanation for this may be that whereas when judging the motivational power of donation solicitations, people do not make strong distinctions between care and harm frames; however, when actually trying to motivate donations during a real crisis, perhaps people tend to favor care frames over harm frames.

Interestingly, while we found substantial congruence between the associations between moral sentiment and donation sentiment and moral frames and donation motivation, no effects of moral frames were found on either hypothetical donation or real donation. That is, people tended to use care and loyalty frames while discussing donation and experimental subjects perceived these frames as particularly efficacious. However, participants experimentally exposed to these frames, relative to control conditions, neither indicated that they would donate more nor actually donated more. This suggests that people, or at least lay-people, may not have particularly reliable insight into the kinds of framing that can motivate charitable donation.

Ultimately, the points of congruence between Study 1 and the subsequent experimental studies indicate that the semantic effects observed in our social media analysis may indeed have psychological implications. Across two studies, participants reported believing that solicitations containing care and loyalty frames would be stronger donation motivators, compared to those containing non-moral and some other moral frames. This is consistent with our assumption that when people tweet about charitable donation they are often attempting to motivate others to donate. That is, it may be that the patterns we observed in Study 1 were partially driven by people's beliefs about what makes an effective tweet. However, when comes to actual charitable donation, it seems that people's perceptions may not be entirely accurate.

Further, this research demonstrates that while careful analysis of social media data can yield robust insights into psychological processes, researchers need to be careful in how they operationalize these processes. The stark absence of evidence for donation effects highlights the importance and difficulty of clearly operationalizing and probing the constructs targeted in natural language analyses. Without our subsequent experimental studies, it would simply not have been possible to determine what the effects observed in study 1 indicated. Under the operationalization of donation motivation, we found largely corroborating effects across two experimental studies. However, under the constructs of hypothetical and real donation, we observed no such effects. By pairing exploratory language analysis with confirmatory experiments, we were able to generate novel hypothesis, but then also winnow these hypothesis down.

Accordingly, we believe that these results highlight the importance of subjecting interpretations of social media analyses to rigorous experimental testing. Operationalizing psychological variables with natural language constructs is difficult and messy. A measure that seems to indicate one thing, may very well indicate something entirely different. Without controlled experimentation, it is practically impossible to interpret the psychological implications of an observational social media study. When paired with experimental methods, however, social media analysis can provide a point of real-world contact that is often otherwise prohibitively expensive for social psychology research. Our view is that in many cases combining social media analysis and laboratory experimentation can afford a greater degree of overall external and construct validity compared to either of these paradigms in isolation.

Data Accessibility Statements

All data reported in this paper is publicly accessible at https://osf.io/crdsj/.

Additional Files

The additional file for this article can be found as follows:

- **Supplementary Materials file.** Supplementary analyses for Studies 1–4. DOI: https://doi.org/10.1525/collabra.129.s1

Notes

containing 100 billion words (available at https://code.google.com/p/word2vec/. To represent each moral domain, we use the same four-word sets used in Garten et al. (2017) (See 1).

[3] *Cohen's d* calculated as $\frac{mean_i}{SD_i}$ where $mean_i$ and SD_i are the model estimates of the mean and standard deviation for foundation *i*. More specifically, SD_i is calculated based on model estimates of the standard error, thus $SD_i = SE_i * sqrt(n)$.

[4] Note: To calculate SD_{pooled}, the mean estimated standard errors of the means from the random intercepts GLMs were converted to mean SDs using the formula $SE * \sqrt{n}$. In hierarchical models, determining the true degrees of freedom is generally not possible, as it lies somewhere between the degrees of freedom for the fully independent model (e.g. the total number moral values estimates – 2) and the degrees of freedom for a fully pooled model (e.g. the total number tweets). We took the conservative approach, and set *n* to *51,928*, the number of tweets in each model.

[5] Because our hypotheses are directional, we pre-registered one-tailed t-tests for this analysis. However, subsequent null hypothesis tests are all two-tailed, regardless of whether the hypothesis is directional.

[6] *Cohen's d* calculated using the 'effsize' R package version 0.7.1.

[7] The planned N was 225; however, an additional 10 participants were collected due to an error in automated data collection with Amazon Mechanical Turk.

Funding Informations

Competing Interests
The authors have no competing interests to declare.

Author Contributions
All authors contributed to methodological and experimental design. Boghrati and Dehghani were responsible for data acquisition and pre-processing. Hoover implemented and interpreted all statistical analyses. Hoover and Johnson drafted the initial manuscript and all authors contributed to critical revision and final approval for submission.

References

Amato, P. R. (1985). An investigation of planned helping behavior. *Journal of Research in Personality*, *19*(2), 232–252. DOI: https://doi.org/10.1016/0092-6566(85)90031-5

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *Journal of personality and social psychology*, *63*(4), 596. DOI: https://doi.org/10.1037/0022-3514.63.4.596

Ashar, Y. K., Andrews-Hanna, J. R., Yarkoni, T., Sills, J., Halifax, J., Dimidjian, S., & Wager, T. D. (2016). Effects of compassion meditation on a psychological model of charitable donation. *Emotion*, *16*(5), 691. DOI: https://doi.org/10.1037/emo0000119

Baumert, A., Gollwitzer, M., Staubach, M., & Schmitt, M. (2011). Justice sensitivity and the processing of justice-related information. *European Journal of Personality*, *25*(5), 386–397. DOI: https://doi.org/10.1002/per.800

Baumert, A., Thomas, N., & Schmitt, M. (2012). Justice sensitivity as resource or risk factor in civic engagement. *Restoring Civil Societies: The Psychology of Intervention and Engagement Following Crisis*, 17–37. DOI: https://doi.org/10.1002/9781118347683.ch2

Bekkers, R. (2006). Traditional and health-related philanthropy: The role of resources and personality. *Social psychology quarterly*, *69*(4), 349–366. DOI: https://doi.org/10.1177/019027250606900404

Bekkers, R., & Schuyt, T. (2008). And who is your neighbor? explaining denominational differences in charitable giving and volunteering in the netherlands. *Review of Religious Research*, 74–96. Retrieved from: http://www.jstor.org/stable/20447529

Bekkers, R., & Wiepking, P. (2011). A literature review of empirical studies of philanthropy: Eight mechanisms that drive charitable giving. *Nonprofit and Voluntary Sector Quarterly*, *40*(5), 924–973. DOI: https://doi.org/10.1177/0899764010380927

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, *20*(3), 351–368. DOI: https://doi.org/10.1093/pan/mpr057

Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799. DOI: https://doi.org/10.1080/01621459.1976.10480949

Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., & Mihalcea, R. (2015). Values in words: Using language to evaluate and understand personal values. In: *International aaai conference on web and social media (icwsm)*, 31–40.

Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Van Bavel, J. J. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*. Retrieved from: http://www.pnas.org/content/early/2017/06/20/1618923114.abstract. DOI: https://doi.org/10.1073/pnas.1618923114

Branscombe, N. R., Spears, R., Ellemers, N., & Doosje, B. (2002). Intragroup and intergroup evaluation effects on group behavior. *Personality and Social Psychology Bulletin*, *28*(6), 744–753. DOI: https://doi.org/10.1177/0146167202289004

Bryant, W. K., Jeon-Slaughter, H., Kang, H., & Tax, A. (2003). Participation in philanthropic activities: Donating money and time. *Journal of Consumer Policy*, *26*(1), 43–73. DOI: https://doi.org/10.1023/A:1022626529603

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's mechanical turk a new source of inexpensive, yet high-quality, data? *Perspectives on psychological science*, *6*(1), 3–5. DOI: https://doi.org/10.1177/1745691610393980

Castillo, M., Petrie, R., & **Wardell, C.** (2014). Fundraising through online social networks: A field experiment on peer-to-peer solicitation. *Journal of public economics*, *114*, 29–35. DOI: https://doi.org/10.1016/j.jpubeco.2014.01.002

Chang, C. T., & **Lee, Y. K.** (2009). Framing charity advertising: Influences of message framing, image valence, and temporal framing on a charitable appeal. *Journal of Applied Social Psychology*, *39*(12), 2910–2935. DOI: https://doi.org/10.1111/j.1559-1816.2009.00555.x

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297–334. DOI: https://doi.org/10.1007/BF02310555

Das, E., Kerkhof, P., & **Kuiper, J.** (2008). Improving the effectiveness of fundraising messages: The impact of charity goal attainment, message framing, and evidence on persuasion. *Journal of Applied Communication Research*, *36*(2), 161–175. DOI: https://doi.org/10.1080/00909880801922854

Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., Graham, J., et al. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, *145*(3), 366. DOI: https://doi.org/10.1037/xge0000139

Dietz, R., Druart, T., & **Edge Research.** (2015). *Donor engagement study: Aligning nonprofit strategy with donor preferences.* Association of Fundraising Professionals. Retrieved from: https://afpmadison.afpnet.org/files/AFP%20Chapters/DC1/Abila%20DonorEngagementStudy.pdf.

Farmer, S. M., & **Fedor, D. B.** (2001). Changing the focus on volunteering: An investigation of volunteers' multiple contributions to a charitable organization. *Journal of Management*, *27*(2), 191–211. DOI: https://doi.org/10.1177/014920630102700204

Fielding, D., Knowles, S., & **Robertson, K.** (2017). *When does it matter how you ask? cross-subject heterogeneity in framing effects in a charitable donation experiment* (Tech. Rep.). Retrieved from: http://www.otago.ac.nz/economics/otago633963.pdf.

Freeman, D., Aquino, K., & **McFerran, B.** (2009). Overcoming beneficiary race as an impediment to charitable donations: Social dominance orientation, the experience of moral elevation, and donation behavior. *Personality and Social Psychology Bulletin*, *35*(1), 72–84. DOI: https://doi.org/10.1177/0146167208325415

Garten, J., Hoover, J., Johnson, K. M., Boghrati, R., Iskiwitch, C., & **Dehghani, M.** (2017). Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis. *Behavior Research Methods*, 1–18. DOI: https://doi.org/10.3758/s13428-017-0875-9

Gómez, A., Brooks, M. L., Buhrmester, M. D., Vázquez, A., Jetten, J., & **Swann, W. B., Jr.** (2011). On the nature of identity fusion: insights into the construct and a new measure. *Journal of personality and social psychology*, *100*(5), 918. DOI: https://doi.org/10.1037/a0022642

Graham, J. (2015). Explaining away differences in moral judgment: Comment on gray and keeney (2015). *Social Psychological and Personality Science*, *6*(8), 869–873. DOI: https://doi.org/10.1177/1948550615592242

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P., & **Ditto, P. H.** (2012). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, *47*, 55–130. DOI: https://doi.org/10.1016/B978-0-12-407236-7.00002-4

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. DOI: https://doi.org/10.1037/a0021847

Gray, K., & **Keeney, J. E.** (2015a). Disconfirming moral foundations theory on its own terms: Reply to graham (2015). *Social Psychological and Personality Science*, *6*(8), 874–877. DOI: https://doi.org/10.1177/1948550615592243

Gray, K., & **Keeney, J. E.** (2015b). Impure or just weird? scenario sampling bias raises questions about the foundation of morality. *Social Psychological and Personality Science*, *6*(8), 859–868. DOI: https://doi.org/10.1177/1948550615592241

Guskin, E., & **Hitlin, P.** (2012). Hurricane sandy and twitter. *Pew Research Journalism Project.* Retrieved from: http://www.journalism.org/2012/11/06/hurricane-sandy-and-twitter/.

Hafer, C. L. (2000). Do innocent victims threaten the belief in a just world? evidence from a modified stroop task. *Journal of personality and social psychology*, *79*(2), 165. DOI: https://doi.org/10.1037/0022-3514.79.2.165

Haidt, J. (2003). The moral emotions. In: Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (Eds.), *Series in affective science. handbook of affective sciences*, 852–870. Oxford: Oxford University Press. Retrieved from: http://psycnet.apa.org/record/2009-07773-044.

Holmes, J. G., Miller, D. T., & **Lerner, M. J.** (2002). Committing altruism under the cloak of self-interest: The exchange fiction. *Journal of experimental social psychology*, *38*(2), 144–151. DOI: https://doi.org/10.1006/jesp.2001.1494

Le, Q. V., & **Mikolov, T.** (2014). Distributed representations of sentences and documents. *arXiv preprint arXiv:1405.4053.* Retrieved from: https://arxiv.org/abs/1405.4053.

Levine, M., Cassidy, C., & **Jentzsch, I.** (2010). The implicit identity effect: Identity primes, group size, and helping. *British Journal of Social Psychology*, *49*(4), 785–802. DOI: https://doi.org/10.1348/014466609X480426

Levine, M., & **Crowther, S.** (2008). The responsive bystander: how social group membership and group size can encourage as well as inhibit bystander intervention. *Journal of personality and social psychology*, *95*(6), 1429. DOI: https://doi.org/10.1037/a0012634

Mair, P., & **Wilcox, R.** (2016). *Robust statistical methods in r using the wrs2 package* (Tech. Rep.). Technical report, Harvard University. Retrieved from: https://

cran.r-project.org/web/packages/WRS2/vignettes/WRS2.pdf.

Martin, J. A. (2013). Disasters and donations: The conditional effects of news attention on charitable giving. *International Journal of Public Opinion Research, 25*(4), 547–560. DOI: https://doi.org/10.1093/ijpor/eds044

Meer, J. (2011). Brother, can you spare a dime? peer pressure in charitable solicitation. *Journal of public economics, 95*(7–8), 926–941. DOI: https://doi.org/10.1016/j.jpubeco.2010.11.026

Mikolov, T., Yih, W.-T., & **Zweig, G.** (2013). Linguistic regularities in continuous space word representations. In: *Annual conference of the north american chapter of the association for computational linguistics: Human language technologies*, 746–751.

Nilsson, A., Erlandsson, A., & **Västfjäll, D.** (2016). The congruency between moral foundations and intentions to donate, self-reported donations, and actual donations to charity. *Journal of Research in Personality, 65*, 22–29. DOI: https://doi.org/10.1016/j.jrp.2016.07.001

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Seligman, M. E., et al. (2015). Automatic personality assessment through social media language. *Journal of personality and social psychology, 108*(6), 934. DOI: https://doi.org/10.1037/pspp0000020

Pennebaker, J. W. (2011). The secret life of pronouns. *New Scientist, 211*(2828), 42–45. DOI: https://doi.org/10.1016/S0262-4079(11)62167-2

Penner, L. A., Dovidio, J. F., Piliavin, J. A., & **Schroeder, D. A.** (2005). Prosocial behavior: Multilevel perspectives. *Annu. Rev. Psychol., 56*, 365–392. DOI: https://doi.org/10.1146/annurev.psych.56.091103.070141

Piferi, R. L., Jobe, R. L., & **Jones, W. H.** (2006). Giving to others during national tragedy: The effects of altruistic and egoistic motivations on long-term giving. *Journal of Social and Personal Relationships, 23*(1), 171–184. DOI: https://doi.org/10.1177/0265407506060185

Publication search | moralfoundations.org. (n.d.). http://www.moralfoundations.org/publications (Accessed: 2017-12-5).

Reber, R., & **Schwarz, N.** (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and cognition, 8*(3), 338–342. DOI: https://doi.org/10.1006/ccog.1999.0386

Reber, R., Winkielman, P., & **Schwarz, N.** (1998). Effects of perceptual fluency on affective judgments. *Psychological science, 9*(1), 45–48. DOI: https://doi.org/10.1111/1467-9280.00008

Reed, P. B., & **Selbee, L. K.** (2002). Is there a distinctive pattern of values associated with giving and volunteering? the canadian case. In: *32nd arnova conference, montreal, canada.* Retrieved from: http://www3.carleton.ca/casr/Volunteers%20patterns.pdf.

Ruths, D., & **Pfeffer, J.** (2014). Social media for large studies of behavior. *Science, 346*(6213), 1063–1064. DOI: https://doi.org/10.1126/science.346.6213.1063

Sagi, E., & **Dehghani, M.** (2014). Measuring moral rhetoric in text. *Social Science Computer Review, 32*(2), 132–144. DOI: https://doi.org/10.1177/0894439313506837

Schein, C., & **Gray, K.** (2017). The theory of dyadic morality: Reinventing moral judgment by redefining harm. *Personality and Social Psychology Review.* DOI: https://doi.org/10.1177/1088868317698288

Schervish, P. G., & **Havens, J. J.** (2002). The boston area diary study and the moral citizenship of care. *Voluntas: International Journal of Voluntary and Nonprofit Organizations, 13*(1), 47–71. DOI: https://doi.org/10.1023/A:1014758113132

Schuyt, T., Bekkers, R., & **Smit, J.** (2010). The philanthropy scale: A sociological perspective in measuring new forms of pro social behaviour. *Social Work & Society, 8*(1), 121–135. Retrieved from: http://www.socwork.net/sws/article/view/28/75.

Slovic, P. (2010). If i look at the mass i will never act: Psychic numbingpsychic numbing and genocidegenocide. In *Emotions and risky technologies*, 37–59. Springer. DOI: https://doi.org/10.1007/978-90-481-8647-1_3

Small, D. A., Loewenstein, G., & **Slovic, P.** (2007). Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organizational Behavior and Human Decision Processes, 102*(2), 143–153. DOI: https://doi.org/10.1016/j.obhdp.2006.01.005

Sussman, A. B., Sharma, E., & **Alter, A. L.** (2015). Framing charitable donations as exceptional expenses increases giving. *Journal of Experimental Psychology: Applied, 21*(2), 130. DOI: https://doi.org/10.1037/xap0000047

Tausczik, Y. R., & **Pennebaker, J. W.** (2010). The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology, 29*(1), 24–54. DOI: https://doi.org/10.1177/0261927X09351676

Todd, S. J., & **Lawson, R. W.** (1999). Towards a better understanding of the financial donor: An examination of donor behaviour in terms of value structure and demographics. *International Journal of Nonprofit and Voluntary Sector Marketing, 4*(3), 235–244. DOI: https://doi.org/10.1002/nvsm.76

Tukey, J. W. (1977). *Exploratory data analysis.* Reading, Mass.

Tusche, A., Böckler, A., Kanske, P., Trautwein, F.-M., & **Singer, T.** (2016). Decoding the charitable brain: Empathy, perspective taking, and attention shifts differentially predict altruistic giving. *Journal of Neuroscience, 36*(17), 4719–4732. DOI: https://doi.org/10.1523/JNEUROSCI.3392-15.2016

Van Lange, P. A., Bekkers, R., Schuyt, T. N., & **Vugt, M. V.** (2007). From games to giving: Social value orientation predicts donations to noble causes. *Basic and Applied Social Psychology, 29*(4), 375–384. DOI: https://doi.org/10.1080/01973530701665223

Västfjäll, D., Slovic, P., Mayorga, M., & **Peters, E.** (2014). Compassion fade: Affect and charity are greatest for a single child in need. *PloS one, 9*(6), e100115. DOI: https://doi.org/10.1371/journal.pone.0100115

Vijayarani, S., Ilamathi, M. J., & **Nithya, M.** (2015). Preprocessing techniques for text mining-an overview. *International Journal of Computer Science & Communication Networks*, *5*(1), 7–16. Retrieved from: http://www.ijcscn.com/Documents/Volumes/vol5issue1/ijcscn2015050102.pdf.

Virtual Social Media Working Group and DHS First Responders Group. (2013). *Lessons learned – social media and hurricane sandy.* Online. Retrieved from: https://www.dhs.gov/publication/lessons-learned-social-media-hurricane-sandy.

Wilhelm, M. O., & **Bekkers, R.** (2010). Helping behavior, dispositional empathic concern, and the principle of care. *Social Psychology Quarterly*, *73*(1), 11–32. DOI: https://doi.org/10.1177/0190272510361435

Yamamoto, Y., Yoo, S., & **Matsui, Y.** (2015). Effects of empathy on fund-raising activities on behalf of victims of the 2011 great east japan earthquake, focusing on the residents in the south kanto area. *The Japanese journal of psychology*, *85*(6), 590–595. DOI: https://doi.org/10.4992/jjpsy.85.13330

Young, L., Chakroff, A., & **Tom, J.** (2012). Doing good leads to more good: The reinforcing power of a moral self-concept. *Review of Philosophy and Psychology*, *3*(3), 325–334. DOI: https://doi.org/10.1007/s13164-012-0111-6

Zagefka, H., Noor, M., Brown, R., de Moura, G. R., & **Hopthrow, T.** (2011). Donating to disaster victims: Responses to natural and humanly caused events. *European journal of social psychology*, *41*(3), 353–363. DOI: https://doi.org/10.1002/ejsp.781

Zlatev, J. J., & **Miller, D. T.** (2016). Selfishly benevolent or benevolently selfish: When self-interest undermines versus promotes prosocial behavior. *Organizational Behavior and Human Decision Processes*, *137*, 112–122. DOI: https://doi.org/10.1016/j.obhdp.2016.08.004