Experimental Economics (2023) 26:1115–1141
https://doi.org/10.1007/s10683-023-09811-z

Using social norms to explain giving behavior

Catherine C. Eckel[1] · Hanna G. Hoover[2] · Erin L. Krupka[3] · Nishita Sinha[4] · Rick K. Wilson[5]

Received: 13 May 2022 / Revised: 12 September 2023 / Accepted: 13 September 2023 /
Published online: 6 October 2023

Abstract

Transfers of resources in dictator games vary significantly by the characteristics of recipients. We focus on social norms and demonstrate that variation in the recipient changes both giving and injunctive norms and may offer an explanation for differences in giving. We elicit generosity using dictator games, and social norms using incentivized coordination games, with two different recipient types: an anonymous student and a charitable organization. A within-subjects design ensures that other factors are held constant. Our results show that differences in giving behavior are closely related to differences in social norms of giving across contexts. Controlling for individual differences in beliefs about the norm, subjects do not weight compliance with the norms in the student recipient or charity recipient dictator game differently. These results suggest that the impact of context on giving co-occurs with an impact on social norms.

Keywords Dictator game · Social norms · Charitable giving

✉ Erin L. Krupka
ekrupka@umich.edu

Catherine C. Eckel
ceckel@tamu.edu

Hanna G. Hoover
h.hoover@northeastern.edu

Nishita Sinha
nishita.sinha@tamu.edu

Rick K. Wilson
rkw@rice.edu

[1] Department of Economics, Texas A &M University, College Station, TX, USA

[2] Dukakis Center for Urban and Regional Policy, Northeastern University, Boston, MA, USA

[3] School of Information, University of Michigan, Ann Arbor, MI, USA

[4] Department of Agricultural Economics, Texas A &M University, College Station, TX, USA

[5] Department of Political Science, Rice University, Houston, TX, USA

JEL Classification D64 · C91

1 Introduction

There are a number of important reasons to understand why people donate to charities. Critically, a deeper understanding can inform mechanism design and theory. Many experimental papers use dictator games to understand individual and institutional factors that drive donations to charities (e.g. Andreoni & Miller, 2002; Vesterlund, 2016). However, recent evidence suggests that injunctive norms (what one "ought" to do) are critical determinants of choice in charitable contributions as well as in dictator games (Drouvelis & Marx, 2021; Krupka & Weber, 2013). Since empirical work has focused almost exclusively on the descriptive norm (i.e. what others are doing), we know very little about the role that injunctive norms (i.e. what one "ought" to do) may play in determining charitable donation behavior.

The importance of norms was first highlighted in studies that sought to understand the pronounced variation in giving that is observed in response to changes in institutional or procedural factors. Most prominently, Bardsley (2008) and List (2007) showed that a small change in procedure—allowing decision makers to take from, as well as give to, a counterpart—dramatically reduced giving. Krupka and Weber (2013) developed a methodology for measuring norms of giving in the two dictator games, and showed that the change in transfers is caused by a change in norms, and norms are strongly impacted by the change in procedure. When dictators can take as well as give, it changes how subjects perceive the game and alters the attendant norms of behavior. Though not directly identified in the same manner as Krupka and Weber (2013) do, previous literature exploring behavioral differences between the standard dictator game and one where the recipient is a charity have posited that the identity of the recipient alters the context and associated social norms (Grossman & Eckel, 2015).[1]

In this paper we combine two streams of research to explore social norms as a potential source of behavioral change in giving across contexts. Using a within-subjects design with a sample of college students, we conduct dictator-game experiments with two recipient types: an anonymous individual (who is also a study participant) and a charitable organization. We identify the impact of recipient type (person or charity) on norms of giving in the dictator game by using the Krupka and Weber (2013) method. The within-subject design allows us to control for important sources of variation and address some concerns regarding demand effects (Bardsley, 2008; List, 2007; Zizzo, 2010).

We show that the injunctive social norms for giving to a charity recipient and a student recipient differ and that behavior in these dictator games tracks those norms. The injunctive norms associated with giving to a student recipient is such that a 50% split with the student recipient is the most appropriate action while transferring

[1] Specifically, Grossman and Eckel (2015) write that "...context makes more apparent the social norm regarding giving".

100% of the endowment to the charity is the most appropriate action. We find evidence that responses to charity norms have lower variance; we interpret this as subjects being more certain about charity norms. Furthermore, controlling for individual differences in beliefs about the norm, subjects do not weight differently compliance with the norms in the student recipient or charity recipient dictator game.

We make contributions to research on charitable donations by eliciting the injunctive norms for a charity as well as the standard dictator game setting. Surprisingly, we know of no other paper which elicits the injunctive norm for a charity recipient and very few papers that use within-subject design to compare norms and behavior when the recipient is an individual or charity.[2] The within-subject design is useful because there is much prior work on norms for transfers in the "standard" dictator game making it a meaningful benchmark for our study. We also extend the literature that tests the impact of transfers in the dictator game based on recipient type; we find similar effect sizes to those observed in charity recipient dictator games using a between-subject setting. This may be informative for future researchers looking to investigate differences in dictator-giving.

We also contribute to the social norms literature by examining norm uncertainty. Prior work looking at norm compliance has not addressed the role of norm certainty as a potential explanation for variation in behavior across settings. We contribute in this paper by showing that controlling for differences in norm certainty across contexts can account for some of the observed treatment effects of context on norms and behavior. In our discussion section we also discuss the limitations on our ability to distinguish between injunctive norms and intrinsic/instrumental concerns.

2 Motivation

Most of the exploration into motives for donations focuses on distinguishing between the extent to which donations are motivated by a concern for others or by a concern for self (Vesterlund, 2016). Motives for giving include pure altruism (Ottoni-Wilhelm et al., 2017), income effects (Cherry & Shogren, 2008; Cherry et al., 2002), and a variety of factors that can produce warm glow, such as reciprocity (Croson, 2008; Falk, 2007), signaling (Andreoni & Bernheim, 2009; Ariely et al., 2009; Bracha et al., 2011; Glazer & Konrad, 1996), social pressure (Andreoni et al., 2017; DellaVigna et al., 2012), and a desire to comply with social norms of giving (Krupka & Croson, 2016), among others.[3]

A social norms framework assumes that information about the context, along with others' normative beliefs and expectations, is mapped into a set of socially

[2] Umer et al. (2022) conduct a meta-analysis of laboratory dictator games where the recipient is either a charity or student. They found only 3 studies that reported money shared with charities as well as students (i.e. used a within-subjects design such as our current study).

[3] Vesterlund (2016) offers a more thorough review.

acknowledged rules for behavior which the actor desires to comply with.[4] As such, norms take features of the context such as "need" or "desert" and map social evaluations and expectations to behavior. Beliefs about deservingness, need, merit and entitlement operate through norms (Drouvelis & Marx, 2021; Krupka & Weber, 2013).[5,6] The empirical evidence shows that dictator behavior is responsive to these context features. For example, Fong (2007) found that donors' beliefs about whether their recipients were poor because of bad luck, lack of effort, or both had large and robust effects on dictator transfers in the expected direction. Engel (2011) conducted a meta-analysis of dictator game behavior and found that if the recipient was "deserving" then they received 8% more. More recently, Umer et al. (2022) conducted a separate meta-analysis, focused only on laboratory studies and studies that employed either student or charity recipients, and found that dictators transferred on average 21% of the endowment to a student recipient and 45% to charity recipient.[7] Whether the analysis focuses on recipients with particular characteristics or charities, authors note that features of the experiment likely affect the norm for behavior as well as giving behavior directly (Engel, 2011; Fong, 2007; Fong & Luttmer, 2009).

Recent evidence suggests that the desire to comply with norms is chief among motives for making transfers to individuals and to charities. Krupka and Weber (2013) show that norms matter for explaining dictator transfers to individuals in a variety of settings. Drouvelis and Marx (2021) tease out multiple, different motives that are likely present in any one decision to donate to a charity. They show that norm compliance is among the most prominent motives.[8] Drouvelis and Marx

[4] Because norms are rules about what is expected of a person "in a given situation," there is not a "pure" norm of appropriateness absent contextual cues. Bicchieri (2016), argues that context elicits norms—that is, the features of a concrete situation are causally relevant to the production of the "psychological process," associated with norms. In this "psychological process," features of the context act to select a normative "script" from memory. The normative script is a "... stylized, stereotyped sequence of actions that are appropriate in this context, and it defines actors and roles. [...] Scripts are the basis of understanding and making sense of events, as they embed knowledge relevant to the present situation". Bicchieri (2016, p. 94).

[5] Deservingness is distinct from need, entitlement and from merit. Deservingness, sometimes referenced as fairness to outputs, is traditionally based on relative equity such that one's "outcomes should be consistent with inputs" (Bendapudi et al., 1996) and the target must be seen as personally responsible for the actions that lead to particular outcomes (Andreoni & Sanchez, 2020; Feather, 1999). Need is typically seen as one's relative outcomes compared to others (Yaari & Bar-Hillel, 1984). Entitlements exist without the personal responsibility criterion of deservingness; rather they articulate outcomes based on a person's rights and justice principles (Feather, 1999). Merit is a consequentialist feature such that, for example, the good that the charity does *merits* some comparative reward. Merit is not required to generate desert, need or entitlement.

[6] One may argue that such instrumental concerns may operate on giving behavior directly and not through social norms. We discuss this possibility in the discussion section.

[7] They look at laboratory studies where the dictator had to earn their endowment or it was a windfall, and when the recipient was a student or charity. Here we focus on their unearned endowment analysis.

[8] They present a comprehensive characterization of donation motives using an experiment that varies treatments between- and within-subjects. By varying giving by others in multiple ways, they show that subject choices are not motivated by the total amount donated, as they would be in the altruistic model. They find that learning what others have done significantly influences donation behavior.

(2021) write, "We argue that our results are most consistent with a model of pure warm glow driven by a preference to comply with an uncertain social norm".

However, little prior work has provided direct empirical evidence regarding the correlation between injunctive social norms and behavior in charity donations; furthermore, when there is empirical evidence testing the role of norms, it has focused on the descriptive norm (i.e. what others have done in the same or similar situation). For example, Frey and Meier (2004) test the effect of informing students about the share of students who, in the past, contributed to two different university funds when paying their tuition bill. Controlling for past donations, they found that students gave more when they were informed that 64% rather than 46% of potential donors made large contributions in the past. There are several other studies on donation behavior showing similar effects arising from treatments that convey information about the descriptive norm (Croson et al., 2010; Kessler, 2017; Martin & Randal, 2008; Shang & Croson, 2009).[9]

Though prior research suggests that norm compliance is a motive for transfer decisions in both the individual and charity contexts, it remains unresolved whether (injunctive) norms differ, are more certain, and/or people adhere more strongly in one of the two contexts.[10] Eckel and Grossman (1996) compare transfer decisions when the object is a person or charity and their results could suggest that norms differ between student and charity recipient contexts. They use a double-blind dictator game with a between-subjects treatment design where subjects either make a transfer decision in the standard dictator game with a student recipient or a charity recipient (the American Red Cross). They find that subjects transfer 31% of their total earnings to a charity as compared to 10.6% to another student. While these results may imply that norms differ when the recipient is a charity or student, it could also be consistent with differences in norm adherence or even norm certainty. However, Krupka and Weber (2013) show that in the context of student dictator games, the desire to comply with norms is constant across several variations the dictator decision context even when injunctive norms vary considerably.[11]

Some prior work provides empirical support that norm certainty can help explain observed changes in behavior and that people appear more certain about norms for charity donations. Chang et al. (2019) find evidence that norms which people are more sure about impact behavior more strongly. They invite both Republicans and Democrats to play dictator games and then very between subject whether the dictator game is described using a "tax" or "neutral" frame. Republican subjects report

[9] See also Agerström et al. (2016), Andreoni (2006), Drouvelis and Marx (2021), Hermalin (1998), Huck et al. (2015), Huck and Rasul (2011), Karlan and List (2020), List and Lucking-Reiley (2002), Potters et al. (2007) and Vesterlund (2003). Krupka and Croson (2016) show that normative cues increase charitable contributions. However, they are not clear on whether the cue is providing information about the injunctive or descriptive norm.

[10] See for example, Kimbrough and Vostroknutov (2016). List (2007) and Bardsley (2008) discuss the sensitivity of giving to the details of the game and context.

[11] For instance, Krupka and Weber (2013) vary the location of the initial endowment and several other aspects of the dictator game and show that norms vary but not the weights placed on compliance with the norms.

norms with greater variance (relative to Democrats) in the neutral frame, which they interpret as greater ambiguity about the norm. They find that when they impose the tax frame, then it is the Republicans whose behavior is more affected by the frame. Echoing this finding, Dreber et al. (2013) use student and Mturk subjects to test whether merely describing dictator games as "taking games" or "giving games" affects behavior. They find that behavior is insensitive to social framing and conclude that the reason framing does not alter behavior is because there is very little norm uncertainty in the game.[12] In the context of charity recipients in dictator games, Grossman and Eckel (2015) offer a parallel interpretation to that of Dreber et al. (2013). They study a dictator game with a charity recipient in a "real donation" lab experiment and vary whether the initial endowment is initially held by the charity recipient or by the dictator. They find that contributions to charities are unaffected by the location of the initial endowment. They suggest that there is less norm ambiguity in the charity context and, as a result, the location of the initial endowment with the charity recipient or dictator does not affect behavior.

Thus, we have suggestive evidence that norms change when we vary the recipient. There is also incomplete evidence that the desire to comply with the norm is unchanged across variations in the dictator game context. And finally, we have some evidence that norm certainty explains some of the observed changes in behavior and that charity norms appear to be robust to variations in the dictator game set up. The latter suggests that charity norms may be more certain than student recipient norms. Because we have little empirical work that identifies injunctive norms for charities, and compares behavior between student and charity recipients it is difficult to know how these factors affect choice.

We address these gaps in the literature by identifying the injunctive norm for transfers to charity recipients and correlating them with changes in behavior. Our focus on injunctive norms is motivated by the surprising absence of prior work identifying injunctive norms for charitable contributions and the relative importance of the injunctive norm.[13] In what follows, we describe the experimental design.

3 Experimental design

The research reported here is part of a larger study where we recruited students from the entering undergraduate classes in 2016, 2017, 2018 and 2019 at Rice University. The aim of the larger project is to examine the evolution of economic preferences (altruism, risk aversion, time preference, competitiveness, loss aversion, ingroup favoritism, among others) across their college years. Subjects participated in

[12] They write, "There is less ambiguity about the social norm in the dictator game...and accordingly framing matters less".

[13] For example, Krupka and Weber (2013) wrote, "... As we will show, (injunctive) social norms concerning the appropriateness of behavior are sufficient for explaining a considerable amount of variation in other-regarding behavior".

numerous laboratory and online studies between matriculation and 2021. Here, we report on the experimental design of a subset of tasks that subjects completed.

The survey wave used in this paper was collected in June and July of 2021. This survey was composed of fifteen modules and had a total of 710 participants. For the present analysis, we use modules 8, 10 and 12.[14] In each module, subjects played a dictator game, guessed what others did in the game and played a coordination game designed to elicit norms for the dictator game they just played.

The order of dictator game modules was partially randomized. For reference, Fig. 1 contains a diagram of the survey flow. For approximately half of the survey participants, module 8 came first, and was later followed by modules 10 and 12. For all other participants, module 10 appeared first, and then followed by modules 8 and 12. Screenshots for the modules are in Online Appendix G.

The dictator game in each module 8, 10, and 12 asked subjects to allocate an endowment of $20 between themselves and a recipient who varied depending on the module. Subjects could choose between eleven possible allocations in the dictator games. The allocations were in $2 intervals (i.e. keep $0 and send $20, keep $2 and send $18 and so on, see Online Appendix G Q11.2 and Q12.7). In module 8, the recipient was an anonymous survey participant who was a freshman at Rice University. In module 10, the recipient was an anonymous survey participant who was from the same entering class as the responding subject.

In module 12, the recipient was a charity focused on COVID-19 relief in India. For the charity, we provided the following information:

> In this task, you have the opportunity to donate for COVID-19 relief work in India. Since March 2021, India has reached devastating levels of infections and mortality, and is currently the epicenter of the pandemic. Over 300,000 new infections have been reported everyday between April 21, 2021 and May 15, 2021. The last 100,000 deaths in India from COVID-19 took place over less than a month's time. India's healthcare system has been under deep stress with many patients having to wait for several hours for treatment. Additionally, the huge surge in the number of patients has led to deep shortage of life-saving oxygen, intensive care units and critical medical supplies such as surgical masks, face shields and PPE. Lack of oxygen and in-time medical care has resulted in hundreds of COVID deaths from respiratory failure. During the second wave, India recorded the highest single day death-toll from COVID-19 in any country since the beginning of the pandemic. To support the ongoing relief work, you have the opportunity to donate to UNICEF USA that is working with its partners to provide oxygen concentrators, diagnostic test kits, PPE and other supplies necessary to treat and contain the spread of infection. More information is available in the following link that you can copy and paste into your browser: https://www.unicefusa.org/stories/india-faces-brutal-covid-19-crisis-unicef-there-help/38520

[14] The other modules included questions regarding COVID-19 precautionary behavior, elicited measures of standard economic preferences, and beliefs.

We chose to include a descriptive vignette for the charitable organization for multiple reasons. First, the experiment is designed to emulate a common presentation format for experiments using student recipients and charity recipients (Fehérová et al., 2022; Jackson, 2022; Sanchez, 2022; Shang & Croson, 2009). Second, per a typical charity solicitation, it includes information regarding the charity's mission and objectives (text was taken directly from the charitable organization's website). A final motivation was that student subjects may know and be familiar with a typical student recipient but might require additional descriptive information about the particular UNICEF USA initiative.

After the dictator allocation decision, subjects were asked to guess the average dictator allocation for that module. This question was incentivized such that if the subject's guess was within $2 of the average allocation decision they received a $2.50 bonus.

The last task of each module had subjects play a variant of the Krupka and Weber (2013) social norm elicitation game. Subjects were asked to consider the dictator game they had just played and to assess how another participant who resided in the same residential college as them, would rate the appropriateness of different allocations.[15] Specifically, they evaluated the social appropriateness of eleven possible allocation decisions on a four point scale, ranging form 'very socially inappropriate' to 'very socially appropriate' (see Online Appendix G for screenshots—particularly Q11.4). For each allocation-evaluation, the subject's response was compared to other study participants in the same residential college. If this task was selected for payment and the subject selected the same social appropriateness rating as the modal response provided by other participants in the same residential college, then the subject earned a bonus of $3.[16] The defining feature of this payment scheme is that subjects are incentivised to coordinate on appropriateness ratings; Krupka and Weber (2013) show that norms create focal points which subjects use in their attempt to coordinate.[17] Thus we interpret responses to capture the social norms for the particular situation.

After all respondents completed the study, a module was randomly selected for each respondent for payment. In the dictator game (modules 8, 10 or 12) respondents knew they could be randomly assigned to either the role of dictator or recipient. Dictators were paid their earnings and recipients, after being randomly assigned to a dictator, were paid what the dictator transferred. In addition, if the module was chosen respondents were paid an additional $2.50 if they guessed within $2 of the average allocation made by dictators. Finally, norms ratings for three different actions (i.e. three different dictator transfers) were randomly selected for payment. The respondent was paid $3 if their rating matched the modal rating. For instances in which a dictator was chosen and the recipient was an incoming freshman at Rice,

[15] Students are randomly assigned to one of twelve residential colleges on Rice campus, and they remain with the college—where they sleep, eat, and study—for their entire time at Rice.

[16] The payment process is described in more detail in the next paragraph.

[17] The Krupka and Weber (2013) mechanism has been used extensively to elicit social norms and has been shown to be robust to competing alternative focal points (Fallucchi & Nosenzo, 2022).

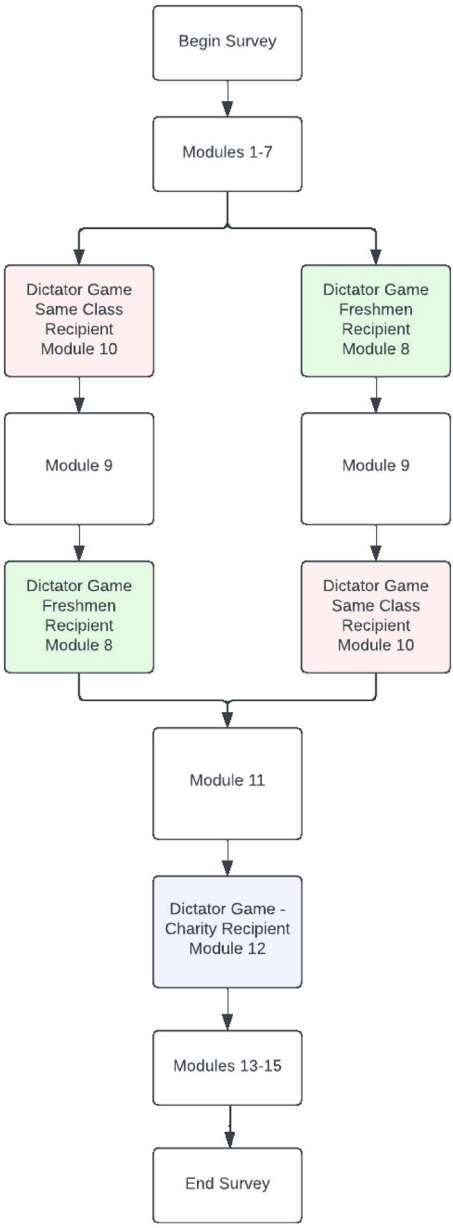

Fig. 1 Depiction the flow of the experimental survey. Subjects were randomized into either the left- or right-hand side. Modules 1–7, 9, 11, and 13–15 are not used in the presented study

then a random freshmen was chosen, an envelope was prepared with the cash, and a brief explanation of the task was provided. Likewise, if the charity module was selected, dictators received their allocation and the charity was sent whatever allocation remained. Respondents were emailed information about their earnings, noting which module was selected for payment, and then paid on-line in the manner they preferred (an example of the email is included in Online Appendix G—payment

options included Venmo, PayPal, and Amazon). Respondents were part of a long-term panel and had previously been paid in a similar manner. Therefore we believe that the payoffs in the experiment were credible for the respondents.

A majority of participants graduated in 2020 (314 participants), with the rest expecting to graduate in 2021 (176 participants), 2022 (123 participants), 2023 (95 participants), 2024 (1 participant) and 2025 (1 participant). Excluding outliers, subjects spent an average of 44.6 min on the survey and earned an average of $21.96.[18] All experimental instructions, as well as experiment procedures and average payments waves are detailed in Online Appendix G.

4 Results

We first compare the dictator transfers among all three treatments. We then present the distribution of norm ratings and test whether norms differ across the treatments. Next, we combine the data on dictator transfers with norms data to understand the relationship between the two. Finally, we analyze the transfer choices of our subjects in a utility framework that brings together concern for one's own monetary payoff and one's social norm compliance.

Result 1: Dictators transfer a significantly larger amount to the recipient in the charity treatment.

Figure 2 presents the distribution of dictator transfers under our three treatments. In treatment 1, when the recipient is a charity, more subjects choose allocations that are favorable to the charity than when the recipient is a student. Over a quarter, or 26.06% (N = 185) of our subjects transfer all their endowment ($20) in the charity treatment. By comparison, only 6.1% (N = 43) and 2.4% (N = 17) of subjects sent $20 to recipients in the Freshmen treatment and same-class treatment, respectively.

Table 1 presents the summary statistics of dictator transfers for all three treatments. The average transfer in the charity treatment is $11.60 or 58% of the initial endowment of $20. The freshmen and same-class average transfer was much lower, $6.54 and $6.44 or 32.7% and 32.2% of the initial endowment, respectively. Using paired t-tests, we find that transfers to the charity were significantly larger than transfers to freshmen and same-class subjects ($t = 12.18$, $p = 0.00$; $t = 13.37$, $p = 0.00$). This finding is also supported by the results of a Wilcoxon signed-rank test, which places no parametric assumptions on the distribution of transfers ($Z = 11.02$, $p = 0.00$; $Z = 11.84$, $p = 0.00$). When we compare freshmen directed transfers to same-class directed transfers, we find no statistical differences between

[18] When we include outliers the average payment remains the same ($21.96); however a small proportion of subjects left the survey and completed it on another day. Including these observations increases the average completion time to 30.28 h.

transfers, using both paired t-tests ($t = 0.63$, $p = 0.53$) and a Wilcoxon signed-rank test ($Z = -0.17$, $p = 0.87$).[19]

To analyze the social appropriateness ratings, we follow Krupka and Weber (2013) by converting the appropriateness rating for each allocation choice into a numerical score. A rating of "Very Socially Appropriate" by a subject is assigned a value of 1, "Somewhat Socially Appropriate" a value of 1/3, "Somewhat Socially Inappropriate" a value of − 1/3 and "Very Socially Inappropriate" is assigned a value of − 1. Recall that these ratings are elicited using an incentivized coordination game. Subjects could earn additional money if their response matched with the modal rating of that allocation choice.

Result 2: The injunctive norm of giving in the charity treatment is significantly higher than in the student treatment.

Table 2 presents the mean appropriateness rating for each allocation for the freshmen and charity treatments. This table contains the means, standard deviations, and the distributions of appropriateness ratings for each allocation choice. The left side of the table contains results from the freshmen treatment and the right side contains results from the charity treatment. In the charity treatment, the mean appropriateness rating monotonically increases in the amount transferred to the charity and peaks at the allocation choice that yields nothing for the dictator and everything for the charity ($0, $20). In the freshmen treatment, however, the appropriateness rating increases up until the even-split allocation ($10, $10), after which it falls. The most appropriate action in this treatment is to equally divide the endowment, whereas in the charity treatment the most appropriate action is to transfer all of the endowment to the recipient.

The second to the last column in Table 2 reports t-statistics for paired t-tests between the norm ratings for each transfer amount across treatments. All allocation appropriateness ratings (except for the ($18, $2), ($16, $4), and ($8, $12) allocations) are significantly different between treatments. For transfer amounts below $10, the difference is positive, with higher levels for the freshmen than the charity treatment. For transfer amounts above $10, the difference is negative. In the last column of the table, we present results from a Wilcoxon rank-sum test comparing the distribution of ratings for each possible transfer amount between the two treatments. The results of these two tests put into perspective the pronounced difference in the social norms of giving to a charity versus a freshmen student. Almost all choices that yield lower payoff for the recipient are considered inappropriate in the charity treatment. In the treatment where the recipient is a freshman student, only the choices that really disadvantage the recipient (give $0 and give $2) are considered inappropriate.

In a similar manner, we report the mean, standard deviations, and distribution of appropriateness ratings for the same class treatment and compare these results

[19] See Online Appendix B for an analysis of social appropriateness ratings at the residential college level.

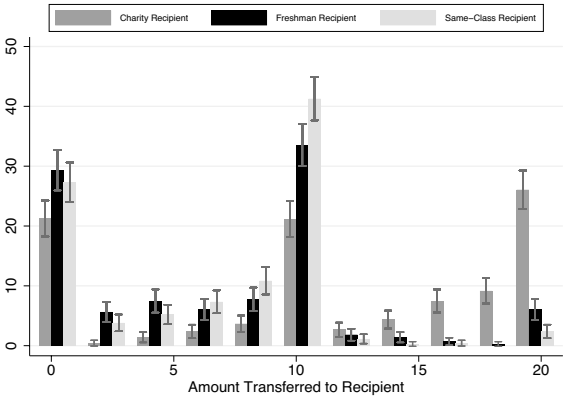

Fig. 2 Percentage of dictator allocation decisions for charity recipient, freshmen recipient and same-class recipient dictator game treatments. Error bars indicate 95% confidence intervals

Table 1 Summary statistics of amount by treatment

	Mean	SD	Min	Max	N
Charity dictator sent ($)	11.60	7.45	0	20	710
Freshmen dictator sent ($)	6.54	5.63	0	20	710
Same-class dictator sent ($)	6.44	4.82	0	20	710

Summary statistics of amount transferred by the dictator by treatment

to the charity treatment in Table 3.[20] The left side of the table reports results from the same class treatment while results of the charity treatment are on the right side. Again, we find that appropriateness ratings peak at the ($10, $10) allocation decision for the student recipient treatment. Appropriateness ratings between the same class treatment and the charity treatment are not different from each other for allocation choices of ($18, $2), ($16, $4), and ($8, $12). Note that the highest mean rating for the charity treatment is 0.84 and corresponds to donating the full endowment to the charity, while in the student treatment the highest mean rating is 0.83, and corresponds to the 50/50 split, indicating that nearly all subjects agree on the appropriateness of these actions.

In the charity treatment, subjects are better able to match their responses to that of the group for the appropriateness ratings of other actions. In most cases, more than 50% of all participants generally agree on the appropriateness rating of different alternative choices. On the contrary, in the same-class treatment, there is less agreement about the appropriateness of allocation choices that strongly favor the recipient. Evidence of lower agreement can be seen by the increased standard

[20] Table D6 in Online Appendix D reports the means, standard deviations, and distributions of appropriateness ratings between the freshmen and same class recipient treatments. Table D6 also reports the t-statistic and Wilcoxon rank-sum test statistic for each allocation amount between the treatments. The results of the Wilcoxon Sign Rank test shows that for almost all allocation choices, the social appropriateness ratings between freshman and same-class recipients are not statistically different from one another.

Table 2 Summary statistics of norm ratings when the recipient is a freshmen student or a charity

Freshmen student recipient (N = 710)								Charity recipient (N = 710)								T statistic	Wilcoxon sign rank test
Potential allocations	Action	Mean	SD	VI (%)	SI (%)	SA (%)	VA (%)	Potential allocations	Action	Mean	SD	VI (%)	SI (%)	SA (%)	VA (%)		
20.0	Give $0	− 0.48	0.60	48	32	14	6	20.0	Give $0	− 0.55	0.58	54	31	9	6	3.410***	3.915***
18.2	Give $2	− 0.37	0.59	35	41	18	6	18.2	Give $2	− 0.37	0.61	38	37	18	7	0.199	0.794
16.4	Give $4	− 0.17	0.52	15	51	27	6	16.4	Give $4	− 0.18	0.60	23	40	28	9	0.438	0.280
14.6	Give $6	0.05	0.48	6	39	47	8	14.6	Give $6	0.00	0.58	13	36	39	12	1.931*	2.252**
12.8	Give $8	0.33	0.44	2	15	63	19	12.8	Give $8	0.18	0.55	7	25	48	18	6.438***	6.809***
10.10	Give $10	0.78	0.36	1	2	24	72	10.10	Give $10	0.53	0.48	2	9	44	45	13.079***	12.352***
8.12	Give $12	0.59	0.47	2	6	43	49	8.12	Give $12	0.59	0.46	2	6	43	49	− 0.268	− 0.305
6.14	Give $14	0.48	0.54	3	14	40	42	6.14	Give $14	0.67	0.44	2	4	36	58	− 8.567***	− 8.341***
4.16	Give $16	0.40	0.61	6	21	31	42	4.16	Give $16	0.74	0.44	2	4	24	70	− 14.113***	− 13.038***
2.18	Give $18	0.33	0.73	15	16	25	45	2.18	Give $18	0.77	0.43	2	3	21	74	− 16.325***	− 14.520***
0.20	Give $20	0.35	0.69	11	18	27	43	0.20	Give $20	0.84	0.39	2	3	14	82	− 19.769***	− 17.021***

We report the mean and standard deviation for each possible allocation by dictator treatment. We also report the percentage of responses of each social appropriateness rating in terms of percentage. The second to last column report the t-statistic of a paired t-test between the charity and student treatment and the last column reports the combined signed-rank sum from a Wilcoxon signed-rank test

Percentages may not sum to 100 due to rounding

VI very inappropriate, *SI* somewhat inappropriate, *SA* somewhat appropriate and *VA* very appropriate

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 3 Summary statistics of norm ratings when the recipient is a same-class student or a charity

Same-class student recipient (N = 710)								Charity recipient (N = 710)								T statistic	Wilcoxon sign rank test
Potential allocations	Action	Mean	SD	VI (%)	SI (%)	SA (%)	VA (%)	Potential allocations	Action	Mean	SD	VI (%)	SI (%)	SA (%)	VA (%)		
20.0	Give $0	− 0.50	0.60	50	30	13	6	20.0	Give $0	− 0.55	0.58	54	31	9	6	2.482**	3.026***
18.2	Give $2	− 0.38	0.58	36	42	17	6	18.2	Give $2	− 0.37	0.61	38	37	18	7	− 0.241	0.228
16.4	Give $4	− 0.18	0.54	17	50	26	7	16.4	Give $4	− 0.18	0.61	23	40	28	9	− 0.039	− 0.151
14.6	Give $6	0.05	0.47	4	41	46	8	14.6	Give $6	0.00	0.58	13	36	39	12	2.211**	1.839*
12.8	Give $8	0.33	0.41	2	14	66	18	12.8	Give $8	0.19	0.55	7	25	48	19	6.837***	6.830***
10.10	Give $10	0.83	0.32	0	1	22	77	10.10	Give $10	0.54	0.49	2	9	44	45	15.906***	14.120***
8.12	Give $12	0.59	0.45	2	6	44	48	8.12	Give $12	0.59	0.46	2	6	43	49	0.000	0.051
6.14	Give $14	0.46	0.54	3	16	40	41	6.14	Give $14	0.67	0.44	2	4	36	58	− 9.342***	− 9.111***
4.16	Give $16	0.37	0.62	7	21	32	40	4.16	Give $16	0.74	0.44	2	4	24	70	− 14.914***	− 13.588***
2.18	Give $18	0.26	0.76	19	15	25	41	2.18	Give $18	0.77	0.43	2	3	21	74	− 17.703***	− 15.488***
0.20	Give $20	0.30	0.70	13	20	27	40	0.20	Give $20	0.84	0.39	2	3	14	82	− 20.866***	− 17.531***

We report the mean and standard deviation for each possible allocation by dictator treatment. We also report the percentage of responses of each social appropriateness rating in terms of percentage. The second to last column report the t-statistic of a paired t-test between the charity and student treatment and the last column reports the combined signed-rank sum from a Wilcoxon signed-rank test

Percentages may not sum to 100 due to rounding

VI very inappropriate, *SI* somewhat inappropriate, *SA* somewhat appropriate, *VA* very appropriate

*$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

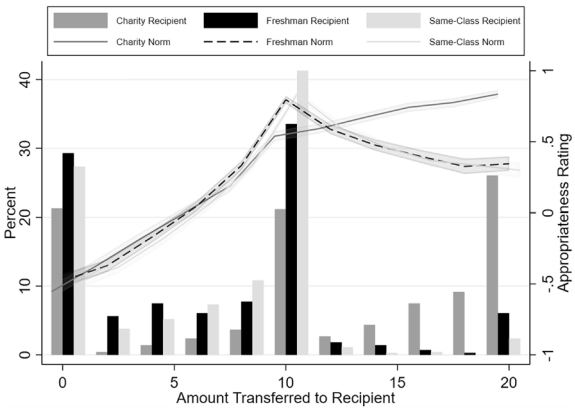

Fig. 3 Bars denote the frequency of dictator allocation decisions in terms of percentage for all three treatments (left y-axis). Lines denote the average social appropriateness rating associated with each allocation decision (right y-axis). Shaded areas denote 95% confidence intervals

deviations of social appropriateness ratings for transfers above ($10, $10) in the same class and freshmen treatment in comparison to the charity treatment. This indicates that our subjects agree more strongly about what the norm is when the recipient is a charity rather than another student.

Result 3: Dictator transfers closely follow the social norms of giving.

In Fig. 3, we overlay the mean and the 95% confidence interval of the appropriateness rating of each allocation choice over the distribution of dictator transfers in all three treatments. As seen in the figure, the modal behavior in each of the two treatments coincides with the peak of mean appropriateness rating. In the charity treatment, the mean appropriateness rating hits an apex at ($0, $20). This allocation choice yields $0 for the dictator and $20 for the recipient: 26.06% of all subjects in the charity treatment gave away all their endowment ($20) and almost half of all subjects (49.72%) transfer more than half of their endowment to the charity. In the freshmen and same class treatment, the mean rating peaks at ($10, $10) which yields $10 each for the dictator and the recipient: 34% and 41% of subjects chose ($10, $10) as their preferred allocation in the freshmen and same class treatment, respectively.

Note that in all treatments, subjects maximize their monetary payoff by keeping the $20 for themselves. However, this payoff-maximizing allocation ($20, $0) is considered more inappropriate when the recipient is a charity than when it is another student, either a freshman or someone in the same entering class. In parallel fashion, we see behavior in the treatments that mirrors the differences in appropriateness rating. More subjects chose the ($20, $0) allocation in the freshmen and same-class treatment than in the charity treatment. Around 29% and 27% of subjects choose to share nothing in the freshmen and same class treatment respectively, in comparison to 21% of subjects in the charity treatment. Note that this happens even though there is only a small difference in the appropriateness rating of the allocation choice ($20, $0) between the treatments: − 0.55 in the charity treatment versus − 0.48 in the freshmen treatment and − 0.50 in the same class treatment.

The equal split allocation of ($10, $10) also has notably different norm ratings between treatments (0.53 in charity versus 0.78 in freshmen and 0.83 in same-class). In the freshman and same-class treatment, this allocation choice is rated as the most appropriate action by the dictator. This implies that any rightward deviation from this action (transferring to the recipient more than $10) yields the dictator both lower payoff and lower norm ratings. As such, we see that only 10.82% and 4.22% of participants in the freshmen and same-class treatment allocated more than $10 to the recipient.

We analyze participants' allocation choices in a social norm utility framework in order to explore the explanatory power of self-interest and norm-compliance motives in determining giving behavior and the differential value of norm compliance between recipient types. We present the utility framework suggested by Krupka and Weber (2013). Let $A = a_1, ..., a_k$ be the set of actions available to a decision maker. Assuming that individuals care both about the monetary payoff from choosing an action a_k and the extent of its social appropriateness, the utility function can be written as:

$$U(a_k) = V(\pi(a_k)) + \gamma N(a_k) \tag{1}$$

In Eq. (1), $V(\pi(a_k))$ is the value a decision maker attaches to monetary returns from taking action a_k and $N(a_k)$ is the social norm rating associated with action a_k. The decision maker's utility is increasing in both monetary payoff and the norm rating of an action. The parameter $\gamma \geq 0$ captures the extent to which a decision maker cares about aligning his/her actions with the social norm. An individual with higher γ will derive more utility from selecting an action collectively perceived as more appropriate than someone with a lower γ. On the contrary, for individuals who do not care about aligning their action with social norms, γ takes a value of 0.

To empirically estimate Eq. 1, we impose a linearity restriction on $V(\pi(a_k))$. As a result, Eq. 1 can be estimated as:

$$U(a_k) = \beta\pi(a_k) + \gamma N(a_k) \qquad \text{(2: Krupka–Weber Model)}$$

We use a conditional logit regression (McFadden, 1973) to identify the parameters—β and γ—in Eq. (2). Note that in a conditional logistic regression where the dependent variable is the selected action, variation stems from the characteristics of the possible actions.[21] In our experiment, these characteristics are the payoffs and norms. However, when we change recipient type, we hold the monetary payoff constant and the only source of variation is variation in norms.

In order to run our regressions, we expand the data set to include all possible choices for the charity treatment and the dictator treatment that the subject first completed, which is either the freshmen treatment or same class treatment. This restriction simplifies the conditional logit analysis and is justified because the freshman

[21] Conditional logistic models are similar to multinomial logistic models. However, conditional logistic models emphasize the characteristics of the alternatives, while multinomial logistic models depend on the characteristics of the individual making the choice. See Hoffman and Duncan (1988) for a comparison of these models.

and same class norms are extremely similar.[22] We transform the data into a panel where we observe a subject's decision under the charity treatment and the other student treatment (2 games of 11 possible choices leads to 22 rows per subject and $710 \times 22 = 15{,}620$ rows in total). We define a binary dependent variable (choice) which takes a value of 1 for the allocation selected and 0 for all other allocations (which were not selected). The social norm for subject i is computed as the average social appropriateness ratings for action a_k by all the participants excluding the social appropriateness rating provided by subject i. This method of computing the "leave one out" mean allows us to exclude the possibility that subjects behave or report social appropriateness ratings in a manner to be self-consistent.

Table 4 reports the results of four different specifications. In all cases, the explanatory variables are the monetary payoff from the chosen action and the elicited appropriateness rating associated with each allocation choice. In the first column we predict choice only as a function of the payoff characteristic of the action, column 2 predicts choice as a function of both payoff and norm characteristics. Although norms may have a different profiles across the recipient-types, it may also be the case that the desire to comply with the norm is stronger (i.e. a person places more weight on norm compliance) depending on the recipient-type (Kimbrough & Vostroknutov, 2016; Krupka & Weber, 2013). We investigate this possibility by including an interaction term of the empirically estimated norm with a charity-treatment indicator variable and report the regression in column 3. In column 4, we run a regression specification where a subject is only observed once and the data is no longer a panel consisting of two treatments. This specification allows for clustering (and matched groups) to be at the subject-level. To do this, we divide the subject pool in half and use the first half's charity-recipient observations and the second half's student recipient observations. This regression gives us a way to examine whether violations of "the identically and independently distributed" assumption necessary for the conditional logit model are generating our results of interest.

Columns 1 through 4 in Table 4 present parameter estimates for the Krupka and Weber (2013) utility framework given by equation 2: Krupka–Weber Model. In column 2 of Table 4, the estimated coefficients for both monetary payoff and appropriateness rating are positive and statistically significant. This implies that subjects place a positive weight on both attributes—they care about their own earnings, as well as conforming to the social norm. Moreover, the influence of social appropriateness on behavior is not just statistically significant, but also large in magnitude. In column 3, the interaction term is positive and significant and could indicate that decision makers gain more utility from taking an action that is seen as more socially appropriate when the recipient is a charity. This is an observation we unpack below. Finally, in column 4, we observe results which are quantitatively similar to those presented in column 3.[23]

[22] We show in the Online Appendix that our selection of comparing the results to the freshmen or same class treatment are similar. See Table A2 and Table A1 in Online Appendix A.

[23] Because we did not elicit beliefs about the distribution of behavior (i.e. the descriptive norm), we cannot include this in our Clogit regressions. However, we find that guesses positively correlated with transfer amounts for freshmen-recipient and same-class-recipients (rho = 0.57, p value = 0.000; rho = 0.54,

Table 4 Conditional logit regression on Krupka–Weber model

	(1)	(2)	(3)	(4)
Monetary payoff (β)	0.0247***	0.168***	0.199***	0.0920***
	(5.32)	(14.16)	(14.42)	(4.96)
Avg. appropriateness rating (γ)		2.130***	1.968***	0.677*
		(15.90)	(12.46)	(2.14)
Avg. appropriateness rating (γ) × charity game			1.015***	0.989***
			(6.56)	(4.04)
Subject-level clustering	No	No	No	Yes
Observations	15,620	15,620	15,620	7810
Pseudo R^2	0.00504	0.0542	0.0634	0.0186
Clusters	1420	1420	1420	710
AIC	6777.7	6444.6	6384.5	3347.7
BIC	6785.4	6459.9	6407.4	3368.6

Results of a conditional logit model using responses from the first student recipient dictator game subjects played and the charity-recipient dictator game. The equation estimated is Equation 2: Krupka–Weber Model. t statistics in parentheses

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Column 4 re-specifies the data such that it is no longer a panel and there is only one dictator game observation per subject

Result 4: Subjects place a positive weight on both payoff and norm characteristics of their choice and the influence of the social norm is large in magnitude.

To compare the likelihood that each model fits the observed data, we use the Akaike and Bayesian information criteria (AIC and BIC). Specifically, smaller AIC and BIC values indicate a better fit of the model to the data.[24] In all specifications that include norms (columns 2–4), the AIC and BIC scores are lower indicating better model fit.[25]

Returning to the significant interaction term in column 3 of Table 4, Krupka and Weber (2013) interpret significance as evidence that subjects' concern for norm compliance is increased in the treatment. An alternative explanation is that norms are less certain in the student recipient case. When social norms are uncertain (such as we see in Tables 2 and 3), then utility from norm compliance may be diminished.

Footnote 23 (continued)
p value = 0.000) but negatively correlated with transfer amounts for charity-recipients (rho = − 0.55, p value = 0.000). Though this is not strictly speaking the descriptive norm, it is suggestive that in the case of transfers to the student, subjects could be following the descriptive norm or the injunctive norm and also believe that others are following these norms. However, when it comes to the charity, their behavior is better predicted by the injunctive norm.

[24] AIC and BIC penalize models for the number of parameters so if norms have no influence on behavior, we expect the model that includes norm characteristics to have larger AIC and BIC values than the model used in column 1.

[25] As an additional analysis, we plot the observed frequencies of allocation choices against the models prediction. This analysis is provided in Online Appendix C.

Similarly, if people are choosing optimally given their own beliefs, and their beliefs are more heterogeneous with student recipients, then this could explain the weaker effect of the mean belief on choice in the student recipient treatments.[26]

Because our elicitation technique provides incentives to accurately report one's belief about the most common response, subjects' own beliefs about the norm may be more likely to influence their choices than the mean belief.[27] Furthermore, subjects would have no uncertainty regarding their own beliefs about the appropriateness and, as such, the utility gained from norm compliance should be equal across treatments if we use their own beliefs about the norms to predict choice.[28]

Result 5: When there is no uncertainty surrounding what is and is not appropriate, compliance with the social norm is not influenced by recipient-type.

The results of this regression specification are provided in Table 5. We find that the interaction between appropriateness rating and the charity-treatment is no longer significant. One way to interpret this is that when there is no uncertainty surrounding what is and is not appropriate, compliance with the social norm is not influenced by recipient-type.[29]

We demonstrate that regardless of the student recipient matching game used in the conditional logit regression analysis (either the freshmen recipient or same class recipient), we arrive at similar results. That is, social norms positively correlate with dictator transfer decisions and that this correlation is not stronger once we control for norm uncertainty.

5 Discussion

Transfers in dictator games are sensitive to demand effects (Levitt & List, 2007; Bardsley, 2008). We adopt the definition in Zizzo (2010) and define an experimenter demand effect (EDE) as changes in behavior by experimental subjects due to cues about what constitutes appropriate behavior. Zizzo (2010) goes on to say that, "... sometimes researchers may need to consciously accept a trade-off between different experimental objectives and constraints, and it may be optimal for them to accept some risk of an EDE as a result, rather than going for a corner solution where such risk is brought to zero". The language we used to describe the charity is more extensive and full of urgency relative to that which we used in the student transfer decisions. As such, the language could provide cues about what is appropriate behavior and could constitute a social source of experimenter demand. However, just as De

[26] See also previous research which discusses the role of norm ambiguity in affecting behavior (Chang et al., 2019; Dreber et al., 2013).

[27] We thank an anonymous referee for directing us to this line of inquiry.

[28] One may be concerned that utilizing the subject's own appropriateness rating as the empirical proxy of the norm would lead to bias in dictator actions. However evidence suggests that elicited norms are stable across different types of actors, notably dictators and third-party spectators (Erkut et al., 2015).

[29] A future avenue of research could be to explore the extent to which subjects may utilize the opportunity of uncertainty to create moral wiggle room and thus decrease transfer amounts.

Quidt et al. (2019) note that, "...while there is evidence of response to framing, this need not be evidence of experimenter demand", it is unclear whether we have differential experimenter demand effects such that any demand effects operated more strongly in the charity case.

For both the charity and student recipient case, we sought to enhance validity by using the common framing adopted in our settings. "An EDE that parallels or helps reproduce an important feature of the real world setup the experiment is trying to model is an EDE that may strengthen the experiment by enhancing its external validity" (Zizzo, 2010, p. 92). In our case, describing the charity as one would encounter it invivo, is an aspect of the parallelism to typical charity solicitations which we sought to establish (Friedman et al., 1994). Furthermore, using the typical language to describe the student dictator games was important to creating parallelism to our bench marking case—the typical laboratory setting where anonymous dictator-recipient pairs are students from the same university. We follow in the footsteps of several other experimental papers who adopted similar approaches (Fehérová et al., 2022; Jackson, 2022; Sanchez, 2022; Shang & Croson, 2009).[30] As Levitt and List (2007) note, scrutiny can exaggerate the importance of prosocial behaviors relative to environments without such scrutiny. Therefore, we focus on effects of treatment conditions relative to control conditions and infer little from the absolute magnitude of the amount given.

The within-subject design also provides us with opportunities to test for a possible marker of experimenter demand. Repeated exposure to the dictator games and norm elicitation could lead to a kind of demand effect whereby subjects are more sure about what the experimenter is after in the last game or the last norm elicitation than they were in the first game or norm elicitation. This type of demand effect would give rise to an order effect whereby we would observe a decrease in the variance of actions in later stages of the experiment. It would also predict that the standard error of the norms elicited in later modules would be smaller than those elicited in earlier modules.

We don't find much evidence of these patterns.[31] For example, we test whether the elicited appropriateness ratings for the charity recipient situation (elicited last) have a lower standard deviation than the appropriateness ratings for the student

[30] For example, in the laboratory experiment reported in Shang and Croson (2009), they use very detailed descriptions to set the stage for their subjects. Here we quote only a bit of it and refer the reader to the Online Appendix. "..imagine that you have been listening to a public radio station every day for the past four years. Another listener, Mary (Tom), listens to the same radio station. This is the only radio station that Mary (Tom) listens to. Every morning Mary (Tom) wakes up, and she (he) turns on the radio. She (He) has her (his) breakfast and prepares for the day with the radio in the background. Every evening, as soon as she (he) gets off work, she (he) turns on the same station and listens to it on her (his) way home. She (He) has been to several station-sponsored events and concerts. This radio station is very important in her (his) life, and if it were to go away, she (he) would miss it".

[31] See Online Appendix E for a through analysis on potential ordering effects. Further, we note that in a d'Adda et al. (2016) they use a within-subject design whereby the norms are elicited from the same subjects who are observed making choices in a bribery game. They test whether the order in which the norm-elicitation task and the bribery game are conducted effects elicited norms and behavior. They find little evidence of order effects in our experiment.

Table 5 Conditional logit regression on Krupka–Weber model—self-reported appropriateness ratings

	(1)	(2)	(3)	(4)
Monetary payoff (β)	0.0247***	0.0785***	0.0765***	0.0113
	(5.32)	(10.16)	(9.88)	(1.13)
Reported appropriateness rating		0.844***	0.933***	0.504**
		(11.57)	(10.16)	(2.87)
Reported appropriateness rating (γ) × charity game			− 0.195	− 0.252
			(− 1.85)	(− 1.30)
Subject-level clustering	No	No	No	Yes
Observations	15,620	15,620	15,620	7810
Pseudo R^2	0.00504	0.0343	0.0349	0.00538
Clusters	1420	1420	1420	710
AIC	6777.7	6580.5	6578.2	3392.7
BIC	6785.4	6595.8	6601.1	3413.6

Results of a conditional logit model using responses from the first student recipient dictator game subjects played and the charity-recipient dictator game. The equation estimated is Eq. 2: Krupka–Weber Model. t statistics in parentheses

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. Column 4 re-specifies the data such that it is no longer a panel and there is only one dictator game observation per subject

recipient situation (elicited prior). We compare the standard deviation of appropriateness ratings between the student recipient and the charity recipient situation (student recipient appropriateness rating std. error is 0.0076, the charity recipient appropriateness rating std. error is 0.0079; f-stat. = 0.9215, df = 7809 p stat. = 0.0003). That is, our result is that the standard deviation of responses did not decrease (the EDE prediction), but rather increased in the norm elicitation for the charity. We can also look at behavior and test whether variance in the latter games is smaller (e.g. because subjects are more sure about what is "demanded of them") than in the first games they played. We don't find differences in the variances when comparing the order in which subjects played the student games, though we do find that variance for the charity dictator games is lower.[32,33]

Ensuring that our results stem from differences in norms rather than differences in salience is also important. We understand salience to be how prominent or

[32] We conduct a two-sample variance test using the transfer amounts of subjects who played Module 10 first and subjects who played Module 10 after playing Module 8 and find no statistical difference (std. error = 0.180, f-stat. = 1.1198, df = 352, 356, p value = 0.287). Similarly, subjects who played Module 8 first and subjects who played Module 8 after completing Module 10 show no statistical difference in standard deviations (std. Error = 0.211, f-stat. = 0.9186, df = 352, 356, p value = 0.4248). When we compare the class recipient or freshman recipient transfers to the charity recipient transfers, we find that class recipient transfers have a significantly larger standard error (f-stat = 0.4186, df = 709, p value = 0.000) as do freshmen-recipient transfers (0.279, f-stat. = 0.571, df = 709, p value = 0.000). Because the charity is last, after both student games, this could be consistent with EDE.

[33] There is also the possibility that the elicited norm itself (i.e. levels) may capture experimenter demand. We acknowledge that this is a possibility.

emotionally striking something is. We offer two supporting types of evidence that speak against salience as an alternative explanation for our findings. Recall, that in our sample subjects transfer 58% of their total earnings to the charity, while only transferring 32.19% and 32.69% to class and freshman subjects respectively, a difference of about 26 percentage points. Eckel and Grossman (1996) test for differences in dictator transfers to a student or a charity (in their case the charity is the Red Cross), they use a between-subject design and they do not describe the charity or its activities. They find that subjects transfer 31% of their total earnings to a charity and transfer between 9.2 and 15% to another student, a difference of 16–21.8 percentage points. We also compare to a meta-analysis (Umer et al., 2022) where salience would vary across the different dictator studies considered. Umer et al. (2022) focus on laboratory experiments that employed either student or charity recipients. They find that when the recipient was a student, subjects transfer about 21% of the endowment but when the recipient was a charity, they transferred about 45% of the endowment, a difference of about 24 percentage points. Both the between-subject design result and the meta-analysis result would minimize the impact of salience but obtain similar results to ours; this suggests that the relative salience of the charity recipient treatment is likely not an alternative explanation to our findings.

Finally, a possible criticism of our experimental design is that by eliciting social norms after subjects choose an allocation amount, subjects may intentionally misreport appropriateness ratings in a way that makes their allocation decision appear more socially appropriate than it may actually be. Erkut et al. (2015) tested whether there are differences between asking subjects their beliefs about the norm (using the Krupka and Weber 2013 method) who either read about the game or just played the dictator game. They found that norms of dictator game giving elicited from "stakeholder" dictators and recipients (i.e. subjects who play the game and do the norm elicitation task) and "spectators" (disinterested third parties who do not play the game and only do the norm elicitation task), are similar to those elicited from disinterested third parties. Their results suggest that norms elicited using the Krupka and Weber (2013) procedure are not malleable to judgment biases associated with the role of the respondents, nor are they affected by respondents' previous experience with the decision setting they are asked to evaluate. Although Erkut et al. (2015) provide suggestive evidence that social desirability bias is likely not present, our regressions calculate the mean norm rating provided by other subjects such that by construction, an empirically estimated norm would have no individual-level "bias" in responses. We also report in the Online Appendix F on a second study where the motivation for self-consistency is muted and we find similar results (Sinha et al., 2021).

Although our experiment closely mimics real-world donation solicitations to charities, the present design is unable to identify the causal chain between recipient-level differences, social norms, and giving behavior. Differences in the recipient may alter dictator giving behavior directly, rather than influencing the social norm which in turn alters behavior. For example, an anonymous student recipient differs from a charity recipient in their inherent deservingness or need. The design we use here does not allow us to distinguish norm abiding behavior from other instrumental motivations and, as such, we are unable to identify a direct causal path between

changes in social norms and changes in giving behavior. What we show is that variation in the recipient changes both giving and injunctive norms in the same direction. In order to address this issue, future research should be conducted which varies recipient-type while keeping constant certain instrumental factors.[34]

6 Conclusion

Social norms are known to influence donation choices. Previous studies have primarily focused on using social information to manipulate perceived giving norms as a policy tool in increasing donations. Instead, in this paper, we explore the explanatory power of injunctive social norms in interpreting the variation in giving behavior to different recipients types. This leads us to several key findings.

First, we replicate results of previous studies, finding that dictator transfers are significantly higher when the recipient is a charity than when the recipient is an anonymous student.[35] Second, we show that the norms of giving differ substantially under the two contexts, in ways consistent with our expectations. Specifically, the most "socially appropriate" action for a dictator when the recipient is a charity is to transfer all of their initial endowment; in contrast, the most appropriate action when the recipient is another student is to divide the endowment equally. Third, we show that giving behavior, unconditional on recipient type, is closely related to social norms. Using conditional logit regression analysis, we demonstrate that the proxied social norm positively correlates with dictator allocation decisions. In all dictator game treatments, the modal behavior receives the highest average appropriateness rating.

Lastly, we find that while conforming to social norms by taking a more appropriate action correlates with motivates giving in both settings, there is significantly stronger norm-compliance in the context of charitable giving. Moreover, when the recipient is a charity, incremental increases in the degree of appropriateness of alternative actions no longer predict behavior, and full conformity to the most appropriate action is more likely to occur.

Overall, our results have significance for both theorists and experimentalists. Currently, there exists no unifying model that explains why giving behavior differs based upon the recipient. A model incorporating social norms, such as one presented in this paper, has the possibility to explain such behavior. That conformity to social norms drives giving behavior is appealing, and has the potential to generate simpler unifying models that can explain giving behavior under different contexts. Moreover, the fact that social norms are followed differently between contexts should influence the way norms are currently modelled in utility-maximizing frameworks. For

[34] We thank an anonymous referee for bringing this issue to our attention.

[35] Eckel and Grossman (1996) believed that their analogous result was due in part to the double-blind procedure they employed: it removed any motive for giving in the individual game, but did not remove the altruism motive for giving in the charity game. We show here that a similar result obtains without the double-blind procedure.

experimentalists and other empirical researchers relying on standard lab measures to learn about the altruistic preference of their subjects, our study illustrates a potential mechanism through which "context" can affect giving behavior. Although our study is not able to identify the direct causes of variation in social norms, measuring norms across experimentally-manipulated contexts can provide additional insight in to the mechanisms behind behavior change. We also validate the Krupka and Weber (2013) norm elicitation strategy in a two-recipient framework which should encourage the use of norm ratings in deriving a priori predictions of donation behavior.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10683-023-09811-z.

Acknowledgements We also wish thank research assistants: Andy Cao, Allegra Hernandez, Carly Mayes, Nanyin Yang, Sora Youn. Lab managers: Economic Research Lab, TAMU, David Cabrera; Behavioral Research Lab, Rice, Annie Pham. This study is funded by the National Science Foundation (For Rice University, SES-2027556; for Texas A &M: SES-2027548; for the University of Michigan: SES-2027513). We also thank, anonymous referees, Tanya Rosenblatt and participants of the Eighth Biennial Conference o Social Dilemmas for their helpful comments and feedback.

Author contributions CCE secured the funding, designed the procedures, provided lab support, collected preliminary data, provided feedback, and co-wrote the preliminary draft. RKW secured the funding, designed the procedures, provided lab support, and curated the data. HGH prepared the manuscript and conducted the formal analysis. ELK led the analysis and manuscript preparation, co-wrote and edited the manuscript, and provided feedback on the preliminary results. NS originated the idea as part of her dissertation, analyzed the first wave of data and co-wrote the preliminary draft, and helped design the second wave of data collection used in the current manuscript.

Funding This study is funded by the National Science Foundation (SES-1534403 and SES-2027556). This study has received IRB approval. Rice University IRB Numbers: IRB-FY2020-278, IRB-FY2021-114.

Data availibility The replication material for the study is available at https://doi.org/10.7302/19bt-f219.

References

Agerström, J., Carlsson, R., Nicklasson, L., & Guntell, L. (2016). Using descriptive social norms to increase charitable giving: The power of local norms. *Journal of Economic Psychology, 52*, 147–153. https://doi.org/10.1016/j.joep.2015.12.007

Andreoni, J. (2006). Leadership giving in charitable fund-raising. *Journal of Public Economic Theory, 8*(1), 1–22.

Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica, 77*(5), 1607–1636.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica, 70*(2), 737–753.
Andreoni, J., Rao, J. M., & Trachtman, H. (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy, 125*(3), 625–653.
Andreoni, J., & Sanchez, A. (2020). Fooling myself or fooling observers? Avoiding social pressures by manipulating perceptions of deservingness of others. *Economic Inquiry, 58*(1), 12–33.
Ariely, D., Bracha, A., & Meier, S. (2009). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review, 99*(1), 544–55.
Bardsley, N. (2008). Dictator game giving: Altruism or artefact? *Experimental Economics, 11*(2), 122–133. https://doi.org/10.1007/s10683-007-9172-2
Bendapudi, N., Singh, S. N., & Bendapudi, V. (1996). Enhancing helping behavior: An integrative framework for promotion planning. *Journal of Marketing, 60*(3), 33–49.
Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford: Oxford University Press.
Bracha, A., Menietti, M., & Vesterlund, L. (2011). Seeds to succeed?: Sequential giving to public projects. *Journal of Public Economics, 95*(5–6), 416–427.
Chang, D., Chen, R., & Krupka, E. (2019). Rhetoric matters: A social norms explanation for the anomaly of framing. *Games and Economic Behavior, 116*, 158–178. https://doi.org/10.1016/j.geb.2019.04.011
Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *American Economic Review, 92*(4), 1218–1221.
Cherry, T. L., & Shogren, J. F. (2008). Self-interest, sympathy and the origin of endowments. *Economics Letters, 101*(1), 69–72.
Croson, R. T. (2008). Differentiating altruism and reciprocity. *Handbook of Experimental Economics Results, 1*, 784–791.
Croson, R. T., Handy, F., & Shang, J. (2010). Gendered giving: The influence of social norms on the donation behavior of men and women. *International Journal of Nonprofit and Voluntary Sector Marketing, 15*(2), 199–213.
d'Adda, G., Drouvelis, M., & Nosenzo, D. (2016). Norm elicitation in within-subject designs: Testing for order effects. *Journal of Behavioral and Experimental Economics, 62*, 1–7.
De Quidt, J., Vesterlund, L., & Wilson, A. J. (2019). Experimenter demand effects. In *Handbook of research methods and applications in experimental economics*. Edward Elgar Publishing.
DellaVigna, S., List, J. A., & Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The Quarterly Journal of Economics, 127*(1), 1–56.
Dreber, A., Ellingsen, T., Johannesson, M., & Rand, D. G. (2013). Do people care about social context? framing effects in dictator games. *Experimental Economics, 16*(3), 349–371.
Drouvelis, M., & Marx, B. M. (2021). Dimensions of donation preferences: The structure of peer and income effects. *Experimental Economics, 24*(1), 274–302.
Eckel, C. C., & Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior, 16*(2), 181–191. https://doi.org/10.1006/game.1996.0081
Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14*(4), 583–610. https://doi.org/10.1007/s10683-011-9283-7
Erkut, H., Nosenzo, D., & Sefton, M. (2015). Identifying social norms using coordination games: Spectators vs. stakeholders. *Economics Letters, 130*, 28–31. https://doi.org/10.1016/j.econlet.2015.02.021
Falk, A. (2007). Gift exchange in the field. *Econometrica, 75*(5), 1501–1511.
Fallucchi, F., & Nosenzo, D. (2022). The coordinating power of social norms. *Experimental Economics, 25*(1), 1–25.
Feather, N. T. (1999). Judgments of deservingness: Studies in the psychology of justice and achievement. *Personality and Social Psychology Review, 3*(2), 86–107.
Fehérová, M., Heger, S., Péliová, J., Servátka, M., & Slonim, R. (2022). Increasing autonomy in charitable giving: The effect of choosing the number of recipients on donations. *Economics Letters, 217*, 110701. https://doi.org/10.1016/j.econlet.2022.110701
Fong, C. M. (2007). Evidence from an experiment on charity to welfare recipients: Reciprocity, altruism and the empathic responsiveness hypothesis. *The Economic Journal, 117*(522), 1008–1024.
Fong, C. M., & Luttmer, E. F. (2009). What determines giving to hurricane katrina victims? Experimental evidence on racial group loyalty. *American Economic Journal: Applied Economics, 1*(2), 64–87. https://doi.org/10.1257/app.1.2.64
Frey, B. S., & Meier, S. (2004). Social comparisons and pro-social behavior: Testing "conditional cooperation'' in a field experiment. *American Economic Review, 94*(5), 1717–1722.

Friedman, S., Friedman, D., & Sunder, S. (1994). *Experimental methods: A primer for economists*. Cambridge: Cambridge University Press.
Glazer, A., & Konrad, K. A. (1996). A signaling explanation for charity. *The American Economic Review, 86*(4), 1019–1028.
Grossman, P. J., & Eckel, C. C. (2015). Giving versus taking for a cause. *Economics Letters, 132*, 28–30. https://doi.org/10.1016/j.econlet.2015.04.002
Hermalin, B. E. (1998). Toward an economic theory of leadership: Leading by example. *American Economic Review, 88*, 1188–1206.
Hoffman, S. D., & Duncan, G. J. (1988). Multinomial and conditional logit discrete-choice models in demography. *Demography, 25*(3), 415–427.
Huck, S., & Rasul, I. (2011). Matched fundraising: Evidence from a natural field experiment. *Journal of Public Economics, 95*(5–6), 351–362.
Huck, S., Rasul, I., & Shephard, A. (2015). Comparing charitable fundraising schemes: Evidence from a natural field experiment and a structural model. *American Economic Journal: Economic Policy, 7*(2), 326–69.
Karlan, D., & List, J. A. (2020). How can Bill and Melinda gates increase other people's donations to fund public goods? *Journal of Public Economics, 191*, 104296.
Kessler, J. B. (2017). Announcements of support and public good provision. *American Economic Review, 107*(12), 3760–87.
Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association, 14*(3), 608–638. https://doi.org/10.1111/jeea.12152
Jackson, K. Kip. (2022). Charitable giving with stochastic income. *Journal of Behavioral and Experimental Economics, 97*, 101826. https://doi.org/10.1016/j.socec.2022.101826
Krupka, E. L., & Croson, R. T. (2016). The differential impact of social norms cues on charitable contributions. *Journal of Economic Behavior and Organization, 128*, 149–158. https://doi.org/10.1016/j.jebo.2016.05.005
Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association, 11*(3), 495–524. https://doi.org/10.1111/jeea.12006
Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives, 21*(2), 153–174. https://doi.org/10.1257/jep.21.2.153
List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy, 115*(3), 482–493. https://doi.org/10.1086/519249
List, J. A., & Lucking-Reiley, D. (2002). The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy, 110*(1), 215–233.
Martin, R., & Randal, J. (2008). How is donation behaviour affected by the donations of others? *Journal of Economic Behavior and Organization, 67*(1), 228–238. https://doi.org/10.1016/j.jebo.2007.08.001
McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.
Ottoni-Wilhelm, M., Vesterlund, L., & Xie, H. (2017). Why do people give? Testing pure and impure altruism. *American Economic Review, 107*(11), 3617–33.
Potters, J., Sefton, M., & Vesterlund, L. (2007). Leading-by-example and signaling in voluntary contribution games: An experimental study. *Economic Theory, 33*(1), 169–182.
Sanchez, A. (2022). Group identity and charitable contributions: Experimental evidence. *Journal of Economic Behavior and Organization, 194*, 542–549. https://doi.org/10.1016/j.jebo.2021.12.032
Shang, J., & Croson, R. T. (2009). A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods. *The Economic Journal, 119*(540), 1422–1439.
Sinha, N. (2021). *Essays on preference for charitable giving*. https://hdl.handle.net/1969.1/195746
Umer, H., Kurosaki, T., & Iwasaki, I. (2022). Unearned endowment and charity recipient lead to higher donations: A meta-analysis of the dictator game lab experiments. *Journal of Behavioral and Experimental Economics, 97*, 101827.
Vesterlund, L. (2003). The informational value of sequential fundraising. *Journal of Public Economics, 87*(3–4), 627–657.
Vesterlund, L. (2016). Using experimental methods to understand why and how we give to charity. *Handbook of Experimental Economics, 2*, 91–151.
Yaari, M. E., & Bar-Hillel, M. (1984). On dividing justly. *Social Choice and Welfare, 1*(1), 1–24.
Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics, 13*(1), 75–98.