Self-serving biases and public justifications in trust games

Cristina Bicchieri & Hugo Mercier

Synthese
An International Journal for Epistemology, Methodology and Philosophy of Science

ISSN 0039-7857

Synthese
DOI 10.1007/s11229-012-0192-x

Springer

Synthese
DOI 10.1007/s11229-012-0192-x

Self-serving biases and public justifications in trust games

Cristina Bicchieri · Hugo Mercier

Received: 19 July 2012 / Accepted: 19 July 2012

Abstract Often, when several norms are present and may be in conflict, individuals will display a self-serving bias, privileging the norm that best serves their interests. Xiao and Bicchieri (J Econ Psychol 31(3):456–470, 2010) tested the effects of inequality on reciprocating behavior in trust games and showed that—when inequality increases—reciprocity loses its appeal. They hypothesized that self-serving biases in choosing to privilege a particular social norm occur when the choice of that norm is publicly justifiable as reasonable, even if not optimal for one of the parties. In line with the literature on motivated reasoning, this justification should find some degree of support among third parties. The results of our experimental survey of third parties support the hypothesis that biases are not always unilateral selfish assessments. Instead, they occur when the choice to favor a particular norm is supported by a shared sense that it is a reasonable and justifiable choice.

Keywords Social norms · Trust game · Self-serving bias · Equality · Reciprocity · Public justification

1 Introduction

Norms of reciprocity are an important part of social life, and so are fairness norms. Though there exists widespread awareness of the importance of these norms in guiding social interactions, little work has been done on situations in which these important norms may be in conflict. For example, reciprocity is crucial in supporting trust, but what if reciprocity creates an inequality, or deepens an existing one? What if the drive towards equality disregards reciprocity, thus hampering trust among the parties? In an original experiment, Xiao and Bicchieri (2010) tested the effects of inequality on

C. Bicchieri (✉) · H. Mercier
University of Pennsylvania, Philadelphia, PA, USA
e-mail: cb36@sas.upenn.edu

Published online: 27 September 2012

reciprocating behavior in trust games. It was observed that when inequality increases, reciprocity loses its appeal.

Since it plays a crucial role in the present experiment, let us describe the design of the original Xiao and Bicchieri (2010) experiment in detail. There were two conditions. The baseline condition was a standard trust game. Two participants were randomly assigned the roles of investor and trustee. Each of them was given 40 'experimental dollars'.[1] The investor could retain her whole endowment, in which case the game ended, or she could transfer $10 to the trustee. If the investor transferred $10 to the trustee, this amount was multiplied by 3, so that the investor then had $30 and the trustee $70, and this was common knowledge among participants. The trustee could now send back any amount of money between $0 and $30 (in $5 increments). In the baseline condition, the most salient amounts to send back were $0 (the trustee behaved selfishly), $10 (the trustee reciprocated) and $20 (the trustee reciprocated and achieved equality, with a final gain of $50 each). The modal answer was to send back $20, which can be explained by the fact that it satisfies both equality and reciprocity norms.

This standard trust game was contrasted with a second, asymmetry condition. In this condition the investor initially received $80 while the trustee received only $40. The choices for both participants were the same, but their effects were different. Here if the investor transferred $10, she achieved equality, with each player holding $70. If the trustee sent back $0, she kept all the transferred money while also satisfying equality. If she transferred $10 or more, she reciprocated but violated equality. In the asymmetry condition, a majority of trustees sent back $0. Their violation of the reciprocity norm may result from a convergence of self-interest and alignment with the equality norm.

In a trust game the interests of investors and trustees are not fully aligned. It is therefore not surprising that trustees support an equality norm when it is in their interest to demand equality, and back reciprocity otherwise. A simple conclusion is that when several norms are present and may be in conflict, individuals will display a self-serving bias, privileging the norm that best serves their interests. In the original paper, Xiao and Bicchieri hypothesized that self-serving biases in choosing to privilege a particular social norm occur when the choice of that norm is *publicly justifiable* as reasonable, even if not optimal for one of the parties. This suggestion is supported by the literature on motivated reasoning. Many experiments have shown that people are apt to find rationalizations for behavior they want to engage in (Sanitioso et al. 1990; Snyder et al. 1979; Valdesolo and DeSteno 2008). Yet their ability to rationalize their behavior is not unlimited: they need to find a reason for their behavior that, they think, would pass muster with the relevant audience (Kunda 1990). While the criteria used to evaluate our own rationalizations are lower than those used to evaluate the reasons given by other people, they do exist (Mercier 2011; Mercier and Sperber 2011).

In the asymmetry condition of Xiao and Bicchieri (2010), most trustees behaved in a way that could be simply explained by self-interest. It is plausible however that some of the trustees who behaved selfishly thought of a justification for their behavior, namely that it respected the equality norm. In line with the literature on motivated

[1] Experimental dollars are only worth 1/6th of their nominal amount in dollars, but for the sake of convenience they will be referred to as '$'.

reasoning, this justification should find some degree of support among third parties. If a justification only satisfied the person producing it, it would be moot.

The nature of the original experiment prevented testing this hypothesis, as the quality of public justification can best be assessed by asking uninvolved third parties' take on the issue. Third parties may hold the view that a reciprocity norm trumps any consideration of equality. This would mean that trustees who sent back $0 in the asymmetry condition are not really trying to engage in justifiable behavior but merely following their self-interest. Or third parties could side with the trustees and agree that it is more important to restore equality than it is to reciprocate (at least when inequality is not strongly justified, as is the case in the original experiment). The results of the new experimental survey support our hypothesis that biases are not always unilateral selfish assessments. Instead, they occur when the choice to favor a particular norm is supported by a shared sense that it is a reasonable and justifiable choice.

2 Experiment

2.1 Design

The design of the current experiment is simple. In two conditions the participants are described one of the Xiao and Bicchieri (2010) conditions. They are then asked questions about what they think the investors and the trustees *will do* (empirical expectations) and *should do* (normative expectations). This is important for several reasons (Bicchieri 2006; Bicchieri and Chavez 2010). On the one hand, we want to measure that a social norm is in fact present. If individuals' normative expectations are mutually consistent, we can be reasonably confident that there exists a shared norm. When we assess the normative expectations of decision-makers (in this case, investors and trustees), we want to distinguish personal normative expectations ("what do you think is fair") and non-personal normative expectations ("what do you believe others think is fair"), since the latter are strongly correlated with behavior, whereas the former are not (Fishbein 1967; Bicchieri and Muldoon 2011). In the case of third parties, there is no need to ask for non-personal normative expectations, as we want to assess whether the non-personal normative expectations of the players are in line with third parties, personal normative expectations.

It is usually the case that, when a norm is present, empirical and normative expectations converge. Shared empirical expectations tell us what 'normal', expected behavior is. So for example in a situation where a norm of reciprocity is salient (as in a simple trust game) we find agreement among empirical and normative expectations. In ambiguous situations, and especially when more than one norm may be relevant, empirical and normative expectations may diverge (Bicchieri and Xiao 2009). This divergence may point to the fact that there is a shared sense that a particular social norm may not be followed in that particular occasion, either because other norms are more salient, or because there is a shared recognition that the incentives to deviate from the norm are too powerful. In our experiment, when different norms compete for salience, assessing third parties expectations allows us to understand the nature of the self-serving choices displayed by trustees in the original game.

2.2 Participants

One hundred and nine participants were recruited through Amazon's Mechanical Turk Website (Buhrmester et al. 2011; DeScioli and Kurzban 2009; Paolacci et al. 2010). Forty percent were male and 60 % female. Participants' mean age was 33.3 years (SD = 12.1). Ninety-five percent had some college education and 61 % had more than a college degree. Participants were paid 10 cents for their participation, a typical amount for a task of this duration on Mechanical Turk.

2.3 Procedure

After accepting to take part in the experiment on Mechanical Turk, participants were directed to a questionnaire designed using the Qualtrics software. Participants were told that they would be described a psychology experiment and that they would have to give their opinions about the behavior of the participants in the experiment. Participants were then randomly assigned to one of two conditions. In the baseline condition participants were described the baseline condition of the Xiao and Bicchieri (2010) experiment, as described above. In the asymmetry condition, participants were described the asymmetry condition of the Xiao and Bicchieri (2010) experiment. In both cases the players were simply identified as 'participant A' and 'participant B', not as 'investor' and 'trustee'. Third parties were not told the actual results of the original experiment.

So as not to render the description too cumbersome, the amounts were described as real dollars and not experimental dollars. In the Xiao and Bicchieri (2010) experiment, participants only gained in real dollars 1/6th of the amount gained in experiment dollars. We ran a version of the current experiment with the real dollars amount (i.e. everything divided by 6), and observed no significant difference in the participants' answers.

After the experiment was described, participants were asked a series of 10 control questions to ensure they had properly understood the experiment. Only those participants who answered all 10 questions accurately were able to proceed to the actual experiment.

Participants answered four questions, two about the behavior of 'participant A' (the investor) and two about the behavior of 'participant B' (the trustee). Their empirical expectations were assessed by a 'will' question: 'what do you think participant A [B] *will* transfer [send back]?' Their normative expectations were assessed by an 'ought' question: 'what do you think participant A [B] *ought to* transfer [send back]?'

3 Results

3.1 Empirical and normative expectations about investors' behavior

3.1.1 Baseline condition

In the original experiment, 65 % of investors in the baseline condition transferred $10. 75 % of third parties believe this is what investors ought to do, whereas only 49 % of

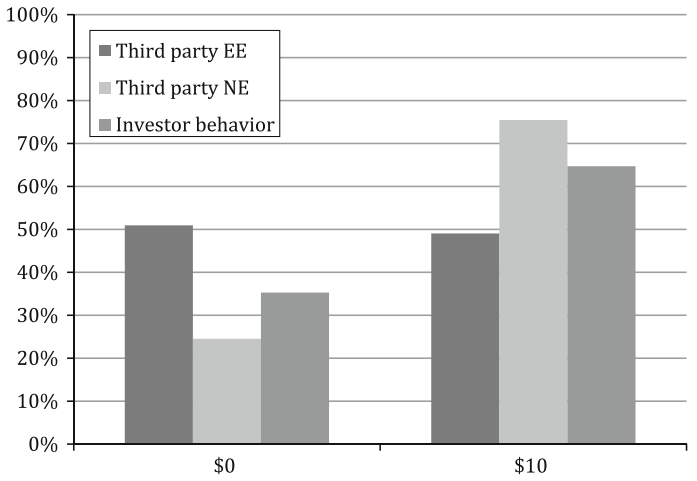

Fig. 1 Empirical and normative expectations of third parties and actual behavior of investors in the baseline condition

them expect this transfer to actually occur.[2] Indeed, we observed a bimodal distribution, with third parties almost equally divided between expecting transfers of zero or $10.

3.1.2 Asymmetry condition

In the asymmetry condition, a similar number of investor (61 %) transferred $10, and 88 % of third parties believed this is what ought to be done, whereas 61 % of them correctly expected this transfer to actually occur.

Figures 1 and 2 plot the expectations of third parties regarding investor behavior as well as the actual investor behavior in the Xiao and Bicchieri (2010) experiment (data provided by the authors). The comparisons across conditions do not yield any significant results (Fisher test, all $ps > .2$). In line with the absence of a statistically significant difference between conditions for actual investor behavior, third parties' empirical and normative expectations did not significantly differ by condition either. Third parties were able to predict relatively accurately investor behavior, as there was no significant difference between third parties' empirical expectations and investor behavior (Fisher test, both $ps > .15$).

There was no statistically significant difference between third parties normative expectations and investor behavior in the baseline condition (Fisher test, $ps > .3$), but a significant difference emerged in the asymmetry condition (Fisher test, $ps < .01$). In this condition, investors invested less than third parties though they should. Finally, third parties' normative expectations about a $10 transfer were significantly higher than their empirical expectations in both conditions (Fisher test, both $ps < .01$).

[2] It is interesting to note that, in the original Xiao and Bicchieri experiment, investors' empirical expectations about what other investors would do were consistent with actual behavior, as 61 % of investors believed that the majority of other investors would transfer $10 in the baseline condition.

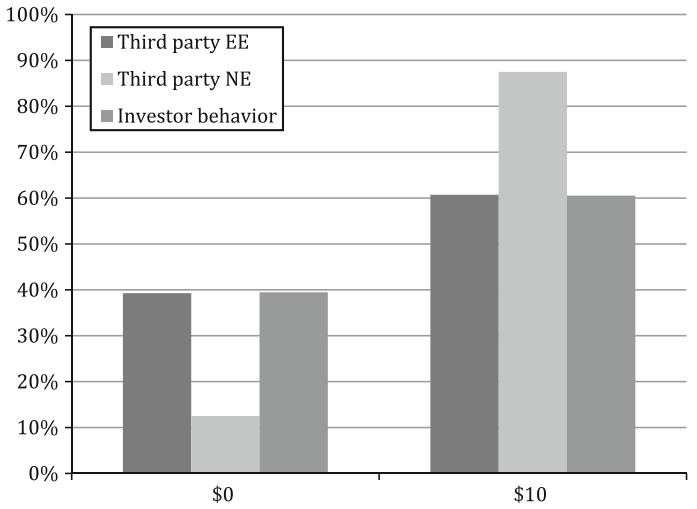

Fig. 2 Empirical and normative expectations of third parties and behavior of investors in the asymmetry condition

Normative expectations can reflect a prudential ought—investors ought to invest because they can expect returns—or a moral ought—investors ought to invest because it is the right thing to do. To evaluate these two possibilities we started by isolating those participants who had a normative expectation that the investors *ought to* transfer \$10. If these third parties expect trustees to send back more than \$10 on average, with an acceptable degree of risk, then the ought could be prudential. By contrast, if these third parties expect investors to return \$10 or less, then the "ought" cannot be prudential and has to be moral. We observe that the (empirically) expected return is \$10.4 (SD = 1.75) in the baseline condition and \$5.20 (SD = 1.39) in the asymmetry condition. Especially in the latter condition, it is very clear that the normative expectation is a moral one, as the expected return would result in a net loss on the part of the investor, who is however normatively expected to reduce the initial state of inequality.

3.2 Expectations about trustee behavior

3.2.1 Baseline condition

Our first analysis simply partitions the results between amounts below \$10—when the investor gets less money than she invested—and amounts equal to or above \$10—when the investor is made whole or gains some money. As can be seen in Fig. 3, third parties had relatively accurate empirical expectations about trustees' behavior (Fisher test, $p > .4$). By contrast, their normative expectations of reciprocity (a back-transfer of \$10 or more) were significantly higher than their empirical expectations (89 vs. 65 %; Fisher test, $p < .01$) and also very high compared to the trustees' actual behavior (Fisher test, $p < .01$).

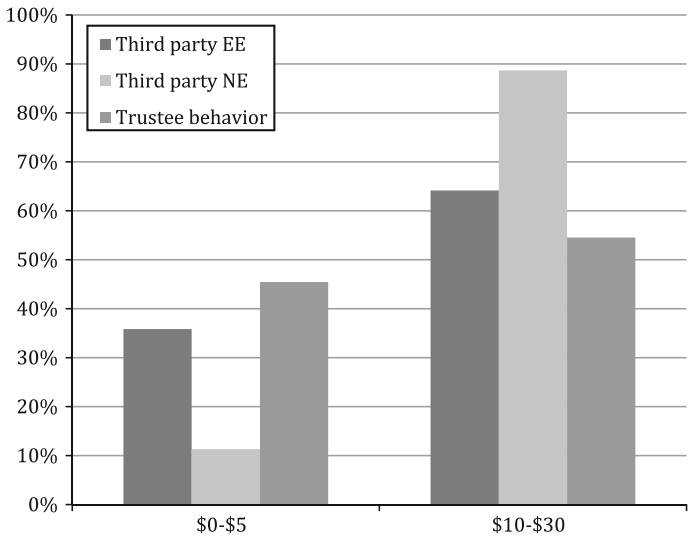

Fig. 3 Expectations of third parties and behavior of trustees in the baseline condition, below and above reciprocation

Table 1 Players' gains as a function of the amount sent back by the trustee in the baseline condition

Amount sent back	Final gains as a function of amount sent back		Description of the outcome
$0	$30	$70	Less than reciprocation
$5	$35	$65	
$10	$40	$60	Reciprocation but less than equality
$15	$45	$55	
$20	$50	$50	Equality
$25	$55	$45	More than equality
$30	$60	$40	

It should be noted that, in the baseline condition of the original experiment (Xiao and Bicchieri 2010), where a majority of trustees (55 %) returned $10 or more, their empirical expectations about other trustees were in line with their actual behavior. Investors, too, expected a majority of trustees to back-transfer $10 or more, as did third parties. Yet both investors and third parties had normative expectations well above their realistic empirical expectations. As we shall see below, these normative expectations clustered around the back-transfer of $20, the one that equalizes the payoffs of investors and trustees. A large majority of trustees themselves, when asked their second order beliefs about investors' normative expectations, correctly assessed that a $20 back-transfer was thought to be the fair choice. Yet only 41 % transferred back such amount.

We thus performed a further analysis, looking at more fine-grained behavior, as explained in Table 1 and plotted in Fig. 4.

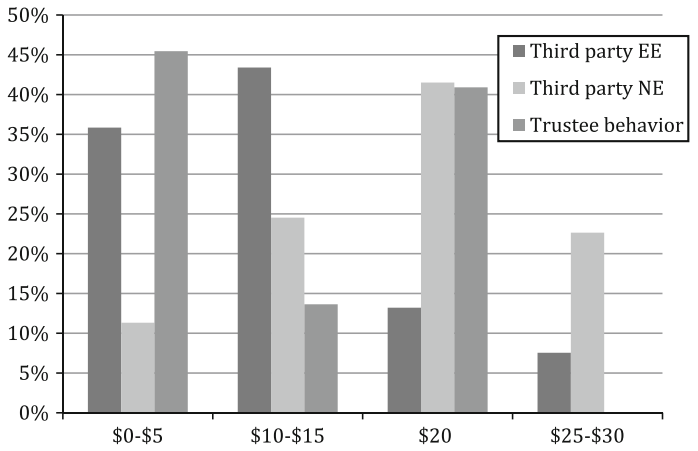

Fig. 4 Empirical and normative expectations of third parties and actual behavior of trustees in the baseline condition, fine-grained analysis

Focusing on amounts equal or greater than \$10, we observed a strong modal response of the trustees at \$20 (41 % of trustees returned \$20). It should be further noted that a back-transfer of \$20 ensures equality among investor and trustee, while at the same time it provides the investor with a gain of \$10, thus safeguarding reciprocity.

This behavior did not match third parties' empirical expectations (Fisher test, $p < .05$), as only 13 % of third parties expected a back-transfer of \$20 to occur. Instead, 43 % of third parties expected trustees to send back a lower amount (\$10 or \$15) more often than the trustees actually did (Fisher test, $p < .05$), as only 14 % of trustees actually sent back \$10 or \$15. By contrast, the modal response of trustees at \$20 fit with third parties' normative expectations (Fisher test, $p = 1$). The only difference at or above the \$10 threshold is that third parties thought trustees ought to give \$25 or \$30 (acting beyond equality) more than they actually did (Fisher test, $p < .05$). In sum, it clearly appears that a sizable majority of third parties (65 %) believes that trustees ought to give back \$20 or more. In this case, both fairness (as equality) and reciprocity norms seem to play a role in third parties' normative assessment, and it is impossible to tease apart the effects of each one.

3.3 Asymmetric condition

Again, we began the analysis by looking at the coarser partition between amounts below \$10 (less than reciprocation) and those equal or above \$10 (reciprocation in various degrees) (see Fig. 5). The differences between third parties' expectations and trustee behavior mirrors those observed in the baseline condition. Third parties' empirical expectations were relatively accurate (Fisher test, $p > .5$) but their normative expectations were significantly higher than both their empirical expectations (Fisher test, $p < .05$) and actual trustee behavior in the case of \$10 or more back-transfer (Fisher test, $p < .05$). Indeed, 57 % of third parties believe that trustees ought to return \$10 or more in the asymmetric case, whereas a lower 33 % expects this to happen.

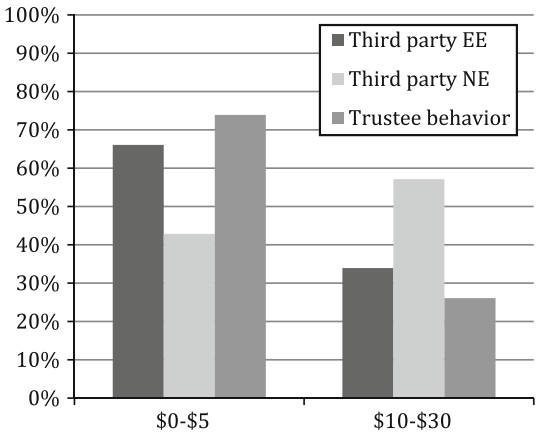

Fig. 5 Empirical and normative expectations of third parties and behavior of trustees in the asymmetric condition, below and above reciprocation

Table 2 Players' gains as a function of the amount sent back by the trustee in the asymmetric condition

Amount sent back	Final gains as a function of amount sent back		Description of the outcome
$0	$70	$70	Equality/selfishness
$5	$75	$65	Less than reciprocation
$10	$80	$60	Reciprocation and more than reciprocation
$15	$85	$55	
$20	$90	$50	
$25	$95	$45	
$30	$100	$40	

These expectations are in line with investors' expectations in the original experiment of Xiao and Bicchieri. 54 % of third parties expect trustees to return zero, comparable with 63 % of investors holding this belief (note: 61 % of trustees return zero in the asymmetric case). As to the normative expectations of both groups, 62 % of investors thought that trustees should return $10 or more, and 57 % of third parties held the same belief. If we look at trustees' second-order beliefs in the original experiment, we observe that 69 % of trustees believe that investors expect no return, a fairly accurate belief. However, when we look at trustees' beliefs about investors' normative expectations, we observe a bimodal pattern. 41 % of trustees believe that investors find a zero return fair, and 51 % believe investors consider fair a return of $10 or more. This split reflects the ambiguity of the situation, since in the asymmetric case a norm of reciprocity is in direct conflict with an equality norm.

The more fine-grained analysis will be conducted differently in this condition as the amounts sent back have different meanings in each condition (see Table 2). The results are plotted in Fig. 6.

From this further analysis it is apparent that the gap between empirical expectations, on the one hand, and normative expectations and trustee behavior, on the other, is driven by differing expectations about giving either $0, or $10 or more. 54 % of third

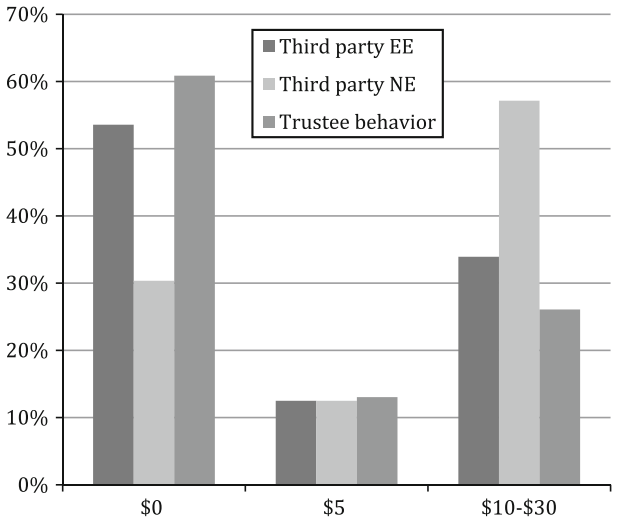

Fig. 6 Empirical and normative expectations of third parties and behavior of trustees in the asymmetric condition, fine-grained analysis

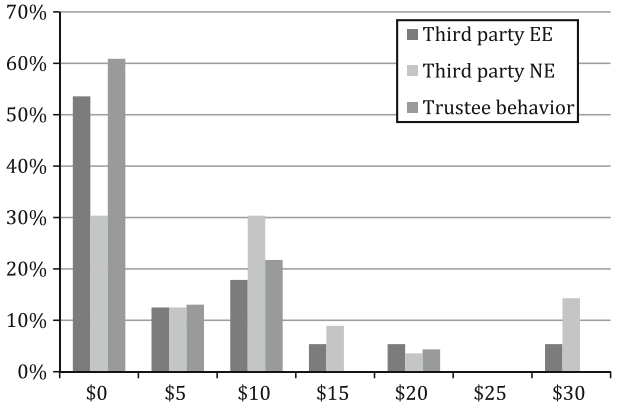

Fig. 7 Empirical and normative expectations of third parties and behavior of trustees in the asymmetric condition, complete distribution of back-transfers

parties expected that a majority of trustees would give back \$0, and indeed 61 % of trustees gave back nothing (Fisher test, $p > .6$). But 57 % thought that trustees ought to give \$10 or more, and only 26 % of trustees displayed reciprocating behavior. It looks as if third parties had realistic expectations about trustees behavior, at odds with the normative demands they imposed on them.

Finally, in the complete distribution of back-transfers (Fig. 7) we observed a peak at \$10 of similar size across third parties' empirical and normative expectations and trustee behavior (all three comparisons, Fisher test, $p > .18$).

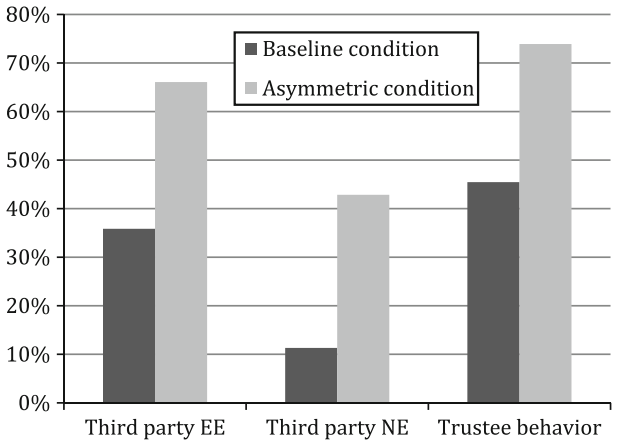

Fig. 8 Percentage of $0 or $5 expectations and behavior, as opposed to $10 or more, across the two conditions

3.4 Comparison of the two conditions

Figure 8 displays the percentage of third parties expectations, normative and empirical, below reciprocation ($0 or $5), and trustee behavior. Expectations of back-transfers lower than $10 were more common in the asymmetric condition than in the baseline condition (Fisher test, third party empirical expectation, $p < .01$, third party normative expectation, $p < .001$, trustee behavior, $p = .07$). The higher normative expectations of zero return in the asymmetric condition (30 vs. 11% in the baseline) are in line with the investors' normative expectations, and with trustees second-order beliefs that a zero return is normatively acceptable for investors (Xiao and Bicchieri 2010).

4 Discussion

4.1 Investor behavior

In both conditions, investor behavior was in line with third parties' empirical expectations. In the asymmetric condition, however, third parties thought investors should transfer $10 more often than they did. This finding is consistent with previous findings showing that empirical expectations tend to have stronger predictive power about (and influence on) behavior than normative expectations (Bicchieri and Xiao 2009). Indeed, in the original experiment of Xiao and Bicchieri, investors were asked how much they expected other investors to invest in both conditions. Investors' empirical expectations about other participants were in line with their actual behavior.

Normative expectations were significantly different from empirical expectations in both conditions, as third parties thought that investors ought to transfer $10 much more often than they in fact expected it to occur. Two considerations support the interpretation of this gap in terms of a moral ought—i.e. third parties believed investors should transfer the money because it was the right thing to do, and not because they

could expect to profit from it. The first is that third parties did not expect trustees to return significantly more than $10. The second is that even though third parties expected significantly less back-transfer in the asymmetric condition than in the baseline condition, they did not adjust their normative expectations of investor transfer accordingly.

We can thus explain the significant gap between normative expectations and actual behavior in the asymmetric condition by the combination of a more pressing moral imperative—due to the unequal endowment—that explains the high normative expectations, and the smaller expected returns that explain the decrease in transfer frequency.

4.2 Trustee behavior

4.2.1 Baseline condition

The overall behavior of the trustees in the baseline condition was more in line with third parties' empirical expectations than with their normative expectations.[3] Again, this is consistent with a stronger influence and predictive power of the former. However, normative expectations were not disregarded by all trustees. If we focus our attention on those trustees who did not act selfishly—those who sent back $10 or more to the investor—we observe that by sending back $20 instead of $10, they followed the normative expectations but not the empirical expectations. One way to interpret this result is as a balance between choosing to act selfishly, completely disregarding any norm, and choosing to follow a norm of reciprocity. Here the strong normative expectation is to give back $20, so those trustees who chose to follow a norm must have felt compelled to follow the most salient norm. By comparison, sending back $10 may have only achieved the worst of both worlds: one would be perceived as having violated the dominant norm, but without the benefit of keeping all (or nearly all) the money. It is noteworthy that third parties' empirical expectations were out of line with players' behavior. This is the only major deviation between behavior and empirical expectations observed in this experiment.

4.2.2 Asymmetry condition

Here again the overall behavior of the trustees is coherent with a superior pull of empirical expectations as opposed to normative expectations, as the trustees behaved in line with the former but not the later. As in the baseline condition, those participants who did not act selfishly followed the dominant norm, here to send back $10.

[3] Incidentally, trustees behavior was consistent with their own empirical expectations of other trustees' behavior, with investors' empirical expectations, and with their own second-degree beliefs about what investors expected them to do.

4.2.3 Comparison of the two conditions

The question that prompted this experiment was the following: in the asymmetry condition, are trustees taking advantage of a potential confusion between norms of equality and reciprocity to behave selfishly? To answer this question, we must look at the difference of trustee behavior in the two conditions and compare this difference to the difference in third parties normative expectations. We know that trustees behave more selfishly when given the opportunity in the asymmetry condition. If third parties' normative expectations were similar across conditions, this would mean that third parties judge selfish acts as harshly in all cases and that the trustees' selfish behavior is not based on different norms. But this is not the case. Across conditions, the change in third parties' normative expectations was commensurate to the change in trustee behavior. Twenty-nine percents more trustees sent back \$0 or \$5 in the asymmetry condition than in the baseline condition, and 33 % more third parties said that such a behavior was normatively justified. We can conclude that trustees are probably simply reacting to a genuine change in the dominant norms rather than trying to take advantage of a change in norms to behave more selfishly.

5 General discussion

People do not always follow their naked self-interest, neither in real life nor in economic games. One of the factors that explain these 'deviations' is norms. When people believe a norm dictates them to behave in a certain way, they are more likely to behave that way, even at a material cost to themselves (Bicchieri 2006). It is not always clear, however, which norm applies in any given context. In Xiao and Bicchieri's variation on the trust game, trustees could abide by one of two norms: a reciprocity norm or an equality norm. In contrast with standard trust games, following the equality norm would allow trustees to also maximize the amount of money earned, at the cost of violating the reciprocity norm. Trustees were indeed more likely to send back \$0 in this variant than in the standard trust game.

The behavior of the trustees could be explained as simply maximizing their monetary payoff. But we know that norms exert a powerful influence on behavior in economic games. It is therefore more likely that norms also impacted their behavior, even when it was undistinguishable from self-interest. In particular, the equality norm could have played a role by allowing trustees to justify their departure from the reciprocity norm and selfish behavior. How could we know that the norm did play a role in the trustees' behavior? A considerable literature in motivated reasoning shows that people have minimal standards when it comes to self-justifications: not every behavior can be adequately rationalized (Kunda 1990; Mercier and Sperber 2011). If a justification is deemed sufficient by the individual producing it, it should also pass muster at least for a portion of its potential audience.

If the trustees relied on a norm of equality to justify their selfish behavior, then a non-negligible share of third parties should deem this behavior to be acceptable. It is indeed what was observed in the present experiment: about a third of third parties thought that trustees *should* return \$0, and 40 % thought that failing to reciprocate (by

sending $0 or $5) was the normatively correct behavior. These percentages are significantly higher than what is observed when third parties judge the behavior of trustees in a standard trust game. The current experiment also confirmed an important finding: behavior is more often in line with empirical expectations (what we think people will do) than with normative expectations (what we think people should do). In all but one instance, the behavior of the investors and the trustees was in line with the empirical expectations, whereas it was often at odds with normative expectations. Asking third parties for their opinions on the behavior of players in economic games can lead to a better understanding of the psychological mechanisms behind these behaviors.

References

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge: Cambridge University Press.

Bicchieri, C., & Chavez, A. (2010). Behaving as expected: Public information and fairness norms. *Journal of Behavioral Decision Making, 23*(2), 161–178.

Bicchieri, C., & Muldoon, R. (2011). Social norms. In *The Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/social-norms/.

Bicchieri, C., & Xiao, E. (2009). Do the right thing: But only if others do so. *Journal of Behavioral Decision Making, 22*, 191–208.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk. *Perspectives on Psychological Science, 6*(1), 3.

DeScioli, P., & Kurzban, R. (2009). The alliance hypothesis for human friendship. *PloS ONE, 4*(6), e5802.

Fishbein, M. (1967). A consideration of beliefs and their role in attitude measurement. In M. Fishbein (Ed.), *Readings in attitude theory and measurement*. New York: Wiley.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*, 480–498.

Mercier, H. (2011). What good is moral reasoning? *Mind & Society, 10*(2), 131–148.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(2), 57–74.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5*(5), 411–419.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59*(2), 229–241.

Snyder, M., Kleck, R. E., Strenta, A., & Mentzer, S. J. (1979). Avoidance of the handicapped: An attributional ambiguity analysis. *Journal of Personality and Social Psychology, 37*(12), 2297–2306.

Valdesolo, P., & DeSteno, D. (2008). The duality of virtue: Deconstructing the moral hypocrite. *Journal of Experimental Social Psychology, 44*, 1334–1338.

Xiao, E., & Bicchieri, C. (2010). When equality trumps reciprocity. *Journal of Economic Psychology, 31*(3), 456–470.