

**To know or not to know? Looking at payoffs signals selfish
behavior, but it does not actually mean so**

Valerio Capraro

Middlesex University London

Jotte Kuilder

University of Amsterdam

In press at the Journal of Behavioral and Experimental Economics

Capraro, V., & Kuilder, J. (in press). To know or not to know? Looking at payoffs signals selfish behavior, but it does not actually mean so. *Journal of Behavioral and Experimental Economics*.

Abstract

Costs and benefits of everyday actions are often not known beforehand. In such situations, people can either make a choice “without looking” at the payoffs, or they can “look” and learn the exact payoffs involved before making the actual choice. Recent studies suggest that the mere act of looking at payoffs will be met with distrust by observers: “lookers” are both less trustworthy and perceived to be less trustworthy than “non-lookers”. Here we extend this line of work by changing the measure of pro-sociality: instead of trustworthiness, we consider altruism. Does “looking at payoffs” signal self-regarding preferences? Do observers’ beliefs match decision makers’ actions? Two experiments demonstrate that: (i) the level of altruism among “lookers” is not different from the level of altruism among “non-lookers”, but (ii) “lookers” are perceived to be less altruistic than “non-lookers”. These results hold both when the measure of altruism is the choice whether to help or not in the so-called “envelope game” (Experiment 1) or when the measure of altruism is the donation in a standard Dictator Game (Experiment 2). In sum, these results uncover a perception gap according to which looking at payoffs signals selfish behavior, but it does not actually mean so.

Keywords: helping, altruism, help without looking, envelope game, perception gap, Dictator Game.

Introduction

Virtually all studies on human pro-sociality assume that decision makers know the exact costs and benefits of a pro-social action beforehand. While this assumption is helpful to develop theoretical models (Fehr & Schmidt, 1999; Bolton & Ockenfels, 2000; Charness & Rabin, 2002; Capraro, 2013) and conduct behavioral experiments (Rapoport, 1965; Kahneman, Knetsch & Thaler, 1986; Camerer, 2003), in many everyday situations people do not actually know the exact payoffs involved beforehand, but can gather this information only in a subsequent stage.

Such situations abound in real life. For example, when a friend asks you to drive her to some store, before making your decision, you can decide to ask for additional information to learn the exact cost of helping (How far is the store? How long will it take?). Similarly, when a friend tells you he is in trouble and needs a temporary loan, before making your decision, you may ask him the exact amount he needs and when he expects to return it. Analogously, before deciding whether to join an ethical cause, you might or might not decide to gather additional information about how much effort (time and money) you need to invest for this cause.

One simple way to formalize this type of situations is by means of a two-stage decision problem with one decision-maker and one observer (see Figure 1). Initially, decision-makers know that they will have to decide between “helping” and “not helping”. They know that helping leads to a fair payoff, while not helping is individually optimal. However, they do not know the exact payoffs. In Stage 1, decision-makers have to decide between “looking at payoffs” and “not looking at payoffs”. In case a decision-maker decides to look, she or he learns the complete payoff structure of the game, that is, she or he learns the payoffs for both players. Then, in Stage 2, decision-makers make their actual choice between helping and not helping.

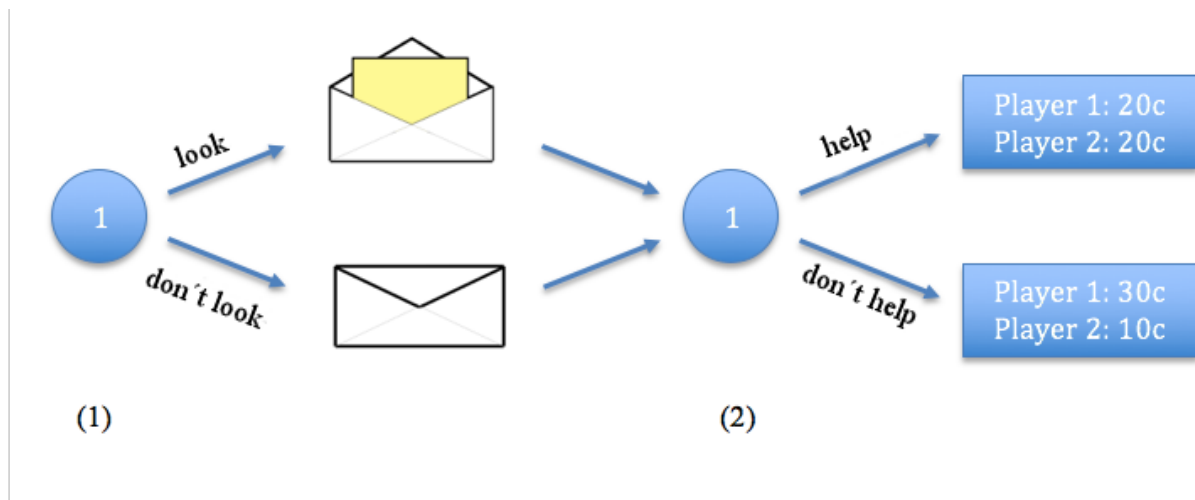


Figure 1. The “envelope game” that we are going to use in Experiment 1. In Stage (1), the decision maker, Player 1, decides whether to look at the payoffs corresponding to “help” and “don’t help”, or not. If Player 1 decides to look at the payoffs, then she or he is informed about the complete payoff structure of the game, that is, she or he learns the payoffs for both players in both situations. In stage (2), Player 1 decides whether to help or not. Player 2 is passive and has no influence on the result of the game.

This decision problem is similar to the envelope game introduced for iterated interactions by Hoffman, Yoeli and Nowak (2015) and for one-shot games by Hilbe, Hoffman and Nowak (2015). These theoretical studies posit that “the mere act of considering one’s strategic options and gathering information about the possible costs and benefits of an action will be met with distrust” (Hilbe, Hoffman & Nowak, 2015).

In agreement with this intuition, Jordan, Hoffman, Nowak and Rand (2016) found that “lookers” are both less trustworthy and perceived to be less trustworthy than “non-lookers” in a standard Trust Game.

Here we aim at extending this line of work by changing the measure of pro-sociality. Instead of asking ourselves whether “lookers” are less trustworthy than “non-lookers”, we ask

ourselves whether they are less *altruistic* than non-lookers, where altruism is defined as unilaterally giving money to an anonymous stranger.

There is indeed one theoretical reason to expect that considering altruism instead of trustworthiness will lead to results which differ from those by Jordan et al. (2016). Hoffman et al. (2015) suggest that looking at payoffs will be met with distrust by observers because looking at payoffs signals that the decision maker is deliberating about the available choices; whereas, “non-looking at payoffs” signals that the decision maker is not deliberating and so he or she will be more likely to use heuristics (i.e., automatic choices that people internalize, because they are payoff-maximizing in real life interactions, and use as default strategies when they have no opportunity to reason about all available options). While both trustworthiness and altruism are not payoff-maximizing in one-shot anonymous laboratory experiments, they behave differently in terms of heuristics internalized from everyday interactions. Since most people’s everyday interactions are with friends, family members, and coworkers, and thus they are repeated, people tend to internalize heuristics that are optimal in iterated games (Rand, Greene & Nowak, 2012; Rand et al., 2014; Rand et al., 2016; Capraro & Cococcioni, 2015). The crucial difference between trustworthiness and altruism is that the former is optimal in the long run, while the latter is not. Specifically, the standard way to measure trustworthiness is through the Trust Game, which is a positive-sum game. Thus, it is optimal in the long-run to be trustworthy, because the benefit of partner’s reciprocal trust overcome the initial cost of trustworthiness. Indeed, a meta-analysis of studies exploring the effect of promoting intuition versus deliberation in positive-sum games, including the Trust Game, shows that subjects tend to internalize trustworthiness as their default strategy (Rand, in press). This logic does not apply to altruism: unilaterally giving money is zero-sum, and thus altruistic behavior is never optimal, neither in one-shot, nor in iterated games. Consequently, the amount of altruism should not depend on whether a person is deliberating

or not. Indeed, a recent meta-analysis of 22 experiments has shown that promoting intuition over deliberation has no effect on altruistic behavior (Rand et al., 2016).

Taking into account this theoretical difference between altruism and trustworthiness, we formulate the following hypothesis.

Hypothesis. Subjects who look at payoffs are not significantly less altruistic than those who do not look at payoffs.

We test this hypothesis by means of two experiments, which demonstrate that indeed (i) “lookers” are not less altruistic than “non-lookers”. Interestingly, in doing so, we also uncover a perception bias according to which (ii) “lookers” are perceived to be far more selfish than they actually are.

In sum, we show that looking at payoffs signals selfish behavior, but it does not actually mean so.

Experiment 1

As mentioned above, we aim to (i) measure a possible change in observers’ beliefs about decision-makers’ levels of altruism, caused by knowing that the decision-maker has decided to look at the cost of helping before making their decision, and (ii) measure whether a possible change in observers’ beliefs corresponds to a change in decision-makers’ actual behavior.

Method

Subjects were living in the US at the time of the experiment and were recruited using the online labor market Amazon Mechanical Turk (Paolacci, Chandler & Ipeirotis 2010; Horton, Rand & Zeckhauser 2011, Paolacci & Chandler, 2014). In none of the studies presented in this paper an a priori power analysis was conducted, but the planned sample sizes

were based on previous studies investigating behavioral changes in games involving pro-social behavior (Capraro, Jordan & Rand, 2014).

Each of 1,088 participants (57% males, mean age=32) was randomly assigned to one of seven conditions and passed standard comprehension questions to make sure they understood the decision problem at hand. Any subjects that did not pass the comprehension questions were automatically excluded from the survey. The seven experimental conditions were as follows.

Received. In this condition all participants were decision makers, to whom we asked to decide between Option A and Option B. Option A would give 20c to both themselves and the person they were paired with (participating in the *Guess Received* condition described below). Option B would give 30c to themselves and 10c to the other person.

Denied. This condition was similar to the *Received* condition, but decision makers (paired with participants in the *Guess Denied* condition) were not told the payoffs corresponding to the two options. Moreover, participants were not given the choice to learn them. The only information they had was that, if they chose Option A, then both themselves and the other participant would receive the same amount of money, while, if they chose Option B, they would get more money at the expenses of the other participant.

Choose. This condition was similar to the *Denied* condition. Participants were told that Option A would allocate the same amount of money to themselves and the other participant, but Option B would maximize their payoff at the expenses of the other participant. After giving this piece of information, we asked participants whether they wanted to know the exact amounts of money corresponding to each of the two options. As in Jordan et al. (2016), finding out the cost of helping was costless. In order to minimize confounding factors (e.g., moral cleansing), participants were not told that their decision whether to look at payoffs or

not would be communicated to an observer. Depending on their choice, these participants were paired with observers in the *Guess Chose Yes* and *Guess Chose No* conditions.

Guess Received. Each participant was grouped together with two other participants, named Person 1 and Person 2 (Person 1 participating in the *Received* condition and Person 2 participating in the *Guess Received* condition). Participants were shown the screenshots of the instructions presented to Person 1 and had to guess Person 1's decision. Correct guesses were incentivized with a \$0.20 prize. We measured beliefs from the point of view of a third-party to avoid that risk-averse people *edge*, that is, they bet on a donation of 0 thinking that, in case they lost, they would at least receive the donation (d'Adda, Capraro & Tavoni, 2015).

Guess Denied. This condition was similar to the *Guess Received* condition, but participants were shown the screenshots of the instructions given to decision makers participating in the *Denied* condition.

Guess Chose Yes. This condition was similar to the previous ones, but participants were shown the screenshots of the instructions given to decision makers participating in the *Choose* condition, who decided to look at payoffs. Observers were communicated also the payoff structure of the game, and not only that the decision maker decided to look at the payoffs.

Guess Chose No. This condition was similar to the previous ones, but participants were shown the screenshots of the instructions given to decision makers participating in the *Choose* condition, who decided *not* to look at payoffs.

Results and discussion

Visual evidence of the results is provided in Figure 2. Apart from the condition in which decision makers were given the choice to look at the payoffs and decided to do so, observers' beliefs about decision makers' behavior look very accurate. The only apparent gap

regards the condition in which decision makers decided to look at the payoffs. The figure suggests that looking at payoffs signals selfish behavior, but it does not actually mean so.

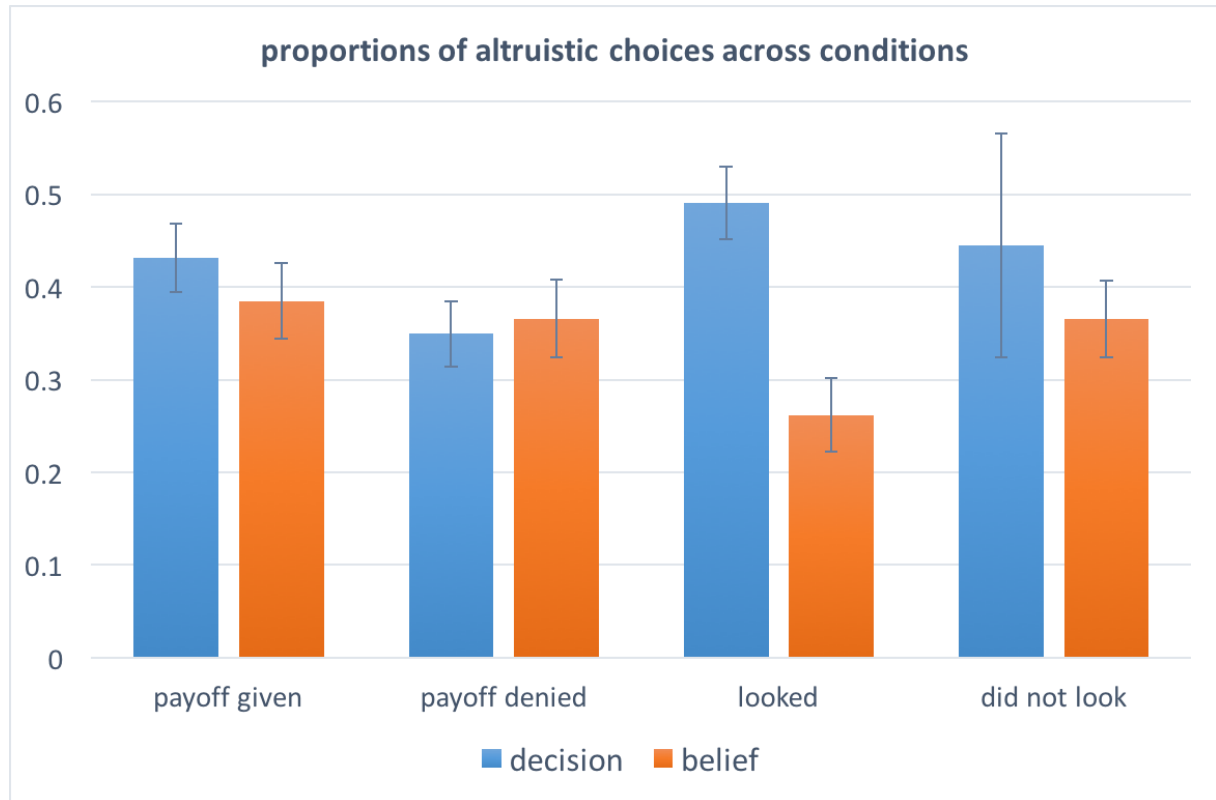


Figure 2. Proportions of altruistic choices across conditions. Error bars denote the standard error of the mean. The first, the second, and the fourth pair of columns provide visual evidence that observers' beliefs about decision makers' behavior match actual choices in all conditions, save the one in which decision makers decided to look at the payoffs before making their decision. The third pair of columns indeed provides evidence that participants looking at the payoffs were perceived to be highly more selfish than they actually are. We refer to the Main Text for detailed statistical analysis.

To confirm this intuition, we now report formal statistical analysis. We define a variable, named “guess”, which takes value 1 if a subject played as observer, and 0 if a subject played as Person 1. Next, we define a variable, named “helping”, which takes value 1

for subjects who played in the role of Person 1 and decided to help, and for subjects who played as observers and guessed that the corresponding Person 1 would help; and takes value 0 otherwise. Logistic regression predicting the probability of “helping” as a function of “guess” shows that in the cases of denied payoffs ($\chi^2(1,320) = 3.817$, coeff = 0.071, $z = -0.299$, $p = 0.765$, difference between means = 2%), received payoffs ($\chi^2(1,324) = 2.123$, coeff = -0.192, $z = 0.841$, $p = 0.400$, difference between means = 5%) and choosing not to know the payoffs ($\chi^2(1,155) = 1.800$, coeff = -0.330, $z = 0.653$, $p = 0.513$, difference between means = 7%), there is no statistically significant difference between the choices of decision makers and observers’ beliefs. Hence the beliefs about Player 1s’ actions are not significantly different from their actual actions. However, results differ when the decision maker chooses to know the payoffs. In this case, there is a strong statistically significant difference between the decisions of the participants acting as Player 1 and observers’ predictions of their choices ($\chi^2(1,289) = 17.90$), coeff = -0.999, $z = 3.901$, $p < .0001$), with a difference between mean of 23%. Taken together, these results provide evidence for a strong gap according to which subjects looking at the payoffs are perceived more selfish than they actually are, and, importantly, this gap is not due to an underlying gap regarding the amount of altruism in others.

Next we examine whether this gap is driven by Player 1’s actual altruism or by the beliefs about Player 1’s altruism, or both. To do so, we compare the *Guess Received* condition and the *Guess Chose Yes* condition by conducting logistic regression predicting “helping” as a function of a dummy variable, which takes value 1 if the subject participated in the *Guess Chose Yes* condition, and 0 otherwise. Results show that observers believe that “lookers” are less altruistic than subjects to whom payoffs are given by default ($\chi^2(1,269) = 9.175$, coeff = -0.567, $z = 2.130$, $p = 0.033$), with a difference between means of 12.5%. Similarly, we compare participants in the *Received* condition and participants in the *Choose* condition, who

decided to look at the payoffs. Logistic regression shows that participants who looked at the payoffs tend to be more altruistic, but not significantly so ($\chi^2(1,344) = 1.556$, $\text{coeff} = 0.241$, $z = 0.217$, $p = 0.266$, difference between means = 6%). These results provide evidence that the aforementioned gap regarding the level of altruism of subjects who decide to look at the payoffs is mainly driven by a gap in observers' beliefs about decision makers' level of altruism.

Finally, we investigate whether not looking at payoffs signals altruistic behavior. Similar analysis as before shows that the level of altruistic behavior among “non-lookers” was not significantly different from the level of altruistic behavior among subjects to whom payoffs were denied by default ($\chi^2(1,271) = 2.768$, $\text{coeff} = -0.003$, $z = 0.012$, $p = 0.990$, difference between means = 0.1%); moreover, the level of perceived altruistic behavior among non-lookers was not significantly different from the level of perceived altruism among subjects to whom the payoffs were denied by default ($\chi^2(1,204) = 2.951$, $\text{coeff} = 0.398$, $z = -0.798$, $p = 0.424$, difference between means = 9%). However, we mention that the proportion of people who decided not to look at payoffs was so small (around 10%), that it is possible that the lack of a significant effect is due to an undesired lack of statistical power.

To summarize, Experiment 1 provides evidence that subjects who look at the payoffs are perceived to be much more selfish than they actually are¹.

Experiment 2

Our first experiment suggests that subjects who look at payoffs are perceived much more selfish than they actually are. One potential explanation for this gap is that the act of looking generates a moral cleansing effect (Sachdeva, Iliev & Medin, 2009). Moral cleansing

¹All these results are robust after controlling for sex, age, and level of education.

theory posits that people have a positive moral conception of themselves and that they strive for balance in their moral acts to maintain this positive concept. In other words, when people do something that they think it is morally wrong, they need to subsequently do something that they think it is morally right to compensate for it (Brañas-Garza, Bucheli, Paz Espinosa & García-Muñoz, 2013; Bandura, 1991; Dunning, Fetchenhauer & Schlösser, 2012; Dunning, 2007). In this light, it is possible that, in our first experiment, some decision makers were aware of the fact that choosing to know every detail of the decision problem, if communicated to observers, would be perceived with distrust. Since this action then tips the moral balance towards bad behavior, these agents may have felt the need to compensate their behavior by helping at the next occasion. Although we tried to minimize moral cleansing effects by not communicating to decision makers that their decision whether to look at payoffs or not would be shown to an observer, it is possible that some decision makers experienced a feeling of deep-rooted, non-consequentialist guilt anyway, and were motivated to subsequently act altruistically.

To avoid this potentially confounding factor, we build on a previous work by Brañas-Garza et al. (2013). Specifically, in Experiment 2 decision makers play two games: they first play an envelope game and then they play a dictator game in the role of dictators. Observers' beliefs are elicited in the dictator game, conditional on whether decision makers decided to look at the cost of helping in the envelope game. Since Brañas-Garza et al. (2013) showed that in games involving altruism (iterated Dictator Game), moral cleansing does have an effect, but such an effect persists for only one round (i.e., the choice at round t is negatively correlated to the choice at round $t-1$), the idea is that any moral cleansing effect, if present in our experiment, should remain confined to the first envelope game and affect only the decision whether to help or not in the envelope game, and not the decision in the subsequent dictator game. Thus, comparing the average donation of dictators who opted for knowing the

cost of helping in the envelope game with the average donation of dictators who opted for not knowing the cost of helping, we can see whether looking at the cost of helping is associated with selfish behavior or not.

Method

A total of 240 brand new subjects (50% males, average age = 35.4 years), living in the US, were recruited on AMT and randomly divided in two conditions, the *active* condition and the *passive* condition.

Participants in the *active* condition played two games. In *Stage I*, they were given \$0.20 and they were told that they could “help” another participant, who was given nothing. Helping means transferring an amount $c > 0$ to a participant in the passive condition, who started with no money. Active participants were also informed that, before making their actual choice, they could find out the exact value of c . Thus, in total, active participants in Stage I made two decisions: first, they decided whether to find out the value of c ; and then, they decided whether to help the other participant. Hence, the Stage I game is strategically equivalent to the envelope game used in Experiment 1. As in Experiment 1 and as in Jordan et al. (2016), finding out the cost of helping was costless. Moreover, as in Experiment 1, participants were not informed that their decision about finding out the cost of helping would be communicated to an observer. In *Stage II*, active participants played a standard Dictator Game (Kahneman, Knetsch & Thaler, 1986) with another passive participant, different from the one they were paired with in Stage I. Precisely, subjects were given another \$0.20 and they were asked how much, if any, they wanted to give to the other player.

Apart from being receivers in both Game I and Game II, participants in the *passive* condition were also asked to guess the behavior of a randomly selected active participant in Game II, conditional on whether that participant chose to look at the cost of helping in Stage I

envelope game. Specifically, passive participants were shown all the screenshots given to active participants in Stage I and Stage II, and, after being informed about whether that active participant chose to look at the cost of helping in Stage 1 envelope game, they were asked to guess the active participant's donation in the Stage 2 dictator game. Correct guesses were incentivized with a \$0.20 reward. Importantly, subjects were informed only about whether active players decided to find out the cost of helping or not, and not about the subsequent decision whether to actually help or not.

All participants (both active and passive) were asked comprehension questions to test for their understanding of the decision problem. Participants failing any comprehension questions were automatically excluded from the survey.

Results and discussion

Figure 3 summarizes the results and provides visual evidence that, in line with Experiment 1, also in Experiment 2 “lookers” were perceived to be far more selfish than they actually were.

In order to formally support this, we begin by exploring whether active participants who looked at the cost of helping were actually more selfish than those who decided to make a choice without looking. To this end, we conduct linear regression predicting Stage II donations of active participants as a function of a dummy variable, which takes value 1 if the subject decided to find out the value of c in Stage 1, and value 0 otherwise. Results show that the average level of altruistic behavior among “lookers” was not statistically different than that among “non-lookers” ($F(1,126) = 0.2808$, $\text{coeff} = 0.2676$, $p = 0.5971$, $r^2 = 0.0022$, difference between means = 2.7%)

Next, we analyze whether looking at the cost of helping was perceived as a signal of selfishness. To this end, we compare the average guess by passive players paired with “active

lookers” with the average guess of passive players paired with “active non-lookers”. Linear regression shows that finding out the cost of helping is a strong signal of selfish behavior ($F(1,110) = 27.7942$, $\text{coeff} = -2.1607$, $p < .0001$, $r^2 = 0.2014$, difference between means = 21.6%)².

In sum, Study 2 finds that the perception gap found in Study 1 - that looking at the cost of helping is a signal of selfish behavior, but it does not actually predict selfish behavior - extends to a context in which moral cleansing is, at least partially, ruled out.

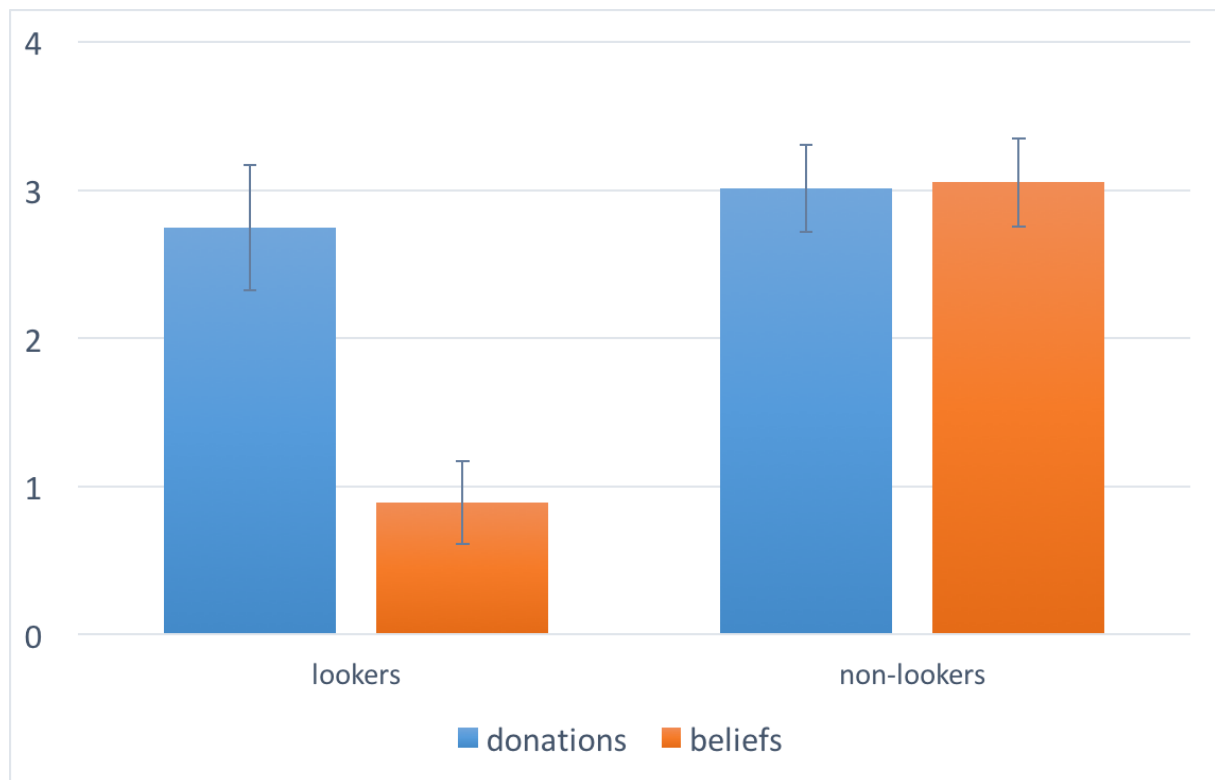


Figure 3. Average choice in dictator game (donation or guess) across conditions. Error bars represent the standard error of the mean. Average donation (normalized such that full donation is equal to 10) among people who decided to look at the cost of helping in Stage I was not different from average donation among people who decided not to look at the cost of helping. However, people who decided to look at the cost of helping were believed to be

²These results are robust after controlling for sex, age, and level of education.

significantly more selfish than those who decided not to look at the cost of helping. We refer to the Main Text for detailed statistical analysis.

General discussion

In daily life, people often do not know the exact costs and benefits of a pro-social action beforehand, but can gather this pieces of information only in a subsequent stage. Starting from this observation, recent studies have started exploring how observers interpret the decision maker's act of "looking at payoffs". Theoretical models propose that looking at costs and benefits will be met with distrust by observers (Hoffman, Yoeli & Nowak, 2015; Hilbe, Hoffman & Nowak, 2015), and, in agreement with this prediction, Jordan et al. (2016) showed that "lookers" are both less trustworthy and perceived to be less trustworthy than "non-lookers".

In this paper we have explored whether these results extend to altruistic behavior. In doing so, we uncovered a perception gap according to which (i) "lookers" are perceived to be *far* less altruistic than "non-lookers" (differences between means always above 20%), but (ii) "lookers" are not significantly less altruistic than "non-lookers".

Specifically, Experiment 1 shows that "lookers" in the envelope game are not significantly less altruistic than subjects to whom payoffs are given by default, suggesting that the mere act of looking is *not* associated with selfish behavior. However, observers do believe that looking at payoffs is associated with selfish behavior, that is, they believe that "lookers" help less than people to whom payoffs are given by default. The size of this perception gap turns out to be enormous, as the difference between percentage of actual helping and belief is around 23%.

One potential explanation for this gap is that looking at payoffs generates a moral cleansing effect (Sachdeva, Iliev & Medin, 2009), which is not anticipated by observers.

Perhaps, “lookers” are aware of the fact that choosing to know every detail of the decision problem will be met with distrust by observers. Since this action tips the moral balance towards bad behavior, some lookers may have felt the need to compensate their behavior by helping at the next occasion.

To avoid this potentially confounding factor, Experiment 2 implements a slightly different design in which the measure of altruism is not taken inside the envelope game, but through a subsequent dictator game. This experimental setup ensures that any moral cleansing effect, if existent, would remain confined to the envelope game (Brañas-Garza et al., 2013). The results of Experiment 2 confirmed those of Experiment 1: (i) “lookers” are perceived to be far less altruistic than “non-lookers”, but (ii) “lookers” are not significantly less altruistic than “non-lookers”. The size of the effect was also comparable to that in Experiment 1 (21%).

Since the measure of altruism in Study 2 was taken in a different task than the one in which the decision whether to look at the payoffs was made, another interesting consequence of Study 2 is that it seems to imply that the decision to look at payoffs is perceived by the other participant as being reflecting some sort of personality trait, not limited to the task in which the decision to look was made. This is in line with the research on the “cooperative phenotype”, suggesting that pro-social choices made in different situations are positively correlated (Capraro et al., 2014; Capraro, Smyth, Mylona & Niblo, 2014; Peysakhovich, Nowak & Rand, 2014; Biziou-van-Pol et al., 2015).

Our findings diverge from previous results by Jordan et al. (2016), who found that “lookers” were less trustworthy and perceived to be less trustworthy than “non-lookers” in a standard Trust Game. As discussed in the introduction, a potential explanation for this difference comes from the analysis of the theoretical reason which has been proposed as the ultimate explanation for why “non-lookers” should be more pro-social than “lookers”. Specifically, Hofmann et al. (2015) propose that non-looking at payoffs signals that decision

makers will be using heuristics internalized from everyday interactions, rather than reasoning about the particular payoff structure of the game. The crucial difference between trustworthiness and altruism is that, while being trustworthy is optimal in repeated interactions and thus it may be internalized as a useful heuristic (because most everyday interactions are with friends, family and coworkers, and thus they are repeated), altruistic behavior, being zero-sum, is not optimal in repeated interactions. Consequently, altruistic behavior should be independent of participants' cognitive reflection mode.

Although this point of view explains why "non-lookers" are more trustworthy than "lookers", but they are not more altruistic than "lookers", it does not straightforwardly explain why "non-lookers" are perceived to be more altruistic than "lookers". One possibility is that observers overgeneralize to the Dictator Game their correct experience about the Trust Game. Of course, this is only one potential explanation, and we hope that future research will help explore this and other potential frameworks.

Another interesting question for future research is understanding why people look at payoffs. Are they considering their own payoff, or the payoff of the other person, or perhaps a combination of both? Are lookers aware of the fact that their behavior can be perceived with mistrust by observers? If so, why do they look at the payoffs? We believe that exploring these questions can highly improve our understanding of human behavior in situations that are common in people's everyday lives.

To conclude, the two experiments presented in this paper uncover a perception gap according to which looking at payoffs signals selfish behavior but it does not actually mean so. Since interpreting others' actions in a correct way is crucial for healthy social relationships, we believe that understanding the origin of this gap and its potential consequences for human behavior is an important direction for future research.

References

- D'Adda, G., Capraro, V., & Tavoni, M. (2015). Push, don't nudge: Behavioral spillovers and policy instruments. Retrieved from SSRN:
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2675498.
- Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes*, 50, 248-287.
- Biziou-van-Pol, L., Haenen, J., Novaro, A., Occhipinti-Liberman, A., & Capraro, V. (2015). Does telling white lies signal pro-social preferences? *Judgment and Decision Making*, 10, 538-548.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90, 166-193.
- Brañas-Garza, P., Bucheli, M., Paz Espinosa, M., & García-Muñoz, T. (2013). Moral cleansing and moral licenses: experimental evidence. *Economics and Philosophy*, 29, 199-212.
- Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interactions*. Princeton: Princeton University Press.
- Capraro, V. (2013). A model of human cooperation in social dilemmas. *PLoS ONE*, 8(8), e72427.
- Capraro, V., & Cococcioni, G. (2015). Social setting, intuition, and experience in laboratory experiments interact to shape cooperative decision-making. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150237.
- Capraro, V., Jordan, J. J., & Rand, D. G. (2014). Heuristics guide the implementation of social preferences in one-shot Prisoner's Dilemma experiments. *Scientific Reports*, 4, 6790.

- Capraro, V., Smyth, C., Mylona, K., & Niblo, G. A. (2014). Benevolent characteristics promote cooperative behaviour among humans. *PLoS ONE*, *9*, e102881.
- Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*, 817-869.
- Dunning, D. (2007). Self-image motives and consumer behavior: How sacrosanct self-beliefs sway preferences in the marketplace. *Journal of Consumer Psychology*, *17*, 237-249.
- Dunning, D., Fetchenhauer, D., & Schlösser, T. M. (2012). Trust as a social and emotional act: Noneconomic considerations in trust behavior. *Journal of Economic Psychology*, *33*, 686-694.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*, 583-610.
- Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*, 817-868.
- Franzen, A., & Pointner, S. (2013). The external validity of giving in the dictator game. *Experimental Economics*, *16*, 155-169.
- Hilbe, C., Hoffman, M., & Nowak, M.A. (2015). Cooperate without looking in a non-repeated game. *Games*, *6*, 458-472.
- Hoffman, M., Yoeli, E., & Nowak, M. A. (2015). Cooperate without looking: Why we care what people think and not just what they do. *Proceedings of the National Academy of Sciences USA*, *112*, 1727-1732.
- Horton, J. J., Rand, D. G., & Zeckhauser, R. J. (2011). The online laboratory: Conducting experiments in a real labor market. *Experimental Economics*, *14*, 399-425.
- Jordan, J. J., Hoffman, M., Nowak, M. A., Rand, D. G. (2016). Uncalculating cooperation as a signal of trustworthiness. Available at SSRN: <http://ssrn.com/abstract=2725550>.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumption of economics. *Journal of Business*, *59*, S285-S300.

- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5, 411-419.
- Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Amazon Mechanical Turk as a participant pool. *Current Directions in Psychological Sciences*, 23, 184-188.
- Peysakhovich, A., Nowak, M. A., Rand, D. G. (2014). Humans display a “cooperative phenotype” that is domain general and temporally stable. *Nature Communications*, 5, 4939.
- Rand, D. G. (in press). Cooperation, fast and slow: Meta-analytic evidence for a theory of social heuristics and self-interested deliberation. *Psychological Science*.
- Rand, D. G., Greene, J. D., & Nowak, M. A. (2012). Spontaneous giving and calculated greed. *Nature*, 489, 427-430.
- Rand, D. G., Peysakhovich, A., Kraft-Todd, G. T., Newman, G. E., Wurzbacher, O., Greene, J. D., & Nowak, M. A. (2014). Social heuristics shape intuitive cooperation. *Nature Communications*, 5, 3677.
- Rand, D. G., Brescoll, V. L., Everett, J. A. C., Capraro, V., & Barcelo, H. (2016). Social heuristics and social roles: Intuition favors altruism for women but not for men. *Journal of Experimental Psychology General: General*, 145, 389-396.
- Rapoport, A. (1965). *Prisoner's Dilemma: A study in conflict and cooperation*. University of Michigan Press.
- Sachdeva, S., Iliev, R., & Medin, D. L. (2009). Sinning saints and saintly sinners: The paradox of moral self-regulation. *Psychological Science*, 20, 523-528.