Bracelets of Pride and Guilt? An Experimental Test of Self-Signaling in Charitable Giving

van der Weele, J.J.; von Siemens, F.A.

DOI
10.1016/j.jebo.2020.02.001

Publication date
2020

Document Version
Final published version

Published in
Journal of Economic Behavior & Organization

Journal of Economic Behavior and Organization 172 (2020) 280–291

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Bracelets of pride and guilt? An experimental test of self-signaling☆

Joël J. van der Weele[a,*], Ferdinand A. von Siemens[b,c]

[a] Faculty of Economics and Business, University of Amsterdam, Roeterstraat 11, Amsterdam 1018WB, the Netherlands
[b] Faculty of Economics and Business, Goethe University Frankfurt, Grüneburgplatz 1, Frankfurt am Main 60323, Germany
[c] CESifo, Poschingerstrasse 5, 81679 Munich, Germany

ARTICLE INFO

Article history:
Received 17 May 2019
Revised 22 January 2020
Accepted 4 February 2020

JEL classification:
C91
C72
D83

Keywords:
Self-signaling
Economic experiments
Cognitive dissonance

ABSTRACT

Self-signaling theory argues that behavior is important to build up or maintain a favorable self-image. We provide a novel test of this argument by manipulating the importance of behavior for future self-image. In two experiments, part of the subject pool is incentivized to wear bracelets as reminders of their initial identity-relevant behavior. We find some evidence that the bracelets increase anticipated memory, which should make behavior more relevant for managing a positive self-image. However, we find no evidence for self-signaling. Instead, our results suggest that participants resolve cognitive dissonance by constructing self-serving rationalizations of their actions that serve as cheap substitutes for self-signaling.

1. Introduction

Building and maintaining a favorable self-image is an important human motivation, see Baumeister (1998) and Fiske (2009). A good self-image may bring both affective and functional benefits, for instance by making it easier to convince others that one is a good interaction partner. How people form and manage their self-image has been the subject of much academic debate. One influential theory is that people construct their image by inferring personal characteristics from their own past behavior, much like an outsider would do, see Bem (1971). This idea seems paradoxical, because the need for inference implies people do not know the preferences driving their own behavior. Moreover, anticipating this, they may manipulate their behavior to "fool" their own self-inferences.

Economic models have illuminated this paradox by framing the decision environment as an intra-personal signaling game (Bodner and Prelec, 2003; Bénabou and Tirole, 2006; 2011). These models posit that at the moment of choice, a person's "true preferences" influence the decision. In the absence of choice, the person does not really know her preferences, and

☆ A previous version of the paper with only the first study was circulated as CESifo Working Paper Nr. 4674 under the title of "Bracelets of Pride and Guilt? An Experimental Test of Self-Signaling in Charitable Giving." We want to thank Luigi Butera, Dirk Engelmann, Holger Herz, Dan Houser, Michael Kosfeld, Devesh Rustagi, Jeroen van de Ven, Roberto Weber, and Christian Zehnder for very useful comments. We also thank Kevin Bauer, Bastian Bohrmann, Jakob Ehlich, Jakob Haerle, Karin Hettwer, and Petra Petersen for great research assistance. Financial support from the research fund administered by the Faculty of Economics and Business Administration at Goethe University Frankfurt is gratefully acknowledged.

* Corresponding author.
E-mail addresses: vdweele@uva.nl (J.J. van der Weele), vonsiemens@econ.uni-frankfurt.de (F.A. von Siemens).

https://doi.org/10.1016/j.jebo.2020.02.001
0167-2681/

she also cannot deduce them by simply imagining what she would do in a given situation. Instead, she makes rational (Bayesian) inferences on the basis of her own past actions. Because the person would like to believe that she has "virtuous" preferences, the result is a game of self-signaling, where both sender and receiver reside within the same person.

Testing self-signaling theory has proven complicated, because it is hard to experimentally vary the informational content that one's own choice conveys to oneself. It is also hard to manipulate the importance of self-image; one cannot separate a person from herself. We thus use a novel technique to manipulate the importance of behavior for self-image construction or maintenance. In two experiments, subjects might have to wear a bracelet for a duration of two weeks. The bracelet is introduced to subjects as a "personal reminder." The bracelet reminder makes it harder to forget bad behavior or poor performance, thus increasing the importance of what happens in the experiment for future self-image. According to the self-signaling model formulated in Bénabou and Tirole (2011), increasing the salience or weight of future self-image should lead to more signaling investments. Because the bracelets are plain and only the subjects know what they refer to, we can separate self-image from social-image concerns.

We use our bracelet design to test self-signaling theory in two separate experiments. In the first experiment, subjects face the moral task of dividing 20 euros between themselves and the German Red Cross. In the "bracelet treatment," subjects can earn an additional 10 euros after two weeks by coming back and presenting an intact bracelet. In the "control treatment," subjects also return after two weeks, but are not required to wear a bracelet. We hypothesize that the bracelet increases incentives to self-signal to future selves and therefore leads to higher donations. Our bracelet manipulation induces a statistically highly significant shift in reported anticipation of remembering the experiment, going from "From Time To Time" to "Often." However, we don't find any difference in giving between the bracelet and control group.

To better understand the origins of this null result, we conduct a second experiment. Male subjects do a physical fitness test – they participate in a push-up competition – and receive feedback on the ranking of their performance. The feedback is manipulated to come from comparison to either a strong or weak reference group, thereby creating exogenous variation in feedback that is independent from performance. We then elicit willingness to pay for not having to wear a reminder bracelet. We hypothesize that, conditional on performance, subjects with negative feedback will have a higher willingness to pay for not having to wear the bracelet. We find no effect of feedback on bracelet valuation. However, as a response to cognitive dissonance or a "sour grapes" effect, the negative feedback reduces the reported importance of the physical fitness test and its reported relation to self-image.

As we discuss in more detail at the end of the paper, our results provide several insights in the theory of self-signaling. While they affirm the importance of self-image management, they also point to a reason why self-signaling is a weak force in many situations: the ability to reinterpret their actions and downplay their diagnostic value for their underlying traits provides subjects with a cheap substitute for self-signaling. Self-signaling may thus affect behavior only in situations with little moral-wiggle room, or by driving people to seek out situations with such wiggle room. It's also possible that the benefits of self-image do not derive from anticipation of the future, but from an instantaneous warm glow. Determining the scope for self-signaling across different contexts is an important challenge for future research.

Our paper adds to the literature on identifying self-signaling. In an early contribution, Quattrone and Tversky (1984) show that people engage in self-signaling about their health status. Several more recent papers have looked at self-signaling in the context of ethical motivations. In Grossman (2015), subjects interact in dictator games where the probability that their decision is implemented varies exogenously. Self-signaling predicts that a lower implementation probability leads to higher donations. This prediction is not reflected in the data. Grossman and van der Weele (2017) show how self-signaling can explain willful ignorance in ethical dilemmas, as demonstrated by Dana et al. (2007) and others. They conduct a series of further experimental tests that support self-signaling. Amongst them is the finding that people are willing to pay for ignorance, and are more open to information after it has become irrelevant to their decisions. In Cueva and Dessi (2012), subjects observe public statements of donations of others, which they argue increase the salience of self-image concerns. They find that observing others raises donations for those with intermediate levels of "social potency." This finding is consistent with self-signaling theory.

Hershfield et al. (2012) document correlations between different ethical attitudes and behavior, and responses to a questionnaire designed to measure the continuity between current and future selves. Tonin and Vlassopoulos (2013) and Murnighan et al. (2001) show how dilution of self-image concerns leads people to behave more selfishly. Mazar et al. (2008) show that people behave more honestly and less selfishly when primed with the importance of moral rules. Shu et al. (2011) document that people misinterpret and forget information that might make them feel bad about their dishonest behavior. Gneezy et al. (2012) show that market participants are eager to profit from favorable deals when the price is given, but dislike to pay peanuts if they are responsible for choosing the price themselves. Most of these studies do not directly target the mechanisms of self-signaling, but rather try to exclude alternative explanations. The studies that target the mechanisms directly yield mixed results. Further direct empirical investigations of the importance of self-signaling for prosocial behavior are needed.

2. Study 1: charitable giving

In this study some participants were incentivized to wear a cloth bracelet. We hypothesize that this bracelet functions as a reminder of the experiment. It thereby increases the salience of future self-image and makes it harder to forget the actions taken in the experiment and the subjects' actions during the experiment. According to the model in Bénabou and Tirole (2011), such reminders should increase self-signaling investments.[1] The use of the bracelet is designed to manipulate two variables in the model. First, the bracelet should increase a parameter s that multiplies the self-image term in the utility function. This parameter measures "the strength of the self-esteem motive" (p.816). Furthermore, "[An] important determinant of s is salience – the extent to which the individual thinks (perhaps prompted by an experimenter or advertiser) about the contribution of A_2 to his future welfare, and how it depends on where his true values really lie" (p.817). Bénabou and Tirole (2011) refer to A_2 as the "relational capital" the individual can build up by behaving in the right way. Second, the bracelet should decrease a parameter λ, where $1-\lambda$ is the probability that future selves infer their true preferences from past actions. In the words of the authors, "$1-\lambda$ should be thought of as the malleability of beliefs through actions, and thus also reflecting the possibility that deeds may themselves be forgotten or repressed, or be uninformative due to situational factors that can be invoked as plausible excuses" (p.815). Thus, $1-\lambda$ increases when actions become more informative or are less likely to be forgotten, see also footnote 10 on p.815 of the article. Proposition 2 on p.823 states that both increases in s and in $1-\lambda$ lead to increased investment in self-signaling.

In designing the study, we put a lot of thought into selecting the best technology to increase the manipulate self-image concerns. We considered phone apps, incentivized periodic logins into a website, or email reminders. However, frequent reminders also emphasize the presence of the experimenter, especially when further data are elicited. Furthermore, email reminders involve anonymity issues, because subjects have to give out their emails, while data transfer via phone apps is privacy sensitive. These technologies therefore blur the boundary between self and social image concerns. By contrast, a bracelet is low-tech, transparent, private, and anonymous. Furthermore, bracelets are common for entrance control at music festivals and all-inclusive holiday resorts, and many people wear similar bracelets for extended periods of time as reminder of the festival or holidays. Thus the bracelets are unlikely to attract a lot of attention from outsiders, and participants may easily understand and anticipate that the bracelet works as reminder of the experiment.

Subjects were recruited from the student population of Goethe University Frankfurt using the online recruitment system ORSEE by Greiner (2015). The email invitations to the experiment announced that subjects might have to wear a cloth bracelet to earn money. We reminded subjects of this announcement before the experiment starts. We thus avoid that some subjects suddenly refuse to wear a bracelet during the sessions. In the end, nobody refused to put on the bracelet.[2]

Upon entering the lab, a randomly chosen orange or yellow cloth bracelet was attached to each participant's wrist by squeezing shut a metal ring with pincers.[3] Because the procedure of attaching the bracelet might affect behavior, we attached bracelets to all participants in both treatments. After the bracelet is attached to their wrists, subjects drew a seat number and an envelope marked "Red Cross." This envelope contained the instructions, a description of the activities of the German Red Cross in Syria, a short questionnaire, and overall 20 euros in one 10-euro bill, one 5-euro bill, and five 1-euro coins. Subjects took their seat at a private cubicle, where they found another envelope marked "For Me." Once all subjects were seated, the experimenter told them to open the envelope marked "Red Cross" and read the instructions and the information about the Red Cross carefully. All questionnaires and instructions can be found in the appendix.

The instructions explained that this is the first of two parts of the experiment. In this first part, the task consisted of dividing 20 euros between the subject and the Red Cross. Subjects took the amount that they want to keep for themselves out of the "Red Cross" envelope and put it into the "For Me" envelope. This feature of the design might emphasize the moral dimension of the donation decision, because taking money from the "Red Cross" envelope could feel almost like taking away money from the charity. In addition, they filled out a short questionnaire. Subjects were told that they can enter the second part of the experiment by reporting to the secretarial office between 14 to 18 days after the first part of the experiment. They were told that upon doing so, the secretary would check if they are eligible to participate in the experiment, whether they had to wear a bracelet (because they come from a corresponding session), ask them to fill out another short questionnaire, and give them an additional ten euros. Subjects are told that this procedure does not take more than two minutes.

[1] Bénabou and Tirole have invoked this framework in other papers and applications (Bénabou and Tirole, 2004; 2006) as well as a recent review paper, see Bénabou and Tirole (2016). All papers have received a large number of citations, and inspired both empirical and theoretical follow-up work, see our introduction and, for example, Grossman (2015) and Grossman and van der Weele (2017). The model shares most fundamental insights with the approach by Bodner and Prelec (2003) – another seminal paper in the literature in economics literature. It also differs in some aspects, like the interpretation of the dual selves assumption and the temporal dimension of decision making, which we discuss further in the conclusion. Note that our experiment corresponds to the model version without "reinvestment stage" discussed in Bénabou and Tirole (2011). This stage allows the authors to study the motivational role of beliefs under imperfect willpower, but is not necessary for self-signaling to matter in our setting.

[2] A working paper contains a more detailed description of the experimental design and results (van der Weele and von Siemens, 2014). For Study 1 we did not apply for IRB approval because at the time the joint ethics commission of the Faculty of Economics and Business Administration of Goethe University Frankfurt and the Gutenberg School of Management & Economics of the Faculty of Law, Management and Economics of Johannes Gutenberg University Mainz had not yet been founded. We applied for and received IRB approval from this institution for Study 2.

[3] We use cloth bracelets that are closed with a metal lock (a picture of the bracelets can be found in the appendix). These bracelets are attached to the wrist of subjects and thus work as constant reminders. The bracelets cannot be taken off without breaking the lock and damaging the bracelet irrevocably. In Germany, the colors yellow and orange have no obvious connotations.

Table 1
Mean donations.

	Control Treatment	Bracelet Treatment	*p*-value
Entire Sample (N = 122)	7.09 (5.95)	6.86 (5.87)	0.97
Women (N = 74)	6.63 (4.90)	7.23 (5.53)	0.64
Men (N = 47)	7.78 (7.33)	6.13 (5.56)	0.59
Indistinct Self-Image (N = 92)	7.71 (6.23)	7.87 (5.84)	0.68

Note: Mean donations for the entire sample and selected subsamples. Standard deviations in brackets. *p*-values from two-sided Mann-Whitney ranksum tests. Note that the number of observations for women and men do not add up, as one subject did not provide a gender identification.

Instructions were identical for both treatments with two exceptions. The first difference was one paragraph printed in bold which in the control treatment reads: "When exiting the room, we will cut off the bracelet from your wrist. Therefore, after the experiment is over, you do not continue to wear the bracelet you are wearing now." In the bracelet treatment this paragraph was replaced by "The bracelet you are wearing will serve as your private reminder of today's experiment. As we explain below, you can earn additional money by wearing the bracelet for two weeks." These paragraphs were read aloud by the experimenter in all but the first session to make sure that subjects understand the situation.[4] The second difference in the instructions was the exclusive announcement in the bracelet treatment that subjects must produce an unbroken bracelet to receive their additional 10 euros in the second part of the experiment.

After about 10 minutes of deliberation time, the experimenter gently urged subjects to put their donation, the questionnaire, and the instructions into the "Red Cross" envelope, and then seal the envelope. They should also mark time and place for the second part of the experiment. Subjects were then called forward one by one, deposited their donation envelope in a box marked "Donations," signed a confirmation that they have received money, and exited the room. In the control treatment, the bracelet was cut from their wrist before exiting. In the questionnaires in the first and the second part of the experiment (taking place two weeks later), subjects answered several questions related to their perception of the bracelet. All instructions and questionnaires can be found in the appendix.

We conducted 6 sessions, 3 for each treatment, in June, 2013. Sessions featured between 17 and 23 subjects. 122 subjects participated in the first part of the experiment, 58 in the control treatment and 64 in the bracelet treatment. Subjects were from all areas of studies with a large minority of 45% studying economics or business studies. 82% of the participants in the first part of the experiment also participated in the second part.

2.1. Results

Self-signaling theory predicts our bracelet treatment to increase donations via increasing anticipated memory of the first part of the experiment. To check whether our treatment manipulation works, we ask in the questionnaire in the first part of the experiment how often subjects expect to remember today's experiment in the coming two weeks. The answers are measured on a scale ranging from 1 ("Never") to 5 ("Very often"). We find a large and statistically highly significant difference in responses: the average score in the control treatment is 3.00 (corresponding to the answer "From time to time") while in the bracelet treatment it is 4.02 (corresponding to the answer "Often"). A Mann Whitney ranksum test yields a *p*-value of 0.00.[5]

Based on self-signaling theory, we expect donations to be higher in the bracelet treatment than in the control treatment, because donations in this treatment should have a stronger impact on future self-image. This prediction is not reflected in the data. We find that subjects donate on average 7.09 euros in the control treatment, and 6.86 euros in the bracelet treatment. Figure 1 shows the distribution of donations in both treatments. The distributions look similar, and an exact Kolmogorov–Smirnov test cannot reject the Null-hypothesis that the distributions of donations are the same (*p*-value of 1.00). When we look at the frequencies of positive donations, we find that 82% make a positive donation in the control treatment, compared to 83% in the bracelet treatment. Furthermore, 7% donate their entire endowment of 20 euro in the control treatment, compared to 9% in the bracelet treatment. Fisher exact tests cannot reject that these frequencies are equal across treatments (*p*-values of 0.82 and 0.75).

Consequently, we do not find that having to wear the bracelet influences donations on aggregate.[6] It is possible, however, that self-signaling might have stronger implications for certain subgroups of participants. For instance, Croson and Gneezy (2009) argue that concerning moral or social decisions, women might react more strongly to changes in the decision context than men. Our treatment manipulation – which arguably changes the moral context of the donation decision – might thus be particularly strong for women. In Table 1 we look at the treatment effect only considering women. A Mann-Whitney

[4] In Session 1 the paragraph emphasizing that the bracelet would be cut off was not read out loud. Two subjects were surprised that their bracelet was removed upon leaving the laboratory. All empirical results are robust to excluding Session 1.

[5] We round all numbers to two digits. We report *p*-value of less than 0.005 as *p*-values of 0.00. All *p*-values are based on two-sided statistical tests.

[6] This non-result also holds in OLS regressions controlling for various subject characteristics, as well as in IV regressions where we use the bracelet as an instrument for anticipated memory, see van der Weele and von Siemens (2014) for more details.

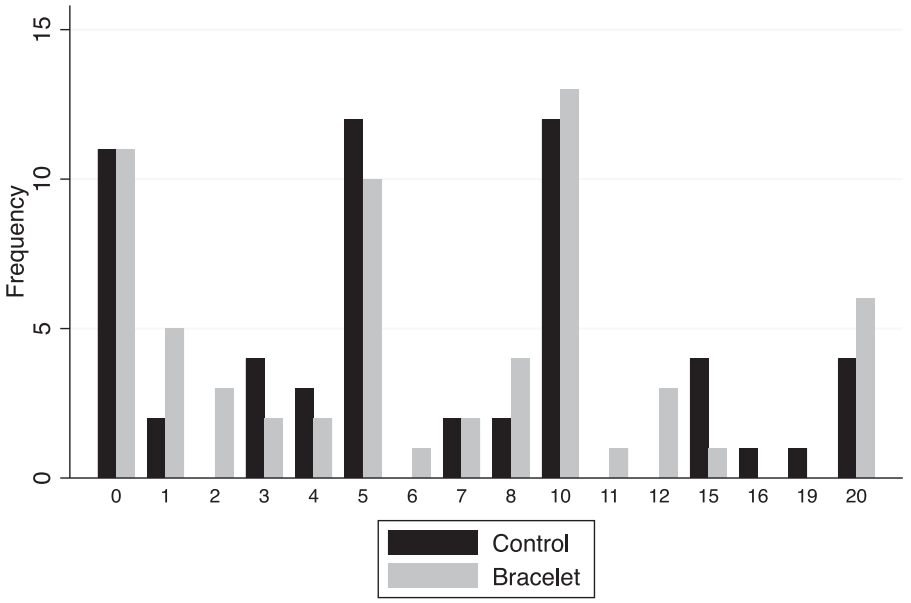

Fig. 1. Distribution of donations.

ranksum test reveals that there is no significant treatment difference (p-value of 0.64). The same holds for men (p-value of 0.59).[7]

Another subgroup that might be more strongly affect by self-signaling are subjects with an undecided self-image. Bayesian updating is stronger for less extreme priors. Self-signaling theory therefore predicts that subjects with an average self-image might have stronger incentives to signal their altruism to themselves (Bénabou and Tirole, 2011). Subjects report on how often they donate to charity as measured on a five-point scale ranging from 1 ("Never") to 5 ("Very Often"). Those who report the extremes "Never" or "Very Often" probably have a very decided self-image and are excluded In Table 1. We do not find a significant treatment effect for the remaining subjects (p-value of 0.68 from a Mann-Whitney ranksum test). The same is true if we only consider subjects who report that they donate to charity "From Time To Time" (p-value of 0.92 from a Mann-Whitney ranksum test). This contrasts with Cueva and Dessi (2012) who report that increasing the saliency of the moral dimension affects the donations of participants with an indistinct self-image as proxied by intermediate social potency.

2.2. Discussion

We do not find any significant treatment effect of having to wear a bracelet on donations, even when looking at subgroups of our participants where self-signaling might make the strongest predictions. We now discuss potential reasons for this null result. We believe we can convincingly rule out some of these concerns, but not others. The latter set of reasons motivate our second study.

First, we do not think that our null result is only due to a lack of statistical power. Figure 1 shows that the distributions of donations are very similar, while average giving is slightly *lower* in the bracelet treatment. Second, it seems unlikely that subjects who anticipate guilt from low donations plan to remove the bracelet after leaving the lab. If this plan translates into action, fewer subjects should return for the second part of the experiment in our bracelet than in our control treatment. This is not the case. We find that 81% of the participants in our control treatment and 83% of the participants in our bracelet treatment return for the second part of the experiment (p-value of 0.82 from a Fisher exact test).

Third, subjects might have strong preferences for wearing or not wearing the bracelet, which might reduce their donations or undermine the influence the bracelet might have on self-image. We can use our questionnaire data to investigate on subjects' perception of the bracelet. To assess whether subjects perceived the bracelet as a gift or a burden, we asked in the bracelet treatment in the questionnaire of the first part of Study 1 whether subjects would have preferred not to wear the bracelet. There is no clear direction to the response, as 17% of the respondents prefer not to wear the bracelet, 34% have no strong opinion, and 48% do not prefer not to wear the bracelet. Preferences for not wearing the bracelet are not significantly correlated with donations (Spearmean's rho of 0.18 with p-value of 0.14). We therefore cannot conclude that participants with the intention to donate little display a stronger dislike of wearing the bracelet. We also look at whether participants in

[7] To control for gender-specific treatment effects, we run control regressions that include the interaction of gender and treatment dummy. This does not affect our results.

our bracelet treatment return earlier or later for the second part of the Study 1 than participants in our control treatment. If the bracelet is a burden (pleasure) to wear, subjects should return earlier (later) in the bracelet treatment. A two-sided exact Kolmogorov-Smirnov test cannot reject that the distribution of return days is the same across treatments (p-value of 0.88). We therefore conclude that participants in our experiment do not, on average, derive (dis)utility from the bracelet.

We now turn to reasons for our null result that we cannot completely rule out and which thus form the motivation for our second study. First, it might be possible that the bracelet reminds subjects of their participation in the experiment, but not of their donation. There are good arguments against this possibility. The donation is the only decision that participants take in the experiment, so they should also remember their action if they think of the experiment. Subjects having to wear the bracelet report in the second questionnaire that they were more likely to think back of the experiment than those not having to wear the bracelet (p-value of 0.00 from a Mann-Whitney ranksum test). And once asked, all subjects report that they remember their donation, although the memory report is not incentivized. Nevertheless, we think we must better account for the possibility that participants (strategically) forget their own behavior. In our follow-up study, we thus measure memory of past behavior in an incentivized manner, and we use color-coding to tie the bracelet more closely to past behavior.

Second, the donation decision might not be very important for subjects' long-term self-image. This may be the case because the experiment involves only 20 euros, subjects already have had ample experience with their own moral behavior so that there is no point in signaling, and people think that they can make up for uncharitable behavior today by behaving more generously at the next occasion.[8] It could also be that the absolute donation is not so informative for self-image, which may be more sensitive to social comparisons and thus to donations relative to other subjects (Festinger, 1954). In our follow-up experiment, we therefore set up a situation in which subjects receive identity-relevant information that they cannot easily get otherwise or make up for later. We also generate a social comparisons by providing relative performance information.

Third, participants who are inclined to donate rather little may develop exculpatory rationalizations, for example that behavior in the artificial laboratory environment has little external validity and consequently does not say much about the true generosity of a participant in real life. Early research on such cognitive dissonance shows how people change their view on the nature of an experimental context in a self-serving way (Festinger and Carlsmith, 1959). If exculpatory rationalizations are easily constructed and convincing enough, the discomfort of being reminded of low donations would be much reduced, and so would the treatment effect. While we cannot prevent subjects from internally acquitting themselves, we include measures to detect this phenomenon in our follow-up study.

3. Study 2: physical exercise

As the discussion above makes clear, we believe there are three main reasons why the increased anticipated memory did not affect charitable giving in Study 1. First, the bracelet may have reminded subjects of the experiment, but not of their behavior in the experiment. Second, the giving task may not have been informative enough to engage in self-signaling, either because subjects regularly encounter giving situations, or because the experiment lacks social comparisons. Third, subjects might have been able to easily construct rationalizations and exculpations for non-giving.

Our objective in this second study is to design a new test that addresses these potential issues. First, we invite only males to take part in a performance task that is specifically designed to be an ego-threat for men: a test of physical strength. To provide social comparisons, we introduce (exogenously manipulated) feedback in the form of explicit relative rankings. To tie the reminder more closely to performance, we use color-coded bracelets to ensure memory of task performance. We measure memory in an incentive compatible manner. Finally, we measure the subjective perception of the importance of task with questionnaire items, in order to detect subjects' rationalizations as responses to cognitive dissonance along this dimension.

In addition, we exogenously vary the self-image threat and measure how this influences the willingness to pay for (not) wearing the bracelet. Thus, we can directly measure the value of the bracelet as a reminder, conditional on whether the reminder has a positive or negative effect on self-image. These multiple changes imply that we cannot directly compare the two experiments. This is a conscious choice. Our aim is not to identify individual reasons for the null result in Study 1, but rather to construct a more powerful test of self-signaling theory.

3.1. Design

Subjects were recruited from the student population of Goethe University Frankfurt using the online recruitment system ORSEE by Greiner (2015). In the invitation email, we did not mention the cloth bracelet, because subjects can avoid having to wear the bracelet. We announced that the experiment involves a push-up contest in which subjects must do 30 push-ups as fast as possible. We also told subjects that they should register for the experiment only if they are able and willing

[8] It is not clear that the theory supports such objections. Bodner and Prelec (2003) argue that every opportunity to behave pro-socially matters, because a favorable self-image requires permanent maintenance. Bénabou and Tirole (2011) suggest that not behaving pro-socially today can make it more difficult to behave pro-socially tomorrow, because a decreased stock in self-image capital reduces incentives to uphold favorable beliefs. Not behaving pro-socially just once thus puts individuals on a "slippery slope."

to do the 30 push-ups. We reminded subjects of the push-ups before the experiment really starts. If they decided not to participate, they are paid a show-up fee and leave the experiment. We thereby avoided that participants refuse to do the task during the experiment.[9]

Upon arrival at the laboratory subjects were led to another room where they signed an informed consent form. They then received instructions that were read out loud by an experimenter. All questionnaires and instructions can be found in the appendix. The instructions explain that this is the first of two parts of the experiment. In this first part of the experiment, the subjects started by participating in a push-up competition. For this they had to complete 30 push-ups as fast as possible. A push-up was correctly performed if the arms are first fully stretched, the straight body is lowered until the breast touches a soft massage ball of 7 cm in diameter, and the arms are fully stretched again. Subjects were allowed to pause if they cannot perform the 30 push-ups in one go. An experimenter takes the time, such that the subject could not see how long it took him to complete the task.[10] Subjects then estimated how long it took them to perform the push-ups.

Subjects were told that after they perform the push-ups, they would be compared to nine other male students of Goethe university who in previous sessions of the experiment also had completed 30 push-ups as fast as possible. Subjects guessed their rank as compared to the reference group of nine other students. To avoid any deception, subjects are informed that there were two different reference groups, and that they were randomly assigned "for experimental reasons."

We conducted 4 session with between 2 and 6 subjects to generate these reference groups, in April, 2018. Overall 18 subjects participated in these reference group sessions, and they earned 10 euros for their participation, independently from the time they took to complete the 30 push-ups. The strong reference group contained the nine best subjects, and the weak reference group the remaining nine worst subjects from the reference group sessions.

Subjects in the main experiment were randomly assigned to one of two treatments, which vary in the reference group used to determine the feedback. Subjects in the "Weak Reference Group" treatment were compared to the weak reference group, subjects in the "Strong Reference Group" treatment were compared to the strong reference group. The assignment to treatments was randomized at the individual level within sessions. We thus create random variation in feedback, allowing us to estimate the causal impact of feedback while controlling for performance.[11]

Subjects received feed-back on their performance in the push-up competition in the form of their ranking in the reference group, from 1 to 10. To further reinforce identity relevance, we provided a normative framing of their performance. We categorized a subject's performance as "good" if it was ranked among the best three, "average" if it was among the middle four, or "bad" if it was among the worst three subjects in the reference group. After receiving the feed-back, subjects themselves had to write down their rank and tick off the corresponding evaluation of their performance to make sure that the information really sinks in.

Participants knew before the push-up competition that they might have to wear a bracelet to be eligible for participation in the second part of the experiment. We stated explicitly that the bracelet is to remind them of their relative performance in the push-up competition. To further load the bracelet with this information, we told subjects before the push-ups to privately assign colors for the cases that their performance is good, average, or bad. We asked subjects to keep this color assignment secret, so that only they know which color corresponds to which performance. We thus cannot check whether subjects in the end select the color that corresponds to their performance, yet we can exclude social image concerns as the colors have no common meaning.

After providing feed-back, we elicited subjects' willingness to pay for not having to wear the bracelet using a non-linear Becker-DeGroot-Marschak mechanism with six different price levels (0, 10, 20, 50, 100 or 200 eurocents). The mechanism is non-linear to allow for finer measurement of small willingness-to-pay. We kept the number of total prices equal to six so that subjects themselves could determine their price with a dice roll. If and only if the dice indicated a higher price as compared to the subjects choice, the experimenter asked the subject to select the bracelet of the color corresponding to their feedback category (good, average, bad). The experimenter then attached this bracelet to the subject's wrist. Finally, subjects filled out a short questionnaire, were paid, confirmed payment, and exited the room.

We tried very hard to ensure that subjects understand all elements of the first part of the experiment. After reading out the instructions, one experimenter thus showed how a correct push-up must be performed, and each subject must do one practice push-up in the group to ensure comprehension. We also explained with two examples how the Becker-DeGroot-Marschak mechanism works. We then played through the mechanism with each subject, where the price is set to 50 eurocents, and answered any remaining questions. Subjects were then led to a waiting space. An experimenter accompanied them there to make sure that subjects do not talk about the details of experiment. Subjects were then called one by one to the room where they completed the first part of the experiment. We thereby avoided spill-over effects.

Like in Study 1, subjects could participate in the second part of the experiment by coming to the secretarial office between 14 to 18 days after the first part of the experiment. Subjects first identified themselves with an individual subject code that they determined themselves in the first part of the experiment. The codes consisted of four letters constructed

[9] Two subjects were surprised that they must do push-ups, and decided not to participate.

[10] One subject failed to complete the 30 push-ups within four minutes. This subject is included in the analysis, with his "completion time" set to 240 seconds.

[11] Subjects were informed that they would be compared to one of two reference groups, one of which is randomly chosen with equal probability. In line with standard experimental practice, we omitted information on the exact constitution of the feedback groups. We did this to avoid experimenter demand effects that might be induced by revealing the treatment variation.

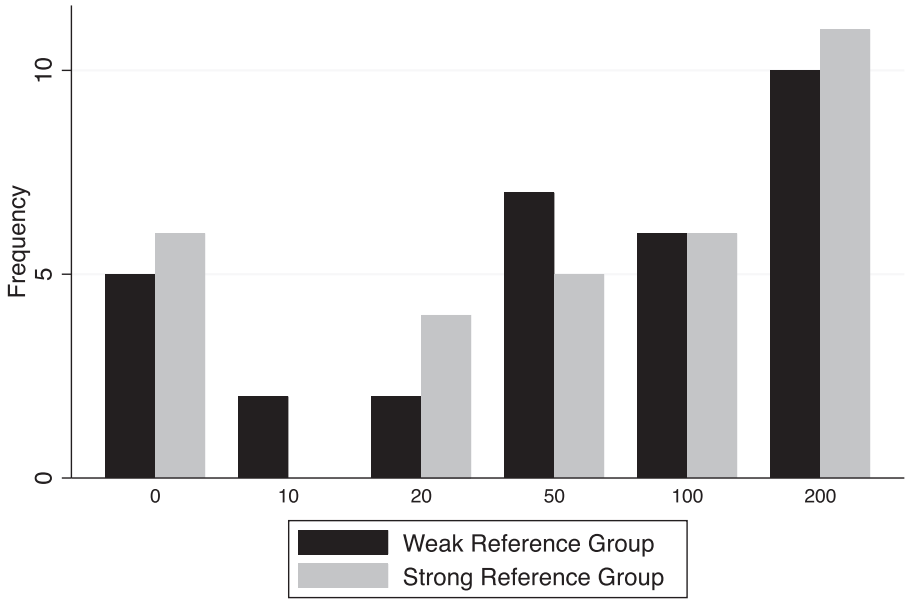

Fig. 2. Distribution of willingness to pay not to wear the bracelet.

from the first names and birth months of the respective subjects' parents. The procedure ensures that participants need not remember, but can easily re-construct their code for the second part of the experiment. The secretary checks the individual codes to see whether subjects are entitled to participate and whether they have to wear a bracelet for participation. Like in the first study, for those wearing the bracelet, earning the 20 euros was conditional on wearing an intact bracelet. Subjects could earn an additional 5 euros from correctly remembering their rank in the push-up competition two weeks ago. They then filled out a short questionnaire, were paid, confirmed payment, and exited the room. This concluded the experiment.

We conducted 18 sessions in April and May, 2018, with between 1 and 7 subjects for the main experiment. Overall 64 subjects participated in the first part of the main experiment, with an equal number in each of the two treatments.[12] Subjects in the main experiment were from all areas of studies, with a majority of 53% studying economics or business studies. 91% of the participants in the first part of the experiment also participated in the second part. Subjects in the first part of the main experiment received 10 euros for participation in the first part of the experiment. They could spend some of this money to avoid having to wear the bracelet. Subjects in the second part of the main experiment received 20 euros for participation. They could earn an additional 5 euros from correctly remembering their rank in the first part of the experiment. All other questionnaire questions were not incentivized.

The design ensures that subjects who are compared to the weak reference group receive on average better feedback than those with the strong reference group. Thus, if they use memories of their past performance in their construction of self-image, as self-signaling theory predicts, they should be more wiling to wear the bracelet. Our main hypothesis therefore is that subjects in the Weak Reference Group treatment display a lower willingness to pay not to wear the bracelet.

3.2. Results

Our randomization appears to have been successful; there are no significant treatment differences in field of study or experience with similar bracelets. An important feature of our design is that the push-ups are performed before the treatment is assigned, and should therefore be independent of it. There is no significant treatment difference in performance, as measured by the time it takes subjects to perform the 30 push-ups (p-value of 0.89 from a Mann-Whitney ranksum test). Our treatment manipulation generates the intended difference in feedback. While those in the Strong Reference Group treatment have an average rank of 8.1, the average rank in the Weak Reference Group is 2.9 (p-value of 0.00 from a Mann-Whitney ranksum test).

Our main outcome variable is the willingness to pay for not having to wear the bracelet. A first indication that there is no effect of the self-image threat on willingness to pay is that there is no significant correlation between performance as measured by the time taken to do the push-ups and willingness to pay (correlation coefficient of 0.04 with p-value of 0.76). The causal analysis involves the effect of the treatment manipulation. Figure 2 displays the distributions of willingness to pay in both feedback treatments. The histogram shows that there are no clear differences between the treatments. The

[12] One female subject participated because her gender was wrongly registered in our subject data base. We exclude this female participant from the data analysis.

Table 2
General motivation in competitions.

	Weak Reference Group	Strong Reference Group
Simply always want to win	0.22	0.00
Want to be proud of myself	0.66	0.74
Want to be admired by others	0.09	0.03
Want to win the prize	0.00	0.16
Am not very motivated	0.03	0.06

Note: Fraction of subjects per treatment who report the respective answer to be their most important motivation in competitive situations.

means are very close (94.1 for the weak reference group and 97.8 for the strong reference group). A Mann-Whitney ranksum test finds no significant differences in willingness to pay (p-value of 0.91). This result also holds in an OLS regression where we control for task performance, and an IV specification where we use the treatment as an instrument for the feedback received, again controlling for task performance.

After the uncertainty in the BDM mechanism is resolved, 17 out of 64 participants end up having to wear the bracelet. As a check on our treatment manipulation, we look at whether these bracelet wearers anticipate remembering the experiment more than the non-wearers. We find that this is the case: their average answer is 3.06 on a five-point scale (where 1 means "Never" and 5 means "Constantly") as compared to 2.77 for the non-wearers. However, a Mann-Whitney ranksum test reveals that this difference is not quite statistically significant (p-value of 0.16). Thus, the evidence that the treatment works as intended is less convincing than in our first study.[13]

Following our discussion of Study 1, we next test whether our non-results may be driven by subjects' rationalizations or cognitive dissonance. One possible rationalization is to attach less importance to the push-up competition upon receiving bad feedback. The randomized feedback in our experimental design means we can identify this causal link. In line with this idea, we find that people with the weak reference group are more likely to report that performing well was important for them. We measure the importance to perform well in the push-up competition on a five-point scale (where 1 means "Very important" and 5 means "Very unimportant"). The mean score is 2.53 for those with the weak reference group vs. 3.28 for those with the strong reference group (p-value of 0.00 from a Mann-Whitney ranksum test).

We also find that feedback affects the answer for the more abstract question of what motivates subjects most in competitions. Subjects choose one of five possible answers. Table 2 summarizes the fractions of subjects in each treatment. In both treatments, a large majority of subjects reports that their most important motivation in competitions is that they themselves want to be proud in their own performance. The fractions are 66% with the weak reference group and 74% with the strong reference group. This indicates that performance in a competitive situation as in the experiment has the potential to affect self-image. Only 3% and 6% of subjects report that they are in general not motivated in competitions. This coincides with the casual observation of the experimenters that subjects really did their best to perform well in the push-up competition.

We find that the treatment changed reported motivations on the task (p-value of 0.01 from a Pearson chi-squared test). 22% of subjects in the weak reference group report that they "simply always want to win." Not a single subject reports this as main motivation in the strong reference group, perhaps because it would have implied that they put in high effort, which makes their poor results more threatening to self-image. By contrast, no single subject with the weak reference group reports to be primarily motivated by the prize, as compared to 16% of the subjects with the strong reference group. Because there was no prize for winning in the current competition, this answer can provide a rationalization for the relative performance in the current competition in this group, by implying that they put in low effort. Thus, Table 2 provides a first hint of self-image management.[14]

We now turn to the second part of the experiment. Attrition in the second part of the experiment is very low, with only 6 participants dropping out, all of whom were among the non-bracelet wearers. Dropout rates were 88% in the weak reference group and 94% in the strong reference group (p-value of 0.67 from a Fisher exact test). There is no correlation between dropout rates and willingness to pay. This shows that subjects are not willing to pay high amount to avoid wearing the bracelet because they expect to remove the bracelet and not return to the second part. With the exception of two subjects, all subjects correctly remember their rank feedback from the first part of experiment. Subjects also report how often they remembered the push-up competition, with answers ranging from 1 ("Never") to 5 ("Constantly"). The average answer is 2.41 for those not having to wear the bracelet, and 3.12 for those having to wear the bracelet (p-value of 0.00 from a Mann-Whitney ranksum test). This mirrors our result in Study 1 and shows that the bracelet works as an effective reminder. We also ask subjects how much they enjoyed remembering their performance in the push-up competition, on a scale from 1 ("Very much") to 5 ("Very little"). Those with the weak reference group report on average 2.21, and those

[13] However, note that unlike in Study 1, those who wear the bracelet in Study 2 are self-selected. It is possible that subjects who would have higher levels of anticipation are willing to pay more not to wear the bracelet, and thus less likely to end up having to wear it.

[14] We originally included the general motivation measure to see for which subjects performing well in the competition is important for their self-image. Our treatment manipulation should be strongest for those who report self-image to be their most important motivation. But even when restricting the analysis to this subsample, we do not find that our treatment affects willingness to pay (p-value of 0.63 from a Mann-Whitney ranksum test). This test, however, might be biased because the treatment affects reported motivation.

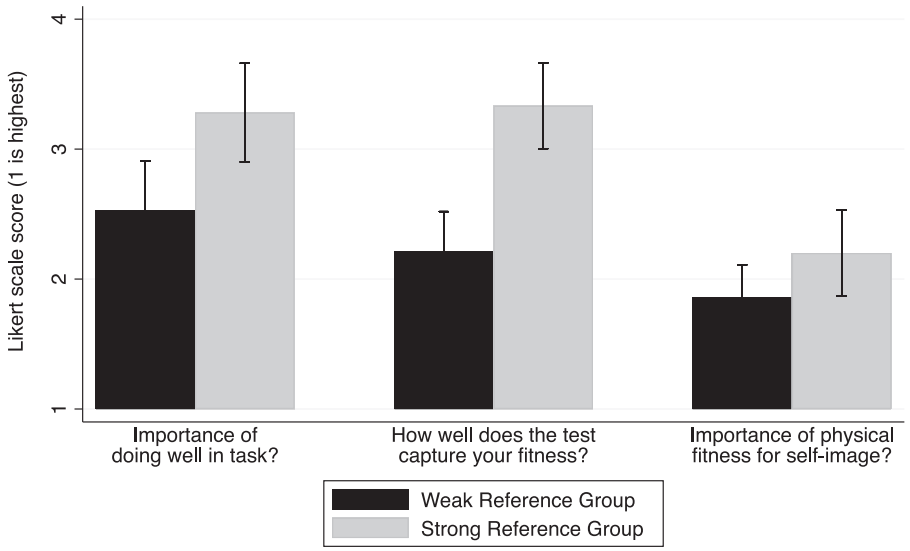

Fig. 3. Responses to Cognitive Dissonance. *Note:* Perception of experimental task as measured by three different questionnaire items. Black bars indicate 95% confidence intervals based on a Students t-distribution.

with a strong reference group report on average 3.4. The difference is statistically significant (p-value of 0.00 from a Mann-Whitney ranksum test). Bad feedback therefore creates negative emotions even in the long run, although we cannot say anything about the channel through which emotions are affected.

To further investigate cognitive dissonance, we also ask subjects whether they thought the push-up exercises adequately captured their physical fitness, on a scale from 1 ("very adequately") to 5 ("very inadequately"). Consistent with our results from the first part of the experiment, we find a treatment effect, with an average score of 2.21 for those with the weak reference group vs. a score of 3.33 for those with the strong reference group (p-value of 0.00 from a Mann-Whitney ranksum test). We find a similar but weaker effect for the question "how important is physical fitness for your self-image" as answered on a 5-point scale from 1 ("Very important") to 5 ("Very unimportant"). The average score is 1.86 for those with the weak reference group, and 2.22 for those with the strong reference group, but the difference is not quite significant (p-value of 0.14 from a Mann-Whitney ranksum test).

Fig. 3 summarizes our results on cognitive dissonance. Given the random variation in feedback in our design, we can rule out omitted variable bias and identify a causal relationship. All results are robust in a regression where we control for performance. The same is true in an IV specification where we instrument feedback with the treatment, again controlling for performance. We find a very systematic and long-lasting response to the cognitive dissonance caused by our negative feedback. Getting bad feedback results in downplaying the importance of the competition, its relevance to one's self-image, and even changes the reported competitive aspects of one's own character.

3.3. Discussion

Study 1 and 2 differ in several dimensions and therefore are not directly comparable. This was a conscious design choice. Nevertheless, we believe that the results of Study 2 are informative for the interpretation of Study 1. For instance, the use of the novel task and color codings does not result in any effect of the bracelet. This suggests that the task and the connection between choice and bracelet are not the main issues responsible for the null-result in Study 1.

By contrast, we do find evidence for our third conjecture, that participants protect their self-image by adopting self-serving perceptions of the task. This effect suggests that the task is relevant to subjects' self-image, and helps explain why we nevertheless do not find any effect on bracelet valuation. It is striking that cognitive dissonance does not only affect the interpretation of the test itself – importance and suitability. It also spills over and influences interpretations about one's own general drivers – motivation in competitive situations, importance of physical fitness – even if the latter effect is weaker. This finding is testimony to the power of cognitive dissonance.

Of course, we can only speculate on the importance of this finding for the finding of Study 1. While we have no means for verification, it seems plausible that making a low donation can be similarly exculpated. Subjects in the bracelet condition probably are able to come up with rationalizations to justify non-giving, thus contributing to the null result.

Finally, the design of our second study is related to Loewenstein and Issacharoff (1994) who compare valuations of people who obtained an object by doing well on a task with the valuations of others. We do not replicate their result that subjects attach more value to an object (in our case a bracelet) if it is associated with a more positive performance. It is hard to

pinpoint the exact reason for this non-replication, because there are many differences between the studies. Note, however, that our study provides a randomization of the feedback, thus controlling for omitted variables.

4. Conclusion

In our experiments, we attempt to test self-signaling theory via a novel manipulation of self-image concerns. We induce participants to wear a bracelet as reminder of their actions or performance in the laboratory. Although reminders increase anticipated and actual memory of the experiment, we do not see an increase in donations (Study 1). We also do not find a change in willingness to pay for not having to wear the bracelet depending on performance feedback (Study 2). As possible explanation for our null results, we find causal evidence that subjects strategically modify the perception of the experimental task. They do so in line with a "sour grapes" motivation, whereby worse performance leads them to downplay the importance of the task, or even to modify their own (reported) preferences.[15]

These results imply several insights for the theory of self-signaling. First is a hitherto unexplored reason why self-signaling is likely to be a much weaker force than social signaling. Participants seemingly use the strategic malleability of task perception to erode the diagnostic value of actions for underlying traits. This cognitive strategy, if anticipated, provides subjects with a cheap substitute for self-signaling. The availability of this substitute is likely to weaken incentives for self-signaling in many situations, as the exploitation of situational ambiguity appears to be widespread, see for example Gino et al. (2016), Exley and Kessler (2018), Gneezy et al. (2016), Schwardmann and van der Weele (2019). By contrast, when signaling to others, the decision maker cannot rely on such a lenient interpretation of her actions, and faces higher incentives for good behavior.

While the sour-grape phenomenon can explain the absence of self-signaling, it also confirms the relevance of self-image management more generally. Thus, our results do not rule out that self-signaling occurs in environments with little ambiguity or room for self-serving interpretations. Self-signaling may also affect behavior by motivating the deliberate *creation* of wiggle room or the sorting into environments with high ambiguity, as demonstrated in Grossman and van der Weele (2017). Determining the scope of self-signaling in the presence of cognitive escape routes is an important challenge for future theoretical and empirical research.

A second insight concerns the temporal aspect of self-signaling. Bodner and Prelec (2003) conceive of self-signaling theory as one where information about the self is simultaneously known and not known by different parts of an individual. In contrast, Bénabou and Tirole (2006, 2011) stress an inter-temporal interpretation that involves the current self signaling to the future self. People are aware of their true underlying preferences the moment they act, but insight is fleeting, so that future selves must infer their true preferences by looking at their own past behavior. The theory does not specify the length of timespan between action and introspection, which may be arbitrarily short (Bénabou and Tirole 2011, p.820). In this case, the instantaneous and inter-temporal interpretation coincide.

As clarifying example, imagine somebody sees a beggar in the street. This person might donate something to feel the immediate affective benefit associated with the reassurance of being an altruistic person. Alternatively, she may do so in order to feel good about herself over the next weeks. As the example makes clear, the two interpretations are not mutually exclusive. Our bracelet design only tests the latter, intertemporal interpretation. The instantaneous interpretation of self-signaling implies our bracelet manipulation should have no effect in our experiment, which only facilitates memory in the future. Our null result thus point in the direction of an instantaneous interpretation of self-signaling theory. However, the temporal aspect of self-signaling, as well as the delineation of "instantaneous" and "past" have not received much attention in the literature, which is a subject for future research.

Relatedly, future work should investigate mechanisms of anticipation and recall. Our Study 1 showed that it is possible to manipulate (anticipated) memory and increase the salience and recall of current actions as parameterized in the Bénabou-Tirole model. However, we do not know whether people correctly anticipated the emotions associated with the recall of their actions, another (implicit) assumption of the self-signaling model.[16] In addition, there is some evidence that memory may itself be distorted to make past behavior seem more favorable, see Saucet and Villeval (2019) and Zimmerman (2020), which may further depress the need for self-signaling.

Finally, from a methodological point of view, the sour-grape effect in Study 2 and the questionnaire answers on anticipated memory show that the bracelet does serve to emphasize self-image concerns. We also find ex-post that bracelets do in fact work as reminders and that people do not object to wearing the bracelet for several weeks for modest rewards. This may inspire their use in future work, including field settings. For instance, independent work has used bracelets to manipulate social image concerns, for example Karing (2018). In doing so, the bracelet may be used to advertise decisions taken inside the laboratory to audiences outside.

[15] Alladi (2018) also documents a similar "sour grapes" phenomenon in his field experiment. He finds that people attach less value to alternatives that they have a low probability of accessing.

[16] We considered asking participants for their anticipated emotions, but decided against it, as it may induce the very anticipation of these emotions. Our results would then be driven by experimenter demand effect, and hence be an artifact of the design.

Declaration of Competing Interest

Both authors declare that they have no relevant financial or material interests that relate to the research described in the paper "Bracelets of Pride and Guilt? An Experimental Test of Self-Signaling".

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2020.02.001.

References

Alladi, V., 2018. Sour Grapes in the Lab and Field: A Test of Access-Based Beliefs. Mimeo, UC San Diego, pp. 1–51.
Baumeister, R., 1998. The Self. In: Gilbert, D., Fiske, S., Lindzsey, G. (Eds.), The Handbook of Social Psychology. McGraw-Hill, New York, pp. 680–740.
Bem, D.J., 1971. Self-perception theory. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology. McGraw-Hill, New York, pp. 1–62.
Bénabou, R., Tirole, J., 2004. Willpower and personal rules. J. Polit. Econ. 112 (4), 848–887.
Bénabou, R., Tirole, J., 2006. Incentives and prosocial behavior. Am. Econ. Rev. 96 (5), 1652–1678.
Bénabou, R., Tirole, J., 2011. Identity, morals, and taboos: beliefs as assets. Q. J. Econ. 126 (2), 805–855.
Bénabou, R., Tirole, J., 2016. Mindful economics: the production, consumption, and value of beliefs. J. Econ. Perspect. 30 (3), 141–164.
Bodner, R., Prelec, D., 2003. Self-signaling and diagnostic utility in everyday decision making. In: Brocas, I., Carillo, J. (Eds.), The Psychology of Economic Decisions. Vol. 1. Rationality and Well-being. Oxford University Press, pp. 105–126.
Croson, R., Gneezy, U., 2009. Gender differences in preferences. J. Econ. Perspect. 47 (2), 448–474.
Cueva, C., Dessi, R., 2012. Charitable Giving, Self-image and Personality. IDEI Working Paper No. 748.
Dana, J., Weber, R.A., Kuang, J.X., 2007. Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. Econ. Theory 33 (1), 67–80.
Exley, C.L., Kessler, J.B., 2018. Motivated Errors. Mimeo, Harvard Business School, pp. 1–72.
Festinger, L., 1954. A theory of social comparison processes. Hum. Relat. 7, 117–140.
Festinger, L., Carlsmith, J.M., 1959. Cognitive consequences of forced compliance. J. Abnorm. Soc. Psychol. 58 (2), 203–210.
Fiske, S.T., 2009. Social Beings: Core Motives in Social Psychology. John Wiley and Sons, Ltd..
Gino, F., Norton, M.I., Weber, R.A., 2016. Motivated Bayesians: feeling moral while acting egoistically. J. Econ. Perspect. 30 (3), 189–212.
Gneezy, A., Gneezy, U., Riener, G., Nelson, L.D., 2012. Pay-what-you-want, identity, and self-signaling in markets. Proc. Natl. Acad. Sci. 109 (19), 7236–7240.
Gneezy, U., Saccardo, S., Serra Garcia, M., van Veldhuizen, R., 2016. Bribing the Self. Mimeo, UC San Diego.
Greiner, B., 2015. Subject pool recruitment procedures: Organizing experiments with ORSEE. J. Econ. Sci. Assoc. 1 (1), 114–125.
Grossman, Z., 2015. Self-signaling versus social signaling in giving. J. Econ. Behav. Organ. 117, 26–39.
Grossman, Z., van der Weele, J., 2017. Self-image and willful ignorance in social decisions. J. Eur. Econ. Assoc. 15 (1), 173–217.
Hershfield, H.E., Cohen, T.R., Thompson, L., 2012. Short horizons and tempting situations: Lack of continuity to our future selves leads to unethical decision making and behavior. Organ. Behav. Hum. Decis. Process. 117, 298–310.
Karing, A., 2018. Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone. Mimeo, UC Berkeley, pp. 1–75.
Loewenstein, G., Issacharoff, S., 1994. Source dependence in the valuation of objects. J. Behav. Decis. Mak. 7 (157-168).
Mazar, N., Amir, O., Ariely, D., 2008. The dishonesty of honest people: a theory of self-concept maintenance. J. Market. Res. 45 (6), 633–644.
Murnighan, J.K., Oesch, J.M., Pillutla, M., 2001. Player types and self-impression management in dictatorship games: two experiment. Games Econ. Behav. 37, 388–414.
Quattrone, G.A., Tversky, A., 1984. Causal versus diagnostic contingencies: on self-deception and on the voter's illusion. J. Personal. Soc. Psychol. 46 (2), 237–248.
Saucet, C., Villeval, M.-C., 2019. Motivated memory in dictator games. Games Econ. Behav. 117, 250–275.
Schwardmann, P., van der Weele, J., 2019. Deception and self-deception. Nat. Hum. Behav. 1–7.
Shu, L.L., Gino, F., Bazerman, M.H., 2011. Dishonest deed, clear conscience: When cheating leads to moral disengagement and motivated cognition. Personal. Soc. Bull. 37 (3), 330–349.
Tonin, M., Vlassopoulos, M., 2013. Experimental evidence of self-image concerns as motivation for giving. J. Econ. Behav. Organ. 90, 19–27.
van der Weele, J., von Siemens, F. A., 2014. Bracelets of pride and guilt? an experimental test of self-signaling in charitable giving. CESifo Working Paper No. 4674.
Zimmerman, F., 2020. The dynamics of motivated beliefs. Am. Econ. Rev. 110 (2), 337–361.