A simple stress test of experimenter demand effects

Piers Fleming · Daniel John Zizzo

Abstract As a stress test of experimenter demand effects, we run an experiment where subjects can physically destroy coupons awarded to them. About one subject out of three does. Giving money back to the experimenter is possible in a separate task but is more consistent with an experimenter demand effect than an explanation based on altruism towards the experimenter. A measure of sensitivity to social pressure helps predict destruction when social information is provided.

Keywords   Altruism · Experimenter demand effects · Social information · Social pressure

JEL Classification   B41 · C72 · C91 · D64

1 Introduction

This paper presents a stress test of experimenter demand effects, which refer to changes in behaviour by experimental subjects due to cues about what constitutes an appropriate behaviour (Zizzo 2010). The idea behind experimenter demand effects is that subjects try to make sense of an unfamiliar experimental environment in order to decide an appropriate response, and in doing so they may be particularly sensitive to whatever cues are provided in such an environment. Our stress test of experimenter demand effects is of particular relevance to the interpretation of results in settings, such as public good games, trust games or bargaining games, that are characterized by *variable*

P. Fleming
School of Psychology and CBESS, University of East Anglia, Norwich NR4 7TJ, UK

D. J. Zizzo (✉)
School of Economics and CBESS, University of East Anglia, Norwich NR4 7TJ, UK
e-mail: d.zizzo@uea.ac.uk

🖄 Springer

surplus. That is, how the overall amount is split across the subjects depends on the actions by the subjects.

We run an experiment where subjects can physically destroy money-equivalent coupons awarded to them[1] as well as, in a different task, explicitly return money to the experimenter. Subjects affected by experimenter demand effects will consider both tasks to be identical where, given their stylized nature and the absence of any simple alternative schema to make sense of them, there is an expectation that they should physically destroy or return some of the assets given to them. The distinctive feature of the destruction task is that the coupons' destruction could not directly benefit the experimenter. This is in contrast to the cash return task or, indeed, any experiment where a money transfer to the budget of the experimenter implicitly takes place if experimental surplus is destroyed or not obtained. In the cash return task, unlike the destruction task, there *can* be a money transfer towards the experimenter, so altruism towards the experimenter may potentially affect behaviour and provide an alternative explanation of behaviour relative to experimenter demand effects.

We are aware of two papers that have tested the possibility of altruism towards the experimenter. In an insightful contribution, Frank (1998) did so in the context of ultimatum games: in the experimental treatment, if the proposer's offer was rejected, the money-equivalent currency (stamps) was physically burned; they found that receivers did not behave differently in this treatment relative to a control where the money implicitly went back to the experimenter. However, the strategic nature of ultimatum games may have meant that feelings of anger out of unfairness may have reduced the relevance of altruism towards the experimenter in this context; furthermore, this experiment did not control for the possibility that some subjects would gain utility from seeing stamps physically burned in the laboratory. Harrison and Johnson (2006) considered the effect of changing the recipients in standard dictator games, and found that giving to the experimenter was intermediate between giving to a charity and giving to another subject.[2] This could be interpreted as evidence of altruism towards the experimenter, but it is equally possible that returning money to the experimenter by the implied transfer of unexploited experimental surplus could *in itself* be due to experimenter demand, i.e. being due to the demands of the experimental decision environment, rather than being due to altruism towards the experimenter.[3]

---

[1] We cannot use money directly because in the U.K. (as in many other countries) it is illegal to destroy money. We later describe experimental checks for our claim that we can treat coupons as equivalent to money in the experiments.

[2] An interesting but unusual feature of their experiment is that, in their baseline, subjects were informed that the experimenters had brought a certain amount of money to the experimental room.

[3] [Harrison and Johnson (2006), footnote 1] are clear that they interpret altruism very broadly in their paper, and are not looking to draw conclusions on the psychological motives underlying behaviour from their experiment. Apart from experimenter demand effects, revealed altruism in dictator games may for example be due to pro-social norms of giving (Guala and Mittone 2010). Below we simply speak of altruism towards the experimenter, but ruling this out as an explanation will also rule out an interpretation of our findings in terms of pro-social norms of giving.

We use an extremely simple setup with minimal strategic concerns or potential for misunderstanding to identify experimenter demand effects against alternative explanations. By providing an equivalent physical task to all subjects regardless of whether coupons are destroyed, we control for the pleasure of physical activity. Our experiment also controls for the potential benefits of coupon destruction for the coupons provider and for the clarity of the instructions. We are then able to separate out experimenter demand effects from altruism towards the experimenter as possible explanations of behaviour. If subjects are driven by altruism towards the experimenter in deciding what to do in experiments with variable surplus tasks, we would expect subjects to return money to the experimenter but *not* to destroy the coupons as this could not benefit the experimenter. Conversely, if experimenter demand effects drive both physical coupon destruction *and* returning money to the experimenter, both should be positive and should be positively correlated to each other.

We have a social information treatment manipulation where we provide summary information about how subjects behaved in a pilot. We also have a partial, albeit imperfect, measure of sensitivity to social pressure by using the Stöber (2001) social desirability scale.[4] Intuitively, under social information the *experimental norm* (what is expected of subjects in the experimental environment) should be clearer and, therefore, subjects who are more responsive to social pressure should destroy more (conversely, subjects who are resistant to social pressure may destroy less). This provides a further stress test of experimenter demand effects.[5]

Section 2 presents the experimental design, Section 3 presents the results and Section 4 briefly concludes. The experimental instructions are available in the appendix.

2 Experimental design

2.1 Design

64 participants completed the social desirability measure online. About one week later participants were invited back to an experimental laboratory session. Participants completed two tasks in counterbalanced order followed by debriefing questions. Questions to check understanding were given ahead of each task, and clarifications were given to subjects who gave any incorrect question.

The two tasks were a (coupons) *destruction task* and a (cash) *return task*. In the destruction task participants were given six 50 pence paper coupons worth £3 and six paper blanks. The coupons were redeemable at university cafés.[6] The participants were

---

[4] An example of question on this scale 'I always stay friendly and courteous with other people even when I'm stressed out' Stöber (2001).

[5] Bischoff and Frank (2011) found only limited evidence for demand effects driven by the way the experimenter is dressed and reads the instructions; Zizzo and Fleming (2011) interpreted evidence from back-to-back dictator and money reduction games in terms of sensitivity to social pressure. Zizzo (2010) contains an overview of research related to experimenter demand effects.

[6] We used coupons rather than stamps as in Frank (1998) because we felt that, with the universal diffusion of mobile and IT forms of communications, particularly among the younger generations, many students rarely make direct use of stamps nowadays (unlike, still, when Frank ran his experiment).

told 'you need to decide how many vouchers to **destroy** by shredding them'. For each coupon not destroyed a paper blank was destroyed instead, using the same shredder, so that the same physical activity took place for all participants, thus controlling for any pleasure from the physical activity.[7] Participants were discreetly observed during this task. Participants kept any non-destroyed coupons for use after the experiment. In the return task participants were given six 50 pence pieces worth £3. The participants were told 'you need to decide how much cash to **return** to the experimenters' and were provided with an envelope into which cash could be returned. Any remaining cash and coupons could be kept as participant payment.[8]

There were two experimental treatments. In a *NoInfo* treatment ($n = 31$), there were no further instructions. In the *SocialInfo* treatment ($n = 33$), subjects were told, truthfully (as we relied on pilot data), the following social information: 'In test sessions… some people destroyed vouchers/returned cash - of those that did, on average (rounded to the nearest 50p) they destroyed/returned £1 of vouchers/cash'.

After the decision tasks participants were asked about their perceived altruism in voucher destruction, 'Do you think destroying vouchers would be beneficial to [campus cafés]?' on a Likert scale from 1 *not at all beneficial* to 7 *extremely beneficial*. As a further check, subjects were also asked about how clear they had found the instructions for each of the two tasks, again on a 1–7 Likert scale.

3 Experimental results

Figure 1 presents histograms with amounts destroyed and returned. 30 % of participants in the *NoInfo* treatment and 35 % of participants in the *SocialInfo* treatment did destroy some coupons. 6.6 and 11.1 % of endowments were, respectively, returned and destroyed in the *NoInfo* treatment, vs. 9.7 and 11.8 % in the *SocialInfo* treatment. The amounts destroyed and returned are comparable (Wilcoxon $p = 0.392$),[9] implying that altruism towards the experimenter (as separate from experimenter demand) is not the primary driver of behaviour, for it cannot explain the destruction activity. While there was more destruction or returning of coupons under *SocialInfo*, this does not achieve statistical significance (Mann–Whitney $p = 0.753$ and 0.258 for destruction and returning, respectively).

Table 1 presents a regression analysis on the amount destroyed. As destruction can only take place in blocks of 50 pence (the value of each coupon), the amount destroyed can take 7 levels, from 0 coupons destroyed to 6 coupons destroyed, and so an ordered probit model is appropriate; Tobit estimates are also reproduced as robustness checks. As independent variables we have CashReturned (the amount returned in the return task) in two of the regressions, SocialInfo (= 1 in *SocialInfo*, else 0), SocDes

---

[7] Physical destruction occurred by using shredders rather than burning the coupons in the light of U.K. health and safety regulations.

[8] Before leaving participants had the opportunity to sell their remaining vouchers using an incentivized Becker et al. (1964) minimum selling price methodology if so they wished. Their average valuations for the 50p coupons was 51 pence (SD = 18.49), confirming that the coupons were appropriately valued by subjects.

[9] All *p* values reported in this paper are two-sided.
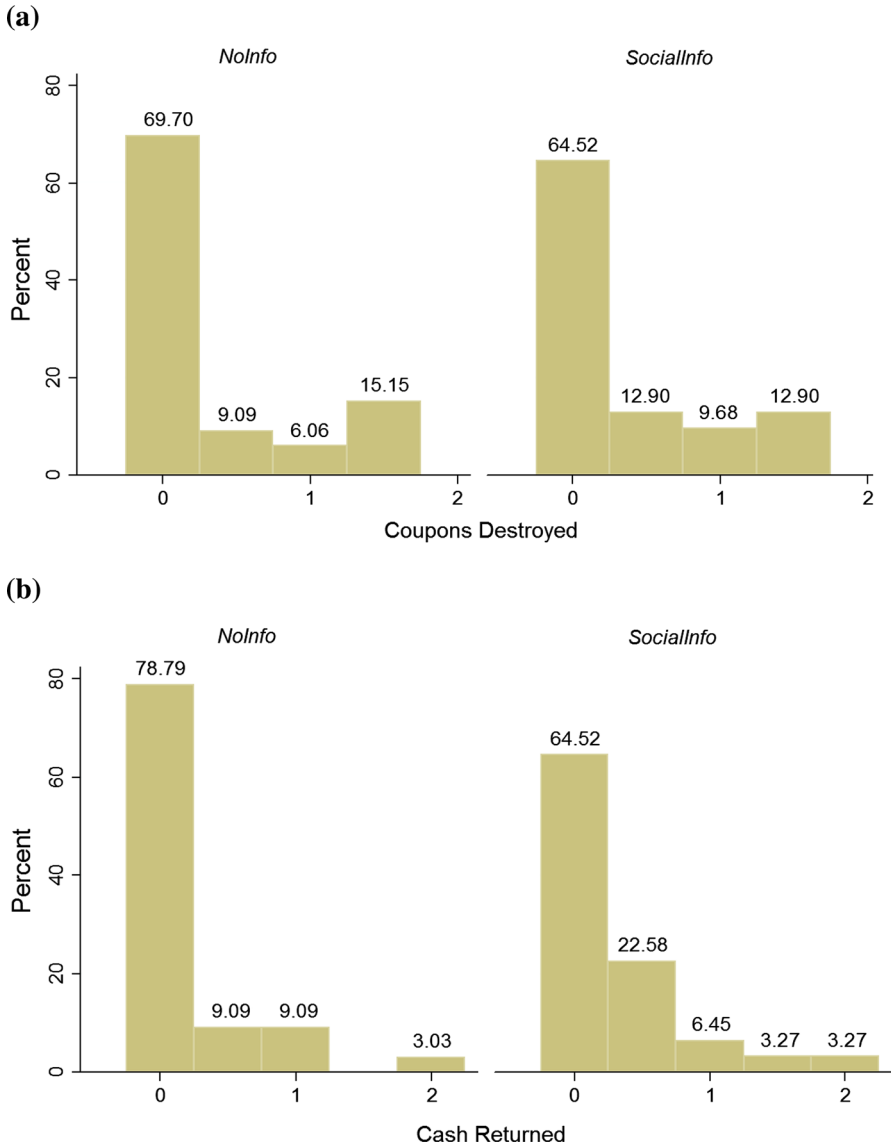
(a)

(b)

Fig. 1 Histograms of destruction and return choices in the NoInfo and SocialInfo treatments. **a** Destruction task. **b** Return task percentages are rounded to the second decimal place

(the Stöber measure of social desirability),[10] SocDesxSocialInfo (an interaction term), Beneficial (stated extent to which the coupons destruction is seen as beneficial towards the university café provider), ClearIDestruction (stated clarity of the destruction task instructions), ClearIReturn (stated clarity of the return task instructions). There is a clear and strong positive correlation between amount returned and amount destroyed,

---

[10] This is centered about the mean to reduce multicollinearity when interacted (Marquardt 1980).

which is consistent with an experimenter demand effect explanation rather than one based on altruism towards the experiment. There is also a positive coefficient on Beneficial, which is to be expected; however, the clarity of the instructions on destruction seems, if anything, to increase rather than decrease destruction,[11] suggesting that confusion was not a problem in our experiment.

SDS17xSocialInfo has a statistically significant positive coefficient. Subjects with a higher value of SDS17 are more socially sensitive to social pressure, vertical—i.e. experimenter demand—and horizontal—i.e. peer pressure. When social information is provided, this provides greater clarity on the social demands of the situation and subjects who are more sensitive to social pressure are then more responsive to destroy more; though, as there is no average effect of social information, this clarity would also seem to yield less destruction from subjects who are less sensitive to social pressure.

4 Conclusions

We considered a stress test of experimenter demand effects. Altruism towards the experimenter is unable to explain the key finding of destruction of coupons by roughly one subject out of three, the same as the fraction of subjects who returned money to the experimenter. While the latter can be explained by altruism, the former cannot; the strong positive correlation between coupons destroyed and money returned to the experimenter is also left unexplained by an explanation based on altruism towards the experimenter. Confusion is unlikely to be behind this as the tasks were simple, subjects' understanding was checked, and variables for clarity of the instructions on the destruction task were either statistically non significant or positively rather than negatively correlated with destruction. Subjects valued the coupons (see footnote 9) and our analysis also controlled for the pleasure of physical destruction by ensuring that the same physical destruction activity took place regardless of how many coupons were destroyed; and for the possibility of benefiting the café coupon providers by having a relevant question that went into our regression analysis.

Social information provided clarity on the experimental norm and, where present, our measure of sensitivity to social pressure predicted the amount of destructive activity, which is consistent with experimenter demand effects. The heterogeneity in degrees to which subjects are sensitive to social pressure is still largely a neglected variable in economic research, and it should not be.

---

[11] In regression 2, the only one not achieving at least $p < 0.1$ significance, $p = 0.106$.

Table 1 Regressions on amount destroyed

| | Ordered probit | | Marginal effects | | Tobit | |
| | Coefficients | | | | Coefficients | |
	(1)	(2)	(1)	(2)	(3)	(4)
CashReturned	1.389*** (0.452)		0.205*** (0.067)		0.902*** (0.301)	
SocialInfo	0.005 (0.346)	0.130 (0.340)	0.001 (0.051)	0.026 (0.067)	0.035 (0.320)	0.163 (0.375)
SocialDes	−0.108 (0.082)	−0.125 (0.083)	−0.016 (0.013)	−0.025 (0.167)	−0.092 (0.075)	−0.122 (0.088)
SocialDes×socialInfo	0.253** (0.121)	0.270** (0.117)	0.037** (0.019)	0.053*** (0.024)	0.189* (0.108)	0.268** (0.126)
Beneficial	0.202** (0.103)	0.212** (0.101)	0.299* (0.164)	0.042** (0.021)	0.179* (0.095)	0.214* (0.112)
ClearIDestruction	0.416** (0.198)	0.299 (0.185)	0.061** (0.031)	0.059 (0.037)	0.302** (0.174)	0.289 (0.201)
ClearIReturn	−0.259 (0.185)	−0.188 (0.187)	−0.038 (0.028)	−0.037 (0.037)	−0.187 (0.163)	−0.151 (0.192)
Constant					−1.843(1.030)	−2.070(1.286)
Log likelihood	−50.351	−56.529			−49.495	−53.836

$n = 63$ for all regressions (one subject failed to answer the questions on the clarity of instructions). The reported marginal effects measure how a marginal increase in each independent variable in the ordered probit regressions 1 and 2 increase the probability of the highest amount being destroyed (i.e. 1.5 coupons, as shown by Fig. 1a). Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix

Sessions vary according to:

- *Order 0 vs Order 1*
- *NoInfo vs SocialInfo [no information vs social information]*

Experimental instructions

Thank you for taking part in this experiment on decision making. Please raise your hand if you have any questions at any point in the experiment. Please do not touch any of the items in front of you unless told otherwise.

There are three stages in this experiment; all of your decisions are confidential. Your decisions will be anonymized for analysis. Your decisions alone will determine your earnings.

You will receive £2 for attending this experiment.

[Order 0]

In stage 1 you will be given £3 worth of vouchers which can be spent in Blend/Zest/Cafe Direct any time this month. You will then decide how much to keep. You can choose to keep all £3.

In stage 2 you will be given £3 in coins. You will then decide how much to keep. You can choose to keep all £3.

Before you leave you will answer some questions.

You will be given the money in advance of completing the experiment so we will now ask you to sign that you have received the cash and vouchers. Please empty the two brown envelopes in front of you (named 'Vouchers' and 'Cash') and count their contents.

Please raise your hand when you are ready to continue.

These questions are only to make sure that you have understood the instructions. Please raise your hand when you have completed them.

1. What is the maximum value of vouchers and money you can choose to leave this experiment with? _____
2. What is the minimum value of vouchers and money you can choose to leave this experiment with? _____
3. Will all of your earnings be in cash? Yes / No

Voucher decision [Order 0: Stage 1]

In the voucher envelope you have been given £3 worth of vouchers which can be spent in Blend/Zest/Cafe Direct on campus this month. There are also an equivalent number of worthless pieces of paper (which you will not keep).

In a moment you will be asked to destroy six of these items (these vouchers or the worthless paper equivalents). We would like you to practice this in a moment.

Please do not destroy anything right away but raise your hand and wait for the experimenter to tell you to proceed.

Practice

In front of the shredder is a blank piece of paper approximately of the same size as a voucher. Please carefully use the shredder to destroy it.

Decision

You have one decision to make in this stage. Namely, you need to decide how many vouchers to **destroy** by shredding them.

You must shred six items. If you decide not to destroy a voucher in the shredder you are asked to destroy a piece of blank paper instead. You keep any vouchers you do not destroy.

[*Treatment SocialInfo only:*

In test sessions of this experiment which we ran in May some people destroyed vouchers - of those that did, on average (rounded to the nearest 50p) they destroyed £1 of vouchers.]

Please now do the following in sequence:

1. Shred six items.
2. Place any of the unshredded blank paper slips back in the envelope.
3. Turn over the page.

Please write down the value of vouchers that you have destroyed (£0–£3): £_____
Please raise your hand to let the experimenter know when you are ready to proceed.

Cash decision [Order 0: Stage 2]

In the cash envelope you have been given £3.
You have one decision to make in this stage. Namely, you need to decide how much cash to **return** to the experimenters. You keep any cash you do not return.
Please now do the following in sequence:

1) Place any cash that you wish to return in the envelope back in the envelope.
2) Turn over the page.

[*Treatment SocialInfo only:*

In test sessions of this experiment which we ran in May some people returned cash—of those that did, on average (rounded to the nearest 50p) they returned £1 of cash.]

Please write down the amount of cash you put in the envelope (£0–£3): £ _____
Please raise your hand to let the experimenter know when you are ready to proceed.

Order 1:

Same instructions as above but with two changes relative to Order 0: (a) Stage 1 of Order 0 becomes stage 2 and Stage 2 of Order 0 becomes Stage 1; (b) corresponding small changes in text.

Final Questions [common for Order 0 and Order 1]

Regarding the decision about vouchers – why did you make the decision you made?

Regarding the decision about cash – why did you make the decision you made?

Do you believe that any shredded vouchers are destroyed and made worthless? Yes / No

Do you think destroying vouchers would be beneficial to Blend/Zest/Cafe Direct?

not at all beneficial						extremely beneficial
1	2	3	4	5	6	7

How clear were the instructions for the decision about vouchers?

not at all clear						extremely clear
1	2	3	4	5	6	7

Is there anything you did not understand for the decision about vouchers?

How clear were the instructions for the decision about cash?

not at all clear						extremely clear
1	2	3	4	5	6	7

Is there anything you did not understand for the decision about cash?

```
┌─────────────────────────────────────────────────────────────┐
│                                                             │
│                                                             │
└─────────────────────────────────────────────────────────────┘
```

Did you feel under any pressure or expectation to destroy the vouchers?

Not at all   Very much

                                                              so

  1   2   3   4   5   6   7

Did you feel under any pressure or expectation to return the cash?

Not at all   Very much

                                                              so

  1   2   3   4   5   6   7

Please make sure that you answered all questions. Then raise your hand to indicate that you are ready to proceed.

Minimum selling price

We will now give you the opportunity to sell your remaining vouchers (six minus the number you destroyed before). We will be willing to pay a random amount X between 1p and 100p for each voucher. You must decide how much you are willing to accept for your vouchers.

If your stated willingness to accept is smaller or equal to the randomly drawn price X, then we will purchase all your remaining vouchers for the randomly drawn price X - not for your stated willingness to accept. You receive the price X for each remaining voucher.

If your stated minimum selling price per voucher is greater than the randomly drawn price X, then we will not purchase any vouchers from you. You keep all of your remaining vouchers.

*Example Imagine that you have 4 remaining vouchers and stated a minimum selling price of 61 pence per voucher. If the randomly drawn price X were, say, 64 pence, we would buy all your vouchers at a price of 64 pence - not 61 pence. Thus you would get paid 4 * 64p = £2.56. If the randomly drawn price X were, say, 53 pence, we would not buy the vouchers from you and you would leave with your vouchers. Because your answer does not affect the amount you get paid, only whether you sell, it is to your advantage to state the minimum that you would be willing to be paid for selling each of your vouchers.*

What would be the lowest price at which you were willing to sell each of your remaining 50p vouchers (for use in Blend/Zest/Cafe Direct)?

My minimum selling price is _____ pence per voucher.

After writing down your minimum selling price please raise your hand to indicate that you are ready to proceed.

To be completed by the experimenter:

Random number drawn: _____

$\boxed{MSP \leq X}$

Sell all remaining vouchers for price X

      Number of remaining vouchers: _____

      Price X: _____

      Total payment for vouchers: _____

$\boxed{MSP > X}$

Leave with all remaining vouchers – no payment

Session: _____

Participant: _____

Demographic Questions

Age: ____

Gender: Female___ Male____

Nationality: _____

Study level: Undergraduate___ Postgraduate____

Do you study or did you study economics as a main university subject? (That is, with economics in the degree title, e.g. BA Banking and Economics)?

Yes___ No____

Finally, please put this sheet and both brown envelopes ('Vouchers' and 'Cash') into the large white envelope provided and give them to the experimenter when you leave. Please close but **do not seal** any of the envelopes.

Please do not discuss this experiment until we have finished testing (the end of this week). Thank you for taking part. If you wish to know more about this experiment, please email Dr. Piers Fleming (p.fleming@uea.ac.uk) and you will be put on a debriefing list. A debriefing note on the experiment will then be emailed to you after all experimental sessions have been run.

References

Becker, G.M., DeGroot M.H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9,* 226-232.

Bischoff, I., & Frank, B. (2011). Good news for the experimenters: Subjects are hard to influence by instructors' cues. *Economics Bulletin, 31,* 3221–3225.

Frank, B. (1998). Good news for the experimenters: Subjects do not care about your welfare. *Economics Letters*, *61*, 171–174.

Guala, F., & Mittone, L. (2010). Paradigmatic experiments: The dictator game. *Journal of Socio Economics*, *39*, 578–584.

Harrison, G. W., & Johnson, L. T. (2006). Identifying altruism in the laboratory. In D. Davis & R. Mark Isaac (Eds.), *Experiments Investigating Fundraising and Charitable Contributors. Research in experimental economics* (Vol. 11, pp. 177–223). Amsterdam and San Diego: Elsevier.

Marquardt, D. W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association*, *75*, 87–91.

Stöber, J. (2001). The social desirability scale-17 (SDS17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, *17*, 222–232.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*, 75–98.

Zizzo, D. J., & Fleming, P. (2011). Can experimental measures of sensitivity to social pressure predict public good contribution? *Economics Letters*, *111*, 239–242.