Evolution and Human Behavior 34 (2013) 35–40

Contents lists available at SciVerse ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

The watching eyes effect in the Dictator Game: it's not how much you give, it's being seen to give something

Daniel Nettle [a,*], Zoe Harper [b], Adam Kidson [b], Rosie Stone [b], Ian S. Penton-Voak [c], Melissa Bateson [a]

[a] *Centre for Behaviour and Evolution & Institute of Neuroscience, Newcastle University*
[b] *School of Psychology, Newcastle University*
[c] *Department of Experimental Psychology, University of Bristol*

ARTICLE INFO

Article history:
Initial receipt 2 July 2012
Final revision received 27 August 2012

Keywords:
Dictator Game
Cooperation
Generosity
Behavioral economics
Watching eyes

ABSTRACT

In a classic study, Haley and Fessler showed that displaying subtle eye-like stimuli caused participants to behave more generously in the Dictator Game. Since their paper was published, there have been both successful replications and null results reported in the literature. However, it is important to clarify that two logically separable effects were found in their original experiment: watching eyes made the mean donation higher, and also increased the probability of donating something rather than nothing. Here, we report a replication study with 118 participants, in which we found that watching eyes significantly increased the probability of donating something, but did not increase the mean donation. Results did not depend on the sex of the participants or the sex of the eyes. We also present a meta-analysis of the seven studies of watching eye effects in the Dictator Game published to date. Combined, these studies total 887 participants, and show that although watching eyes do not reliably increase mean donations, they do reliably increase the probability of donating something rather than nothing (combined odds ratio 1.39). We conclude that the watching eyes effect in the Dictator Game is robust, but its interpretation may require refinement. Rather than making people directionally more generous, it may be that watching eyes reduce variation in social behavior.

1. Introduction

In a landmark study, Haley and Fessler (2005) showed that placing subtle eye-like stimuli in the participants' environment during an experimental economic game called the Dictator Game (DG) caused them to be more generous towards an anonymous other person than when no such stimuli were displayed. It was already known that people behave more generously when under the gaze of real human eyes (Bull & Gibson-Robinson, 1981; Kurzban, 2001). This is not surprising given that human cooperative behavior is thought to be largely maintained by the social sanctions and reputational costs which tend to fall on those who are not sufficiently prosocial and whose behavior comes to be known to others (Barclay, 2004; Fehr & Gachter, 2002; Milinski, Semmann, & Krambeck, 2002; Sylwester & Roberts, 2010; Wedekind & Milinski, 2000). However, the important achievement of Haley and Fessler (2005) was to show that eye-like images were sufficient to increase prosocial behavior, even though they were artificial and not particularly realistic, and conflicted with the explicit information on their anonymity which participants were given. This suggests the existence of automatic cognitive mechanisms for detecting social gaze and regulating social behavior accordingly (Izuma, 2012), which can be activated by cues only minimally similar to real observers.

Haley and Fessler's paper has provoked enormous interest and a great deal of research on the proximate psychological mechanisms regulating prosocially oriented behaviors. There have been a number of successful replication studies using the DG or close variants of it (Keller & Pfattheicher, 2011; Oda, Niwa, Honma, & Hiraishi, 2011; Rigdon, Ishii, Watabe, & Kitayama, 2009), or showing similar effects using different games (Burnham & Hare, 2007 using the Public Goods Game). There has also been a replication using a non-monetary laboratory task (Bourrat, Baumard, & McKay, 2011), and a series of field studies showing that eye images can enhance generosity and prosocialbehavior in naturalistic settings (Bateson, Nettle, & Roberts, 2006; Ekström, 2011; Ernest-Jones, Nettle, & Bateson, 2011; Francey & Bergmüller, 2012; Powell, Roberts, & Nettle, 2012). On the other hand, there have also been several laboratory studies that reported no watching eyes effect, two using the DG as used by Haley and Fessler (Raihani & Bshary, 2012; Tane & Takezawa, 2011), and another the Trust Game (Fehr & Schneider, 2010). There are important procedural differences between these studies which may explain the heterogeneity of findings. For example, Tane and Takezawa's (2011) study was conducted in the dark, where perceived observability is low and thus a watching eyes effect was not expected. Nonetheless, a fair summary of the current state of the literature is that the evidence is mixed. It is

* Corresponding author. Henry Wellcome Building, Framlington Place, Newcastle, NE2 4HH, UK.
E-mail address: daniel.nettle@ncl.ac.uk (D. Nettle).

1090-5138/$ – see front matter
http://dx.doi.org/10.1016/j.evolhumbehav.2012.08.004

therefore important to synthesize the existing data on the effect, and to search for factors moderating it which explain discrepancies between studies.

The current paper has two components. The first component is a novel empirical study of the watching eyes effect using the DG paradigm. This is partly for replication, and partly to explore sex of participant and sex of eyes as possible moderating factors. Although Haley and Fessler (2005) found no interaction between sex of participant and experimental condition, Rigdon et al. (2009) found that the watching eyes effect was only present in the male participants. This may have been because female participants were significantly more generous than males in the control condition, and so there is less scope for eyes to increase their giving. The potential importance of sex of eyes is suggested by data in the study by Bateson et al. (2006). Bateson et al. used alternating presentations of different eye images and control images, and found the greatest contributions when male rather than female eyes were displayed. However, their study was not designed to test for the importance of sex of eyes, and was unable to determine whether this was due to chance, or to some feature other than sex of the particular images used (e.g. age, or obliqueness of gaze).

In our study, then, we individually administered a DG in a laboratory where no-one was able see at the moment the participant made their anonymous decision, but where was displayed an ostensibly irrelevant poster which featured either an image of eyes or a control image. To minimize the impact of any idiosyncratic features of the stimulus eyes other than their sex, we used created composite male and female faces, and to maximize the chance of detecting an effect of sex of eyes, we morphed the female eyes to be extremely feminine, and the male ones to be extremely masculine. The eyes from these morphed images were then used in the poster. We sought to determine whether the watching eyes effect described by Haley and Fessler (2005) could be found, and whether it was moderated by sex of participant, sex of eyes, or the interaction of the two.

The second component of our paper is a meta-analysis of the studies to date which have looked for the watching eyes effect in the DG. Here, we sought to determine whether the balance of evidence to date supports the robustness of the effect in this context, or suggests that the original findings were false positives. We also sought to clarify exactly what the nature of the effect is. Haley and Fessler (2005) in fact reported two different results. First, the mean donation was significantly higher in the eyes than the control conditions. We call this the *increased mean donation effect*. Second, participants in the eyes conditions were significantly more likely to donate something as opposed to nothing. We call this the *increased probability of donation effect*. Note that finding the increased probability of donation effect does not automatically entail finding the increased mean donation effect. For example, if donations in the control condition had a larger variance than in the eyes conditions, there might be an increased probability of donation in the eyes conditions, but a mean donation that is no greater (or even, that is less) as compared to the control condition. There is some evidence that such situations do occur: Rigdon et al. (2009) found an increased probability of donation effect without an increased mean donation effect. Moreover, an underappreciated feature of Haley and Fessler's original study is that the variation in donations was substantially smaller in the eyes than in the control conditions (coefficients of variation (CV): eyes conditions 0.74, control conditions 1.15).

Despite the two effects being potentially distinct, subsequent studies have not consistently tested for both of them in their data (for example, Oda et al., 2011, Tane & Takezawa, 2011, and Raihani & Bshary, 2012 only report on the increased mean donation effect, and not the increased probability of donation effect). We seek to rectify this in our meta-analysis by systematically testing for both effects in all published studies.

2. Original study: Methods

2.1. Participants and design

Participants were 118 students of Newcastle University (57 males, 61 female), with 95% falling in the 18–25 age range. Some of the participants could count this study towards a certified amount of research participation as part of their BSc Psychology, but the financial incentives were identical for all participants. Participants were randomly assigned to enter the laboratory in one of three conditions: when control posters were displayed (control), when the masculinized eyes posters were displayed (male eyes), or when feminized eye posters were displayed (female eyes). Participants were assigned to the three conditions in the ratio of 4:3:3 respectively. This represented a compromise between the optimal statistical power for comparing the eyes conditions to the control condition, and the eyes conditions to each other. The study was approved by the psychology subcommittee of the Faculty of Medical Sciences ethics committee at Newcastle University.

2.2. Posters

Male and female average composite faces were constructed from individual images of 29 male and 57 female white European faces, each of which was delineated with 219 points allowing shape, colour and texture to be averaged using standard methods (Tiddeman, Burt, & Perrett, 2001). Full details of the standardized photography setup are available in Stephan, Penton-Voak, Perrett, Tiddeman, Clement, and Henneberg (2005). The composites were then caricatured in shape to generate an artificially masculinized male face and a feminized female face (with the differences between the male and female faces exaggerated by 50% in the appropriate direction, see Perrett et al., 1998). Laminated A4 posters were created, consisting of the irrelevant message 'Please do not eat or drink in the cubicle' plus the eyes from either the masculinized or feminized composite face, or the logo of the institution as a control image (Fig. 1). Five identical posters were displayed around the laboratory fixed to the walls at a height of approximately 1.5 m from the floor. One was within the cubicle where the participant sat, slightly to the right of the direction the participant would naturally face.

2.3. Dictator game procedure

Participants were met by an experimenter and led into the laboratory which had been preset with the appropriate posters. After instruction and filling in a consent form, the participant was screened from the experimenter's view with curtains. On the desk in the experimental cubicle were a preliminary questionnaire and DG instructions, a cardboard box, and two envelopes, one marked 'For you to take away' and the other marked 'For a random student'. Also on the desk was £5 in the form of 4x £1 coins and 2x £0.5 coins. Participants were instructed to put whatever money they wanted to keep into the 'take away' envelope, which they could then pocket, and whatever money they wished to donate to a random student in the 'random student' envelope, which they were then instructed to place into the cardboard box. Participants were informed that the 'random student' envelope would be given to a member of the student body. In fact, the recipients were students who participated in the experiment on a subsequent occasion who received this additional envelope unexpectedly after they had completed their experimental session. We did not make this explicit in advance, as participants often attend sessions in groups of friends who wait whilst each participates in turn, and we were concerned that dictators' perceptions of anonymity might be compromised if they knew who the next person coming in was. 'Unknown recipient' procedures, where the recipient is not specified as another participant in the same experiment, have been

Fig. 1. The posters used in this study *in situ* in the laboratory.

used in DG research before, and lead to comparable results to when there are multiple simultaneous participants and dictators know that someone else in the session is the recipient (Koch & Normann, 2008). After making their decision, participants were asked to rate how anonymous their decision felt on a seven point scale from strongly agree to strongly disagree.

2.4. Data analysis

We first tested for an increased mean donation effect by combining the male eyes and female eyes conditions and comparing both to the control. We used a General Linear Model with amount donated to the random student as the outcome variable and condition (eyes vs. control), sex of participant, and their interaction as predictors. As donated amounts were not normally distributed, we also used non-parametric Mann–Whitney U tests to confirm the General Linear Model analysis. We then tested for an increased probability of donation effect using a binary logistic Generalized Linear Model in SPSS version 19, with donation (0 = nothing, 1 = something) as the outcome variable, and condition (eyes vs. control), sex of participant and their interaction as predictors. We then went on to compare donation behavior in just the two eyes conditions, using sex of eyes, sex of participant and their interaction as the predictor variables. For amount donated, we did this with a General Linear Model, again confirmed with non-parametric Mann–Whitney U tests. For probability of donation, there were too few cases to fit a Generalized Linear Model, so we used Fisher's exact test to examine the proportions of individuals donating something rather than nothing in the male versus female eyes conditions, both overall, and for each sex of participant separately.

3. Results

The data from the study are available in 'csv' format as Electronic Supplementary Material (available on the journal's website at www.ehbonline.org). Eighty-six of the 118 participants (73%) donated something to the random student, with the remainder taking the whole £5 for themselves. The mean amount donated was £1.56 (s.d. £1.47) overall, and £2.15 (s.d. £1.30) amongst those who donated something. Participants' ratings of how anonymous they felt did not differ significantly between the eyes and no-eyes conditions (Eyes: M 4.55, s.d. 1.94, No eyes: M 4.33, s.d. 1.91, $t_{116} = 0.52$, $p = 0.53$), or between the sexes (Female participants: M 4.38, s.d. 1.73, Male participants: M 4.54, s.d. 2.11, $t_{116} = 0.47$, $p = 0.64$).

3.1. Donation behavior in combined eyes conditions versus control condition

Fig. 2 shows the donations participants made for the control conditions versus the two eyes conditions combined. In a General Linear Model with amount donated as the outcome and condition (eyes vs. no eyes), participant sex, and their interaction as predictors, there was a significant effect of sex ($F_{1,114} = 4.13$, $p = 0.05$), with a marginal mean of £1.83 for female participants and £1.27 for male participants. There was no significant effect of condition ($F_{1,114} = 0.03$, $p = 0.86$), or condition by sex interaction ($F_{1,114} = 0.06$, $p = 0.81$). Non-parametric Mann–Whitney U tests confirmed that there was no difference in amount donated between the eyes and no eyes conditions, either overall ($U = 1505$, $Z = -1.03$, $p = 0.30$), in the male participants ($U = 348.5$, $Z = -0.71$, $p = 0.48$), or in the female participants ($U = 396.5$, $Z = -0.86$, $p = 0.39$). Thus, there was no evidence of an increased mean donation effect.

Using a binary logistic Generalized Linear Model to predict donating something versus nothing, with participant sex, condition and their interaction as predictors, we found significant effects of participant sex (Wald $\chi^2 = 5.47$, $p = 0.02$) and of condition (Wald $\chi^2 = 4.26$, $p = 0.04$), but no significant sex by condition interaction (Wald $\chi^2 = 0.09$, $p = 0.77$). The sex effect was due to female participants being more likely than male participants to donate something (82% of women did so versus 63% of men). The condition effect was due to participants in the eyes conditions being more likely

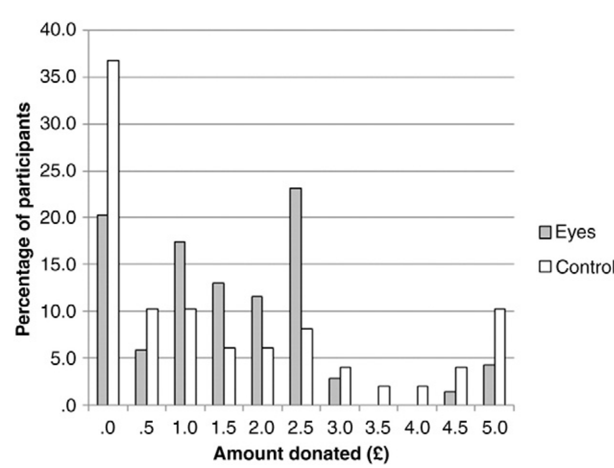

Fig. 2. Histogram showing Dictator Game allocations for participants in the eyes conditions and control condition.

to donate something than participants in the control condition (80% versus 63% of participants). Thus, as Fig. 2 suggests, there was an increased probability of donation effect: eyes posters shifted people away from donating nothing towards being more likely to donate something. However, it did not make them donate any more if they do donate. In fact, participants who did donate something donated non-significantly more in the control condition than in the eyes conditions (Eyes: M £1.98, s.d. £1.08; Control: M £2.44, s.d. £1.61, $t_{84}=1.56$, $p=0.12$). This is driven by the fact that 22% of the participants in the control condition donated more than half of their stake, whereas only 9% of the participants in the eyes conditions did so. This difference was near-significant (Fisher's exact test, $p=0.06$). A consequence of the greater number of zero donations and the greater number of large donations in the control condition is that the standard deviation was significantly smaller in the eyes conditions than the control condition (CVs: eyes 0.79, control 1.13; Levene's test, $F=9.24$, $p=0.01$).

3.2. Donation behavior in male versus female eyes conditions

In a General Linear Model with amount donated as the outcome, there was a significant effect of sex of participant ($F_{1,65}=4.46$, $p=0.04$), with female participants donating more than male participants (marginal means: Female participants £1.89, Male participants £1.27). There was no effect of sex of eyes ($F_{1,65}=0.03$, $p=0.86$), or sex of participant by sex of eyes interaction ($F_{1,65}=0.93$, $p=0.34$).Non-parametric Mann–Whitney U tests confirmed that there was no effect of sex of eyes on amount donated overall ($U=557$, $Z=-0.46$, $p=0.64$), amongst the male participants ($U=142.5$, $Z=-0.07$, $p=0.95$), or amongst the female participants ($U=118.5$, $Z=-1.17$, $p=0.26$). There was no significant difference in the proportion of participants donating something between the male and female eyes conditions (male eyes: 26/35; female eyes: 29/34; Fisher's exact test, $p=0.37$). This was also true in each sex of participant considered separately (male participants, 12/17 and 12/17, Fisher's exact test, $p=1$; female participants, 14/18 and 17/17, Fisher's exact test, $p=0.10$).

4. Original study: Discussion

In an original laboratory study using the Dictator Game, with ostensibly irrelevant posters on the walls that featured either eyes or a control image, we found no evidence of an increased mean donation effect, but clear evidence of an increased probability of donation effect. 80% of participants exposed to eyes posters donated something to the other party, compared to 63% of those exposed to the control posters. Thus, our findings agree with those of Rigdon et al. (2009) in replicating one but not the other of the effects reported by Haley and Fessler (2005). We also found that donations in the control conditions were more variable overall than those in the eyes conditions. This replicates the results of Haley and Fessler (2005) in this regard, and the coefficients of variation were remarkably similar (Haley and Fessler: eyes 0.74, control 1.15; this study: eyes 0.79, control 1.13).

We found main effects of sex of participant on both mean donation and probability of donation, with women donating more, and being more likely to donate. This replicates sex differences which are often observed in the Dictator Game (Bolton & Katok, 1995; Eckel & Grossman, 1998), but are not universally found (Haley & Fessler, 2005 did not observe them). However, there was no interaction between sex of participant and condition. Thus, sex of participant is unlikely to be an important moderating factor in explaining when watching eyes effects will or will not be found. We found no evidence that the sex of the eyes is important, or any interaction between sex of eyes and sex of participant. This suggests that the apparently greater impact of the male eyes in the study by Bateson et al. (2006) may have been chance, or to do with features of those particular images other than their sex.

5. Meta-analysis of studies published to date

Using citation searches as well our knowledge of the literature, we have been able to locate five studies of the watching eyes effect in the Dictator Game in addition to Haley and Fessler (2005) and the study presented here. These seven studies present a mixed picture, with four presenting themselves as positive findings of an effect, one as a positive finding in a sub-group of participants, and two as negative findings. However, as discussed in the Introduction, studies are not all consistent in whether they test separately for an increased mean donation effect and an increased probability of donation effect. Thus, we extracted data from all papers to do this systematically. Where data were not reported in the papers, the authors have kindly supplied them. Our aim was to tabulate which of the two effects were present in which studies. In addition, we pooled all data from the seven studies to examine whether the set of results so far overall support the existence of either or both watching eyes effects. Doing this collapses studies which have important procedural differences which likely affected the results. Nonetheless, it is an informative way of assessing the current state of evidence.

The results of the meta-analysis are presented in Table 1. For each effect, we state whether the effect was found in the study (‘yes’ means that it was statistically significant or near significant ($p<0.1$), not just that the difference was in the predicted direction). Where the relevant comparison is not reported for the probability of donation we have carried out our own analysis using Fisher's exact test on the numbers of participants donating nothing and something in the eyes and control conditions respectively. Note that this exercise affects the conclusions one should draw from the studies in some cases. For

Table 1
Meta-analysis of the seven studies of the watching eyes effect in the Dictator Game published to date. For explanation, see text.

	Sample size	Increased mean donation effect		Increased probability of donation effect	
		Present?	Cohen's *d*	Present?	Odds ratio
Haley and Fessler (2005)	124	Yes	0.30	Yes	3.35
Rigdon et al. (2009)	113	No	0.18	Yes	1.91
Oda et al. (2011)	61	Yes	0.59	No[a]	0.97
Keller and Pfattheicher (2011)[b]	100	Yes[c]	0.06	Yes[c]	0.89
Tane and Takezawa (2011)[b]	80	No	−0.27	No	0.29
Raihani and Bshary (2012)[d]	291	No	−0.16	Yes	2.32
This study	118	No	0.03	Yes	2.28
Overall	887	No	0.04	Yes	1.39

[a] Donating something was almost universal in this study (all but 4 participants), and so there was no power to detect an increased probability of donation effect.
[b] Studies 1 and 2 have been combined.
[c] Effect was moderated by the individual-difference trait prevention-focus. Effect not significant if prevention focus not included in model.
[d] Based on comparing just the ‘eyes’ and ‘control’ conditions, though a condition with images of flowers was also run. If the control and flower conditions are pooled to make a super-control condition, then $p=0.07$ for the increased probability of donation effect, and the odds ratio becomes 2.00.

example, Raihani and Bshary (2012) present their study as a failure to find an eyes effect, but did not test for an increased probability of donation effect. In fact, such an effect is present, with 90.2% of participants donating something in the eyes condition versus 79.9% in the control condition (Fisher's exact test, $p = 0.02$). For the increased mean donation effect, two studies find it overall, a further one finds it in a subgroup of participants, and four find no evidence of it. For the increased probability of donation effect, four studies find it overall, a further one finds it in a subgroup of participants, and two find no evidence of it.

Table 1 also gives effect sizes for each study. For the mean donation effect, this is Cohen's *d*, the difference between the means expressed in pooled standard deviations (Cohen, 1988). Thus, zero is no effect, a positive number indicates an increased mean donation under eyes, and a negative number a decreased mean donation under eyes. For the probability of donation effect, the effect size measure is the odds ratio of donating something under eyes compared to under control conditions. Thus, 1 indicates no effect, >1 indicates that eyes increase the odds of donating, and <1 indicates that eyes decrease the odds. As well as computing effect sizes for each study individually, we were able to compute them for the seven studies combined. For the increased mean donation effect, the combined data show that this is not significantly different from zero (pooled $d = 0.04$, $t_{887} = 0.60$, $p = 0.55$). For the increased probability of donation effect, though, there is a significant effect in the pooled data (percentage of participants donating something: eyes 79.4%, control 73.4%, Fisher's exact test, $p = 0.04$). This represents an odds ratio of 1.39 (95% CI 1.02–1.91), which would be considered a moderate effect by most definitions. Removing from this analysis Oda et al. (2011), where almost all participants donated something, and Tane and Takezawa (2011), who placed participants in the darkness where the feeling of observability is low and thus a watching eyes effect is not expected, increases the odds ratio to 1.60 (95% CI 1.14–2.23).

It may not be appropriate to pool these very different studies to get an idea of the overall size of the watching eyes effect. For example, the study of Raihani and Bshary (2012) was carried out over the internet, whereas all others were completed in a laboratory, and that of Tane and Takezawa (2011) was carried out in the dark, whereas all other laboratory studies were performed in the light. Nonetheless, our analysis shows that if these procedural differences are set aside and the data simply pooled, the picture that emerges is that watching eyes robustly increase the odds of donating something, but do not tend to have any effect on the mean amount donated.

6. General discussion

In this paper, we have presented an original study and a meta-analysis of existing studies of the watching eyes effect in the Dictator Game. Both parts of the paper point to the same conclusion: watching eyes do not robustly increase mean donations in the Dictator Game, but they do robustly increase the odds of donating something rather than nothing to the other player. This has important implications methodologically, and interpretively.

Methodologically, these results show that it is essential when replicating or extending the watching eyes effect in experimental economic games to report the correct statistics. Although Haley and Fessler (2005) were explicit in their original paper that their effect was largely driven by an increased probability of donating something under eyes compared to in control conditions, Oda et al. (2011), Tane and Takezawa (2011), and Raihani and Bshary (2012) did not report the statistics required to assess whether such effects occurred in their experiments. No study of the watching eyes effect should be considered to have tested for it without examining the probability of donation. Moreover, it would be useful to report tests for differences in the variances of donations in eyes versus control condition. Another methodological implication is that statistical power for searching for a watching eyes effect is optimised by tasks in which about half of people donate nothing and about half donate something under control conditions. If donation is much rarer or much more common than this, then many participants will need to be run to detect a change in the probability of donation. Dictator Game scenarios with low stakes often lead to the majority of participants donating, and thus they are not very suitable for measuring the eyes effect economically.

The second implication is interpretative. The results presented here could be taken to mean that the probability of donation is simply a more robust outcome measure for the DG than the mean amount donated, which is likely to be affected by subtle framing or anchoring effects which erase any effects of the experimental manipulation. However, we feel that our results imply something more specific about the psychological consequences of watching eyes. For watching eyes to increase the probability of donating something without also increasing the mean donation, there must be some individuals who are high donors under control conditions who actually donate less under eyes. In other words, eyes do not make people more generous across the board. Such an effect is suggested by the smaller variance in donations under eyes which was present in Haley and Fessler (2005) and replicated in our study presented here. It is also consistent with the findings of Keller and Pfattheicher (2011). They found large positive watching eyes effect amongst participants high in prevention focus, but no effect in the sample overall, suggesting that the trends were in the opposite direction amongst those in prevention focus.

Thus, rather than seeing watching eyes as always making people more generous, it might be more accurate to say that they make people more resistant to extreme strategies, such as giving nothing or giving an oddly large amount(cf. Kummerli, Burton-Chellew, Ross-Gillespie, & West, 2010). There could be an adaptive logic to such behavior; when others may be watching, it pays to be neither a bad interaction partner nor foolishly generous. Alternatively, giving something but not too much may be perceived as normative, and eye images may make people more inclined to follow norms. The field findings of Bateson et al. (2006), Ernest-Jones et al. (2011) and Francey and Bergmüller(2012) can also be interpreted in this way: what these studies showed is that people followed a local norm (paying for coffee, clearing/sorting one's litter) more when eye images were displayed. Interpretations of the watching eyes effect in terms of resistance to extreme strategies or adherence to norms are substantially different to those discussed in earlier papers. Admittedly, those interpretations are not obviously consistent with the findings of Oda et al. (2011) that watching eyes make people anticipate a reward rather than fear a punishment for prosocial behavior, or those of Powell et al. (2012) that watching eyes made people more generous in a context where giving was rare, non-normative and visible. Future experiments need to be designed to tease apart the effects of watching eyes on people's willingness to adopt an extreme strategy, their propensity to follow a local norm, and their generosity *per se.*

Supplementary Materials

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.evolhumbehav.2012.08.004.

Acknowledgments

We would like to thank Johannes Keller, Stefan Pfattheicher, Ryo Oda, Nichola Raihani and Daniel Fessler for responding helpfully and promptly to our requests regarding the data from their studies.

References

Barclay, P. (2004). Trustworthiness and competitive altruism can also solve the “tragedy of the commons”. *Evolution and Human Behavior*, *25*(4), 209–220.

Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology Letters*, *2*, 412–414, http://dx.doi.org/10.1098/rsbl.2006.0509.

Bolton, G. E., & Katok, E. (1995). An experimental test for gender differences in beneficent behavior. *Economics Letters*, *48*(3–4), 287–292.

Bourrat, P., Baumard, N., & McKay, R. (2011). Surveillance cues enhance moral condemnation. [Article]. *Evolutionary Psychology*, *9*(2), 193–199.

Bull, R., & Gibson-Robinson, E. (1981). The influences of eye-gaze, style of dress and locality on the amounts of money donated to a charity. *Human Relations*, *34*, 895–905.

Burnham, T. C., & Hare, B. (2007). Engineering human cooperation — does involuntary neural activation increase public goods contributions? *Human Nature*, *18*, 88–108, http://dx.doi.org/10.1007/s12110-007-9012-2.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsbaum, NJ: Lawrence Erlbaum Associates.

Eckel, C. C., & Grossman, P. J. (1998). Are women less selfish than men? Evidence from dictator experiments. *The Economic Journal*, *108*(448), 726–735.

Ekström, M. (2011). Do watching eyes affect charitable giving? Evidence from a field experiment. *Experimental Economics*, *15*, 530–546.

Ernest-Jones, M., Nettle, D., & Bateson, M. (2011). Effects of eye images on everyday cooperative behavior: a field experiment. [Article]. *Evolution and Human Behavior*, *32*(3), 172–178, http://dx.doi.org/10.1016/j.evolhumbehav.2010.10.006.

Fehr, E., & Gachter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137–140.

Fehr, E., & Schneider, F. (2010). Eyes are on us, but nobody cares: are eye cues relevant for strong reciprocity? *Proceedings of the Royal Society B: Biological Sciences*, *277*, 1315–1323, http://dx.doi.org/10.1098/rspb.2009.1900.

Francey, D., & Bergmüller, R. (2012). Images of eyes enhance investments in a real-life public good. *PLoS One*, *7*(5), e37397, http://dx.doi.org/10.1371/journal.pone.0037397.

Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching? Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, *26*, 245–256.

Izuma, K. (2012). The social neuroscience of reputation. *Neuroscience Research*, *72*(4), 283–288, http://dx.doi.org/10.1016/j.neures.2012.01.003.

Keller, J., & Pfattheicher, S. (2011). Vigilant self-regulation, cues of being watched and cooperativeness. [Article]. *European Journal of Personality*, *25*(5), 363–372, http://dx.doi.org/10.1002/per.797.

Koch, A. K., & Normann, H. T. (2008). Giving in dictator games: regard for others or regard by others? *Southern Economic Journal*, *75*(1), 223–231.

Kummerli, R., Burton-Chellew, M. N., Ross-Gillespie, A., & West, S. A. (2010). Resistance to extreme strategies, rather than prosocial preferences, can explain human cooperation in public goods games. [Article]. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(22), 10125–10130, http://dx.doi.org/10.1073/pnas.1000829107.

Kurzban, R. (2001). The social psychophysics of cooperation: nonverbal communication in a public goods game. [Article]. *Journal of Nonverbal Behavior*, *25*(4), 241–259, http://dx.doi.org/10.1023/a:1012563421824.

Milinski, M., Semmann, D., & Krambeck, H. J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature*, *415*, 424–426.

Oda, R., Niwa, Y., Honma, A., & Hiraishi, K. (2011). An eye-like painting enhances the expectation of a good reputation. *Evolution and Human Behavior*, *32*(3), 166–171, http://dx.doi.org/10.1016/j.evolhumbehav.2010.11.002.

Perrett, D. I., Lee, K. J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D. M., & Akamatsu, S. (1998). Effects of sexual dimorphism on facial attractiveness. [Article]. *Nature*, *394*(6696), 884–887, http://dx.doi.org/10.1038/29772.

Powell, K. L., Roberts, G., & Nettle, D. (2012). Eye images increase charitable donations: evidence from an opportunistic field experiment in a supermarket. *Ethology*, http://dx.doi.org/10.1111/eth.12011.

Raihani, N. J., & Bshary, R. (2012). A positive effect of flowers rather than eye images in a large-scale, cross-cultural dictator game. *Proceedings of the Royal Society B: Biological Sciences*, http://dx.doi.org/10.1098/rspb.2012.0758.

Rigdon, M., Ishii, K., Watabe, M., & Kitayama, S. (2009). Minimal social cues in the dictator game. *Journal of Economic Psychology*, *30*, 358–367, http://dx.doi.org/10.1016/j.joep.2009.02.002.

Stephan, C. N., Penton-Voak, I. S., Perrett, D. I., Tiddeman, B. P., Clement, J. G., & Henneberg, M. (2005). Two-dimensional computer generated average human face morphology and facial approximation. In M. Marks, & J. Clement (Eds.), *Computer Graphic Facial Reconstruction*. Amsterdam: Elsevier.

Sylwester, K., & Roberts, G. (2010). Cooperators benefit through reputation-based partner choice in economic games. [Article]. *Biology Letters*, *6*(5), 659–662, http://dx.doi.org/10.1098/rsbl.2010.0209.

Tane, K., & Takezawa, M. (2011). Perception of human face does not induce cooperation in darkness. *Letters on Evolutionary Behavioral Science*, *2*, 24–27.

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research [Article]. *IEEE Computer Graphics and Applications*, *21*(5), 42–50, http://dx.doi.org/10.1109/38.946630.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, *288*(5467), 850–852.