A closer look at the size of the gaze-liking effect: A preregistered replication

Jason Tipples[1]

Anna Pecchinenda[2]

[1]School of Social, Psychological & Communication Sciences

Leeds Beckett University

[2]Department of Psychology, Sapienza University of Rome

Requests for reprints should be addressed to: Jason Tipples, Psychology Group, School of Social, Psychological & Communication Sciences Leeds Beckett University, [CL 815], City Campus, Leeds, LS1 3HE, UK. Tel: +44 (0)113 812 23002 | Email: W.Tipples@leedsbeckett.ac.uk

Abstract

This study is a direct replication of gaze-liking effect (Bayliss, Paul, Cannon, & Tipper, 2006) using the same design, stimuli and procedure. The gaze-liking effect describes the tendency for people to rate objects as more likeable when they have recently seen a person repeatedly gaze toward rather than away from the object. However, as subsequent studies show considerable variability in the size of this effect, we sampled a larger number of participants (N = 98) than the original study (N = 24) to gain a more precise estimate of the gaze-liking effect size. Our results indicate a much smaller standardised effect size (d_z = 0.02) than that of the original study (dz = 0.94). Our smaller effect size was not due to general insensitivity to eye-gaze effects because the same sample showed a clear (d_z = 1.09) gaze-cuing effect - faster reaction times when eyes looked toward vs away from target objects. We discuss the implications of our findings for future studies wishing to study the gaze-liking effect.

A closer look at the size of the gaze-induced object desirability effect: A

preregistered replication

Eye gaze can act as a powerful social cue, triggering automatic shifts in attention (Friesen & Kingstone, 1998; Driver et al., 1999; Langton & Bruce, 1999). In addition to shifting attention, research has shown that observing someone consistently gazing in the direction of a specific object can affect liking for that object. Specifically, one innovative study showed that people rate objects more likeable that they have recently seen a person gaze repeatedly toward rather than away from (Bayliss, Paul, Cannon, & Tipper, 2006). This effect appears to reflect an implicit process (but see; Bry, Treinen, Corneille, & Yzerbyt, 2011) as participants do not report being aware that gaze direction has been deliberately manipulated to look toward certain objects, and away from other objects. Moreover, and in keeping with the idea of a specialized social perception process, the first report showed that the effect is specific to gaze cues – the effect of cue validity on liking was absent for arrow cues (Experiment 2; Bayliss et al., 2006).

Further studies corroborate the initial finding. Specifically, varying attributes of the face that are relevant to social interaction including expression (Bayliss, Frischen, Fenske, & Tipper, 2007) trustworthiness (King, Rowe, & Leonards, 2011; Treinen, Corneille, & Luypaert, 2012) and attractiveness (Strick, van Baaren, Holland, & van Knippenberg, 2009) leads to a modulation of the gaze-induced liking effect. For example, the gaze-induced liking effect is recorded when the observed faces appear smiling but not when they appear disgusted (Bayliss et al., 2007). Also, in keeping with the social nature of the effect, one study (Capozzi, Bayliss, Elena, & Becchio, 2015) found that the liking effect was relatively larger with multiple faces compared to a single face condition (although the latter condition failed to replicate the original effect). Moreover, the specific sequence of gaze shifts between the observer and object matter. Object ratings are higher when the sequence of gaze shifts indicate an intention to communicate: first gaze toward the observer, then gaze toward object and finally, gaze back toward the observer (van der Weiden, Veling, & Aarts, 2010). However, research indicates that the effect of gaze on object is not restricted to human gaze – one study (Corneille, Mauduit, Holland, & Strick, 2008) reported an effect of seeing a dog's head gazing toward or away from an object on an

index of liking. Overall, these results indicate that seeing a person or animal repeatedly gaze toward an object can increase liking for that object.

Replication to quantify the size of the gaze liking effect is needed

Replication of the effect of eye gaze on object liking is needed because studies have also highlighted the considerable variability in the size of the gaze liking effect. The effect for eye gaze in Figure 3 of the original study (Bayliss et al., 2006) showed that the effect was in the predicted direction for 22 out of the 24 participants. This is a large effect (d_z = 0.94) when considered in terms of the effect size benchmarks for dependent means suggested by Lakens (Lakens, 2017). This effect is also slightly larger than estimates of one of psychology's most reliable effects -the Stroop effect (aggregated Cohen's d = 0.88; Ebersole et al., 2016). At the time the gaze-induced liking effect was published there was no precedent for expecting this effect and, given the inherent uncertainty that accompanies small samples (Tversky & Kahneman, 1975) this may not be the best estimate of the effect size. Subsequent studies reported widely varying estimates of the effect size using identical or even smaller sample sizes. Specifically, the significant findings for the gaze-liking effect range from d_z = 0.35 (Ulloa, Marchetti, Taffou, & George, 2015) to 1.22 (for trustworthy faces; King et al., 2011). Moreover, some studies have reported non-significant results in conditions that were expected to show an effect: 1) for gaze cues on ratings for garage objects (Ulloa et al., 2015) and 2) for pointing gestures on ratings for kitchen and garage and other objects (Ulloa et al., 2015). Overall, there is considerable variability in the reported effect sizes and research would benefit from a more precise estimate of the effect size.

Replication design

To estimate the effect size from a larger sample we used the same design, stimuli and procedure of the original study (Bayliss et al., 2006). Participants completed 6 blocks of trials in which they looked at pictures of faces with eyes directed either toward or away from specific garage and kitchen target objects. Participants were required to categorise the objects as either "kitchen" (to indicate that they were usually found in the kitchen) or "garage" (to indicate that they were usually found in the garage). In a 6th final block, the target decision was followed by a rating task in which participants rated how much they liked each target objects.

Bayesian data analyses

We adopted a Bayesian approach to model the data. The benefits of the Bayesian approach have been described in several recent books and articles (Gelman et al., 2014; Kruschke & Liddell, 2017; McElreath, 2016). The Bayesian approach is suited to the goal of the current study –estimating the size of the effect of gaze on liking – because it permits probabilistic statements about effect sizes. For example, we can say that the most likely gaze liking effect size is d_z = 0.35 and that small effect sizes (e.g., d_z < 0.1) are less than 1% probable given the data, model and priors. The Bayesian modelling approach uses Markov Chain Monte Carlo (MCMC) sampling to estimate a range of probable values for model parameters. MCMC requires checks for chain convergence. For brevity, MCMC chain convergence diagnostics are published in the online supplement.

Justification for priors

The Bayesian approach requires specification of priors. Considering the variability in effect size we did not have strong expectations for the effect size. Given our uncertainty about the effect size we choose only weakly informative priors for all our analyses. Weakly informative, rather than uninformative priors were chosen to prevent the MCMC exploring unrealistic parameter values.

Sample size justification

Selecting a sufficiently large sample to replicate the gaze liking effect requires an idea of the original effect size. This is challenging because as stated above, previous studies have reported varying effect sizes. Our solution was to base our sample size on being able to detect the smallest significant effect size reported in the literature; d_z = 0.35 (Ulloa et al., 2015) with a specific level of precision: We wanted our most likely estimates to fall within a region no larger than d_z = .5 (plus or minus d_z = 0.25). Although the study by Ulloa et al was not a direct replication of the original study (it included a different set of faces and additional conditions) the sample size was larger (N=36) than in the original study, making it a reasonable candidate for the smallest effect size of interest.

Simulation

The results of our simulation show that 96 participants are needed to achieve our study goal - detecting an effect size d_z = 0.35 with the 95% most likely estimates falling within a region no larger than d_z = .5 (plus or minus d_z = 0.25). This sample size is based on

simulation using the "BEST" package by (Kruschke & Meredith, 2017). The results indicate that for a sample size N = 96, the 95% most credible values of the effect size d_z = 0.35 fall within a maximum effect size width d_z = .5 on over 95% of the 1000 simulated experiments.

Method

Participants

The final sample consisted of 98 participants; 86 were female (Mean age = 21.2; SD age = 1.82) and 12 were male (Mean age = 21.9; SD age = 14.9). All participants were undergraduates who participated in partial fulfilment of course credits. The experiment was conducted in compliance with institutional ethical guidelines. The study was approved by the Department of Psychology Ethics Committee.

Materials and Apparatus

Targets All the stimuli used in the original study were kindly supplied by the first author (Bayliss, et al., 2006; Bayliss, et al., 2007). Target objects were 36 images of household items: 18 belonging to the "garage" category (i.e., screwdriver), and 18 belonging to the "kitchen" category (i.e., teapot). Two versions of each image were used; one oriented to the left and one to the right. Objects were in four different colours (red, green, yellow, and blue). When presented at a viewing distance of approximately 60 cm, the target-object subtended 6.0° by 2.4° of visual angle.

Faces The face stimuli were 40 unfamiliar faces with neutral or moderately positive facial expressions. The faces varied between 6.3 and 7.6 cm in width and between 8.5 and 11.5 cm in height and were framed in windows measuring 9.8 x 14.0 cm. The faces were presented in full colour on a light grey background.

When presented at a viewing distance of approximately 60 cm, the central face-cues subtended 18.0° by 11.2° of visual angle and object-probes subtended 7.1° by 2.4° of visual angle, whereas the centre-to-centre distance between central face-cues and probes was 12 cm. One each trial, faces were presented first with straight gaze (for 1500 ms), followed by the face-cue with averted gaze (for 500 ms). Three versions of each face were used: 1) one with gaze straight ahead, 2) one with the pupils averted leftwards, and 3) another with the pupils averted rightwards. The eye regions on these faces were as in the original study, and measured approximately 0.5 cm high, varying between 4.0 and 5.0 cm in width. Which face was used as cue on each trial was randomised. The fixation cross will measure 0.8 x 0.8 cm.

E-Prime Version 2.0 Professional software (Schneider, Eschman, & Zuccolotto, 2002) was used for task presentation and data collection. Stimuli were presented on a Pentium IV computer via a 17" LCD monitor (1366 x 768 pixels, 60 Hz). Responses were collected using a standard USB-keyboard with timing error less than 1 ms.

Design and Procedure

Upon arrival to the laboratory, participants completed the informed consent form after which, they sat in front of a computer in a quiet and well-lit room. The position and height of the seat was adjusted to guarantee that all participants have their head and eyes centered on the fixation cross presented on screen. Instructions were presented on screen and participants completed 16 practice trials, followed by 432 trials divided into 6 blocks of 72 trials. Each block consisted of 72 experimental trials with equally probable factorial combinations of eye gaze direction (left, right), probe position (left, right), and probe category (kitchen, garage). In each block, there were 72 trials (36 valid and 36 invalid). Following Bayliss et al., (2006; 2007), although there are four possible colors for the object-probes, for each participant only two colors were used.

Each trial started with a fixation cross (2000 ms) followed by the face cue (500 ms) with eyes gazing to either the left or right. A probe the appeared to the left or right of the face and remained on screen until response or 3000 ms elapsed. The ITI was set to 500 ms. Participants were informed that cue direction was not predictive of probe position and that their task was to indicate as quickly and accurately as possible whether the object used as probe belongs to either the kitchen or garage. Participants responded by pressing the keys "u" and "b" of the keyboard labelled "Kitchen" and "Garage". These keys were chosen to be perpendicular to the left/right probe position. At response, a tone sounded to give feedback on performance (a "bell" for correct and a "buzzer" for incorrect/timeout). Following Bayliss et al. (2006; 2007) in the sixth block the procedure changed slightly and participants were informed that after responding to the probe, the question "How much did you like that object?" would appear at the top of the screen with a 9-point scale, in which 9 indicated "Like very much," and 1 indicated "Don't like at all". Participants rated the object by pressing the selected number (1-9) on the keyboard using their left hand. Each block consisted of 72 experimental trials comprised of equally probable factorial combination of cue direction (left, right), probe position (left, right), and probe category

(kitchen, garage). In total, participants were exposed 6 times (once per block) to 72 object-probes consistently presented in the valid condition, with object colour, object category, and response-key assignment counterbalanced across 12 different versions of the experiment.

Counterbalancing

Assignment of target-object colour pairing to either a valid or invalid cue was counterbalanced across participants. For example, if in one version, yellow objects (half kitchen and half garage objects) were presented with valid cues and blue objects with invalid cues, in another version the opposite assignment was used (blue objects were presented with valid cues and yellow objects presented with invalid cues). Pairing of objects to specific valid or invalid cues was counterbalanced with different versions of the experiment so that objects that were presented with valid cues in one version, appeared with invalid cues in another version of the experiment. Finally, key assignment (U vs B) to probe category (garage vs kitchen) was counterbalanced across different versions of the experiment.

Results

Data preparation and exclusion

Exclusion criteria set out in the pre-registration plan stated that participant's data would be excluded if their mean (or median for the RTs) was either 3 times the interquartile range above the third quartile or 3 times the interquartile range below the first quartile for the sample. Following these criteria, the data of a single participant - who made 40% errors - was removed from all analyses. The raw data, analyses scripts and pre-registration plan can be found on the Open Science Repository https://osf.io/n8e65/
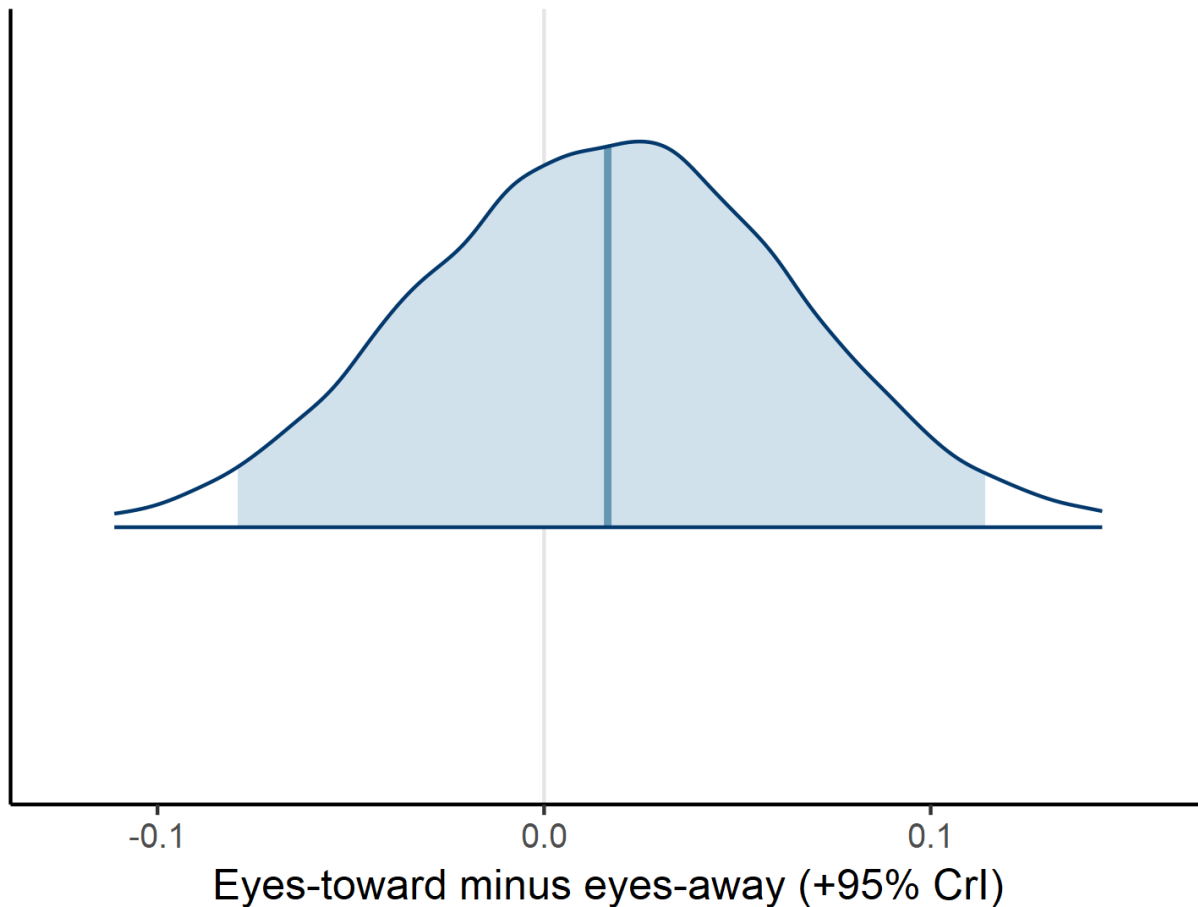
Object ratings

Figure 1: Posterior density distribution of the gaze-liking effect [eyes-toward (valid) minus eyes-away (invalid)]. Shaded region indicates 95% most probable values for the gaze-liking effect.

There was a very small difference indicating increased liking for objects presented in the eyes-toward (*M*=4.752; *SD* = 1.04) compared to objects presented in the eyes-away condition (*M*=4.750; *SD* = 1.03). Bayesian analyses supports the observation that the effect is very small and most likely close to zero. Specifically, the 95% most credible estimates for the gaze-liking effect are shown as the shaded area in the posterior distribution in Figure 1. The line through the middle is the estimated most plausible value (the mean) for the gaze-liking effect. For our sample, the results indicate a very small effect size for a difference (0.01) on a 9-point scale with zero as a probable value within a small (precise) range of

estimated values (95% CrI [-0.08, 0.10]). This difference is less than half a scale point. The standardised effect size (Cohen's d_z) for the effect is 0.02. Identical analyses with target items as a random effect also indicated a small difference (mean difference estimate =0.09) that also included zero as a probable value (95% CrI [-0.01, 0.20], Cohen's d_z = 0.29). Taken together, the results indicate that the gaze-liking effect is very small.

To allow comparison with the analyses conducted in previous research we also conducted a one-sample, one-tailed t-test against zero for the gaze-liking effect. The test was not significant either by-subjects, $t(97) = 0.03$, $p = 0.48$ or by-items, $t(35) = 1.62$, $p = 0.11$[1]

---

[1] Studies wishing to plan sample size based on power analyses based for this effect size should note that such analyses will provide an estimate of the number of items needed to achieve a given level of power. For the frequentist t-test reported here with items as the sole random effect, a future study wishing to achieve power = .80 with alpha = 0.05, would require 93 items (93 target objects). To produce an idea of the number of participants it is necessary to model the effect with both items and participants as random effects within a single model. To do this we used a multilevel model with mean item ratings regressed onto condition (eyes-toward, eyes-away) with random effects for both participants and items. The estimated effect size for condition (eyes-toward, eyes-away) for this analysis was smaller (Cohen's d_z =0.006) than the estimate from the model with participants as a single random effect.

Reaction times

Gaze cuing RT effect - posterior distribution

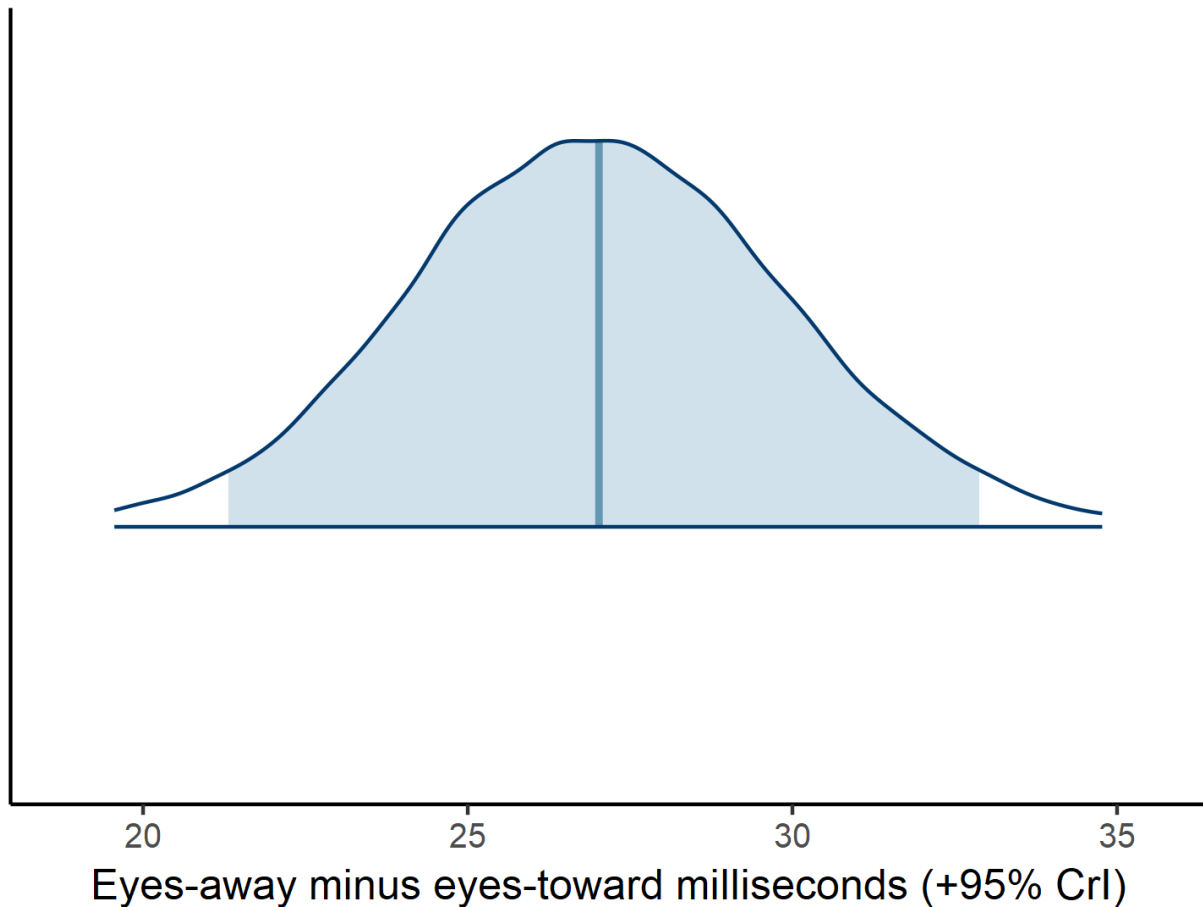Eyes-away minus eyes-toward milliseconds (+95% CrI)

Figure 2. Posterior density distribution of the gaze-cueing effect [eyes-toward (valid) minus eyes-away (invalid)]. Shaded region indicates 95% most probable values for the gaze-cuing effect.

The results are in keeping with the RT analyses of the gaze-cuing RT effect in previous studies and studies of the gaze-cuing effect more broadly. Specifically, median reaction times were shorter for targets in the eyes-toward (M =678, SD = 78) compared to the eyes away (M =706, SD = 82) condition. Bayesian analyses supports this observation. As shown in Figure 2, the 95% most credible values for the eyes-away minus eyes-toward difference ranged from 21 to 32 milliseconds with 27 milliseconds as the most probable value (Cohen's d_z = 1.09). A one-sample, one-tailed t-test for the gaze-cuing effect was significant, $t(97) = 7.47$, $p < 0.00001$.

Accuracy

The mean percentage correct (96%) was similar to that reported in the original study (Bayliss et al; 95.8%). For blocks 1 to 5 accuracy rates were at ceiling for both the eyes-toward (*M* =96%, *SD* = 0.03) and eyes-away (*M* =95%, *SD* =0.03) conditions. For the rating blocks, percentage accuracy was also at ceiling (M = 97%; SD= 0.04). The high accuracy rates were not normally distributed and therefore, we applied an arc sine transformation to the mean proportion correct before statistical testing. A Bayesian one-sample t-tests indicated that zero was a probable value for of eyes-away, eyes-toward difference for blocks 1 to 5, (95% CrI [-0.009 , 0.016]). The accuracy rates were very high for the rating block and consequently, the model failed to converge. To address this issue, we modelled the number of errors in the rating block as a zero-inflated Poison-distributed random variable. Distributional analyses indicated an adequate fit for this type of distribution. We regressed the number of errors (mostly zero) onto condition (eyes-toward, eyes-away) in a multilevel Bayesian regression model that included varying-intercepts for participants. Weakly regularising priors were used for all effects. The model converged. The analyses indicate that zero was a probable value for the eyes-toward, eyes-away difference, β = -0.07 (95% CrI [-0.41, 0.27]). Similarly, one-sample, two-tailed t-tests for the eyes-away, eyes-toward difference were not significant for either blocks 1 to 5, *t*(97) = 0.54, *p* = 0.58 or the final rating block, *t*(97) = -0.26, *p* = 0.79.

Discussion

The aim of this registered replication was to estimate the size of the gaze-liking effect - the effect whereby people rate objects as more likeable when they have recently seen a person repeatedly gaze toward rather than away from the object (Bayliss, et al., 2006). Using a relatively large sample size (N= 98) our results indicate that the effect is very small. This has practical implications for future studies. Based on our estimated effect size, future researchers planning to conduct frequentist NHST aiming to achieve power =0.80 and alpha = .05 (one-tailed test) will require at least 3000 participants to detect the gaze-liking effect. In contrast, also based on our analyses, researchers wishing to replicate the gaze-cuing effect using the same method used here with the same levels of desired power will need 15 participants. The fact that we recorded a clear gaze-cueing effect indicates that small effect size estimate for the gaze-liking effect was not due to our

participants being generally insensitive to the direction of eye-gaze in experimental manipulations.

Our study included 2 differences that were not noted at the pre-registration stage. Specifically, in the original study each participant saw a variety of red, yellow, blue, and green stimuli that could serve as either cued or uncued objects. For example, participants in the original study might have seen eye gaze toward a blue screwdriver and, toward a yellow iron and away from a red hairdryer. Our design allows to rule out any likely effect of seeing a specific colour with a valid or invalid cue but it reduces the number of colours experienced by each participant to 2 and it does differ from the likely experience of participants in the original study. Therefore, even though the current study includes a valid manipulation of the eye gaze direction to target relationship — our study is a clear test of the original hypothesis — the difference in the randomisation of colour could have affected the results. Although, it is worth noting that another study (Ulloa et al, 2015) used a consistent mapping of colour to eye-gaze validity and recorded the gaze-liking effect. Considering this, we think it unlikely that this difference in procedure would have led to the large decreased in effect size for the current (statistically powerful) study.

A second difference between the current and original study (Bayliss et al., 2006) is that participants in the original study verbally stated their rating to the experimenter, whereas in the present study participants entered their ratings on a keyboard. Our study may have increased working memory load relative to the original study and this may have affected the results. However, we note that in another study (Manera, Elena, Bayliss & Becchio, 2014) that recorded the gaze-liking effect, participants also entered their ratings on a keyboard. In short, we do not think these differences are likely to have produced the large reduction in effect size between the original study and current research.

What then might have been the reason for the large decrease in reported effect size? One possible reason for the decline in the effect size is the file drawer effect (Rosenthal, 1979) – the tendency to publish only positive (significant) results and not publish non-significant or non-confirmatory findings. Since the publication of the original study in 2006 there has been a marked increased focus on power analyses, transparency, statistical techniques, study pre-registration and data sharing. Together these methods seem likely to reduce file drawer and associated research problems.

The results do not rule out the possibility that observing averted eye gaze has a larger effect on object liking. It is possible that the original task is not very sensitive and that a different task might reveal a larger effect for both eye gaze and arrow cues. Alternatively, the current task is sensitive but uncontrolled differences between the current and original study might explain the difference in effect size. Although we think the latter unlikely, it is good that future studies with clear hypotheses about these uncontrolled differences are now able to describe and test these by using a pre-registration protocol.

Reference

Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). *Affective evaluations of objects are influenced by observed gaze direction and emotional expression*. *Cognition* (Vol. 104). https://doi.org/10.1016/j.cognition.2006.07.012

Bayliss, A. P., Paul, M. A., Cannon, P. R., & Tipper, S. P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. *Psychonomic Bulletin & Review*, *13*(6), 1061–1066. https://doi.org/10.3758/BF03213926

Bry, C., Treinen, E., Corneille, O., & Yzerbyt, V. (2011). *Eye'm lovin' it! The role of gazing awareness in mimetic desires. Journal of Experimental Social Psychology* (Vol. 47). https://doi.org/10.1016/j.jesp.2011.03.023

Capozzi, F., Bayliss, A. P., Elena, M. R., & Becchio, C. (2015). One is not enough: Group size modulates social gaze-induced object desirability effects. *Psychonomic Bulletin & Review*, *22*(3), 850–855. https://doi.org/10.3758/s13423-014-0717-z

Corneille, O., Mauduit, S., Holland, R. W., & Strick, M. (2008). Liking products by the head of a dog: Perceived orientation of attention induces valence acquisition. *Journal of Experimental Social Psychology*, *45*, 234–237. https://doi.org/10.1016/j.jesp.2008.07.004 Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze Perception Triggers Reflexive Visuospatial Orienting. *Visual Cognition*, *6*(5), 509–540. https://doi.org/10.1080/135062899394920

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., … others. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, *67*, 68–82.Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, *5*(3), 490–495. https://doi.org/10.3758/BF03208827

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2014). *Bayesian Data Analysis, Third Edition (Chapman & {Hall/CRC} Texts in Statistical Science)*. Chapman and Hall/CRC.

King, D., Rowe, A., & Leonards, U. (2011). I Trust You; Hence I Like the Things You Look at: Gaze Cueing and Sender Trustworthiness Influence Object Evaluation. *Social

Cognition, *29*(4), 476–485. https://doi.org/10.1521/soco.2011.29.4.476

Kruschke, J. K., & Liddell, T. M. (2017). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, 1–29. https://doi.org/10.3758/s13423-016-1221-4

Lakens, D. (2017). Equivalence Tests. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177

Langton, S. R. H., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*(5), 541–567. https://doi.org/10.1080/135062899394939Manera, V., Elena, M. R., Bayliss, A. P., & Becchio, C. (2014). When seeing is more than looking: Intentional gaze modulates object desirability. Emotion, 14(4), 824–832. http://doi.org/10.1037/a0036258

McElreath, R. (2016). *Statistical rethinking: a Bayesian course with examples in R and Stan*.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin, 86*(3), 638-641. http://dx.doi.org/10.1037/0033-2909.86.3.638

Strick, M., van Baaren, R. B., Holland, R. W., & van Knippenberg, A. (2009). Humor in advertisements enhances product liking by mere association. *Journal of Experimental Psychology: Applied*, *15*(1), 35–45. https://doi.org/10.1037/a0014812

Treinen, E., Corneille, O., & Luypaert, G. (2012). *L-eye to me: The combined role of Need for Cognition and facial trustworthiness in mimetic desires*. *Cognition* (Vol. 122). https://doi.org/10.1016/j.cognition.2011.10.006

Tversky, A., & Kahneman, D. (1975). Judgment under Uncertainty: Heuristics and Biases. In *Utility, Probability, and Human Decision Making* (pp. 141–162). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-010-1834-0_8

Ulloa, J. L., Marchetti, C., Taffou, M., & George, N. (2015). Only your eyes tell me what you like: Exploring the liking effect induced by other's gaze. *Cognition and Emotion*, *29*(3), 460–470. https://doi.org/10.1080/02699931.2014.919899

van der Weiden, A., Veling, H., & Aarts, H. (2010). When observing gaze shifts of others enhances object desirability. *Emotion (Washington, D.C.)*, *10*(6), 939–43. https://doi.org/10.1037/a0020501

Figure Captions

Figure 1 Posterior density distribution of the gaze-liking effect [eyes-toward (valid) minus eyes-away (invalid)]. Shaded region indicates 95% most probable values for the gaze-liking effect.

Figure 2. Posterior density distribution of the gaze-cueing effect [eyes-toward (valid) minus eyes-away (invalid)]. Shaded region indicates 95% most probable values for the gaze-cuing effect.

Figure 1

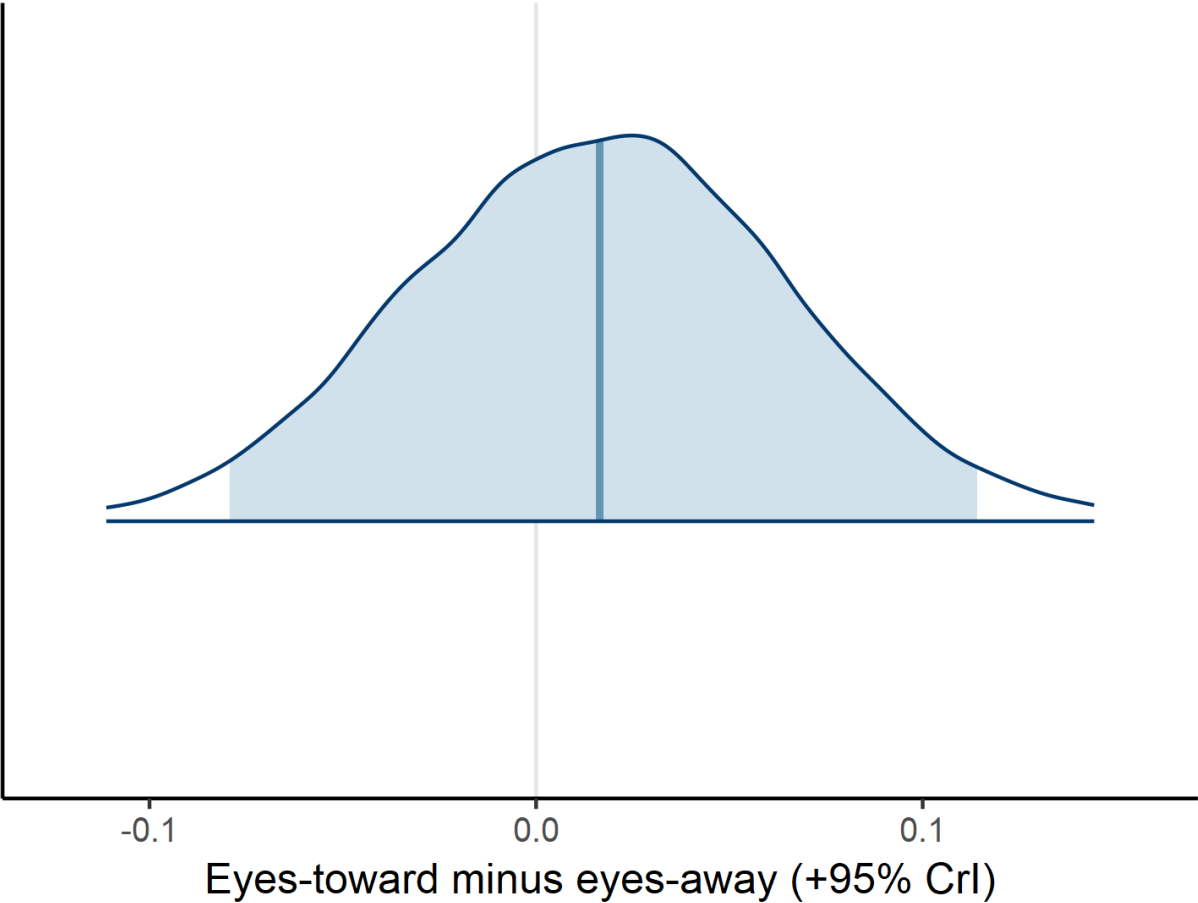Gaze liking rating effect - posterior distribution

Eyes-toward minus eyes-away (+95% CrI)

Figure 2

Gaze cuing RT effect - posterior distribution

Eyes-away minus eyes-toward milliseconds (+95% CrI)