Evolution and Human Behavior 38 (2017) 561–571

Contents lists available at ScienceDirect

Evolution and Human Behavior

journal homepage: www.ehbonline.org

Original Article

Effect of artificial surveillance cues on reported moral judgment: Experimental failures to replicate and two meta-analyses

Stefanie B. Northover [a,b,*], William C. Pedersen [c], Adam B. Cohen [b], Paul W. Andrews [a]

[a] Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada
[b] Department of Psychology, Arizona State University, 651 E. University Drive, P.O. Box 871104, Tempe, AZ 85287, USA
[c] Department of Psychology, California State University, Long Beach, 1250 Bellflower Boulevard, Long Beach, CA 90840, USA

ARTICLE INFO

Article history:
Initial receipt 3 August 2016
Final revision received 20 December 2016

Keywords:
Surveillance cues
Cues of being watched
Observation cues
Eyespots
Moral judgment
Meta-analysis

ABSTRACT

Several papers have reported that artificial surveillance cues, such as images of watching eyes, cause anonymous participants to behave as if they are actually under surveillance, thus increasing moral behavior. In a series of four experiments, we found no evidence that artificial surveillance cues impact reported moral judgment, self-rated possession of positive traits, or religiosity. Two small meta-analyses, both comprising six experiments investigating the effect of artificial surveillance cues on moral judgment, provided mixed conclusions. One meta-analysis produced a mean effect size not significantly different from zero and the other produced a mean effect size on the edge of significance. On the whole, artificial surveillance cues have inconsistent effects, or possibly no effect, on moral outcomes.

1. Introduction

When people are under surveillance, they tend to behave more prosocially than they otherwise would (Kurzban, 2001; Kurzban, DeScioli, & O'Brien, 2007; Piazza & Bering, 2008; Satow, 1975; van Rompay, Vonk, & Fransen, 2009). Even artificial cues of surveillance, such as stylized images of eyes, have apparently increased prosocial behavior in lab and field experiments (e.g., Haley & Fessler, 2005; Pfattheicher & Keller, 2015). Participants seemingly behave like they are being watched when they are exposed to artificial cues of being watched, even though participants are consciously aware that they are not actually being watched. Outcomes in such experiments have included donating to charity (Pfattheicher, 2015), hand washing (Carbon & Hesslinger, 2011), and picking up litter (Ernest-Jones, Nettle, & Bateson, 2011).

However, surveillance cue effects sometimes fail to replicate (Carbon & Hesslinger, 2011; Matsugasaki, Tsukamoto, & Ohtsubo, 2015). Two meta-analyses investigating the impact of artificial surveillance cues on generosity produced small mean effect sizes that were not significantly different from zero (Northover, Pedersen, Cohen, & Andrews, 2017). Although artificial surveillance cues may not impact generosity, more work should be done to investigate additional behavioral outcomes. The goal of the present paper is to investigate the effect of artificial surveillance cues on moral judgment, an outcome sufficiently different from generosity to warrant separate consideration.

Our primary measure of moral judgment is that used by Bourrat, Baumard, and McKay (2011), who asked participants to rate the moral acceptability of two misdeeds: returning a lost wallet but keeping the money, and falsifying information on a résumé (Schnall, Haidt, Clore, & Jordan, 2008). Participants who were exposed to an image of watching eyes rated the transgressions more harshly than participants exposed to an image of flowers, suggesting that the artificial surveillance cue caused the participants to respond like they were truly under surveillance.

In many cases, reported surveillance cue effects are seemingly conditional on features of the environment, qualities of the surveillance cue, participant traits, or methods of data analysis (Northover et al., 2017). Although many moderating variables have been proposed, findings are inconsistent. One potential moderator is the masculinity or femininity of the surveillance cue. In a field experiment conducted by Bateson et al. (2006), anonymous people contributed more money to an honesty box, used to collect funds for coffee, when masculine eyes were displayed than when feminine eyes were displayed (but see Carbon & Hesslinger, 2011). Matland and Murray (2015) also found a greater

* Corresponding author at: Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada.
E-mail addresses: stefanie.northover@asu.edu (S.B. Northover), bill.pedersen@csulb.edu (W.C. Pedersen), adamcohen@asu.edu (A.B. Cohen), pandrew@mcmaster.ca (P.W. Andrews).

http://dx.doi.org/10.1016/j.evolhumbehav.2016.12.003
1090-5138/

effect from masculine eyes. However, other studies found no significant difference between masculine and feminine eye cues (Nettle et al., 2013; Panagopoulos, 2014).

Another potential moderator is the familiarity of the surveillance cue. A familiar face may induce the feeling of being watched by a member of the community. People are more likely to behave prosocially in less densely-populated areas (Korte & Kerr, 1975; Levine, Martinez, Brase, & Sorenson, 1994; Rushton, 1978; Yousif & Korte, 1995). This may be because the less dense the population of a community, the easier it is to build and maintain a reputation. Therefore, people should behave more prosocially when they are observed by familiar individuals than they do when they are observed by unfamiliar individuals.

We conducted four experiments investigating artificial surveillance cue effects. Initially, we were interested in multiple dependent variables – religiosity, positive traits, and moral judgment (Experiment 1). However, we turned our attention to a single dependent variable – moral judgment – when we were unable to conceptually replicate Bourrat et al.'s (2011) moral judgment results. Experiments 2, 3, and 4 were increasingly precise replications of Bourrat and colleagues. We also investigated the femininity/masculinity and the familiarity of the surveillance cues as possible moderators. In addition, we evaluated several moderating variables in post hoc fashion after multiple experiments failed to replicate the findings of Bourrat et al. These variables included the length of surveillance cue exposure, the location of the surveillance cue, whether the experimenters drew attention to the surveillance cue, and the location of the experiment. None of our experiments resulted in significant surveillance cue effects.[1] In addition to our experiments, we conducted small meta-analyses of the six studies which investigated the effect of surveillance cues on Bourrat and colleagues' moral judgment task.

2. Experiment 1

2.1. Surveillance cue traits

In addition to exploring effects of surveillance cues generally, we investigated different attributes of surveillance cues: familiarity and masculinity/femininity.

2.2. Dependent measures

In Experiment 1, we investigated two dependent measures in addition to moral judgment: self-rated possession of positive traits and religiosity. If surveillance increases the likelihood of reputation-boosting behavior, then any traits that are desirable in social exchanges may be displayed or exaggerated. Thus, a watched individual may behave in a way that implies the possession of positive traits such as kindness, honesty, generosity, or reliability.

Additionally, people who are being watched may wish to appear religious. Religion tends to be associated with morality and trustworthiness (Edgell, Gerteis, & Hartmann, 2006; Farkas, Johnson, Foleno, Duffett, & Foley, 2001; Hall, Cohen, Meyer, Varley & Brewer, 2015; Tan & Vogel, 2008), whereas atheists tend to be viewed as untrustworthy (Gervais, Shariff, & Norenzayan, 2011) and incite negative feelings in others (Pew Research Center, 2014). Although the Canadian province of Ontario, from which our sample came, is not a particularly religious region (23.14% of people claimed no religious affiliation in a 2011 census; Statistics Canada, 2013), a meta-analysis conducted by Sedikides and Gebauer (2010) showed a significant positive correlation between intrinsic religiosity and socially desirable responding among Canadians. The authors proposed that this relationship exists because religiosity can be used by people to self-enhance. If the authors are correct, their findings suggest that religiosity is valued by Canadian culture. Therefore, Canadian participants may exaggerate their religiosity when they feel like they are being watched.

2.3. Method

2.3.1. Participants

We recruited 338 psychology students from McMaster University, located in southern Ontario. Participants were given course credit for their participation. The mean age of the participants was 19.1 years; there were 83 men, 253 women, and 2 of unreported gender; about 50% were White, 40% Asian, 6.5% Middle Eastern, and 5% indicated some other ethnicity.

2.3.2. Procedure

Each participant was seated alone in a small room with the door closed, isolated from other people to provide for anonymity and privacy. The participants' task was to complete a computer questionnaire made up of three parts designed to measure religiosity, self-rated possession of positive traits, and moral judgment. The computer screen was split into two frames. The left frame contained the questionnaire, which was administered through LimeSurvey (www.limesurvey.org). The contents of the right frame depended on which of four conditions the participant had been randomly assigned to – familiar face, unfamiliar face, chair (an image control condition), or no image (blank screen). For the familiar face condition, the image was of a celebrity's face. For the unfamiliar face condition, the image was of the face of a person who was not well known in North America. For the chair condition, the image was of a chair on a white background.

The cover story told to the participants was, "We're studying simultaneous processing of various types of visual stimuli. All conditions will have words. Some conditions will also have images. Some conditions will *not* have images. At the end of the experiment, you'll be asked questions about any images you see if you have them, so *please pay careful attention to them*."

To ensure the experimenters were blind to condition, the experimenters clicked a button on the computer screen as soon as they were finished giving directions to each participant. Immediately after clicking the button, the experimenters left the experiment room. Clicking the button started a ten second countdown, then the right frame loaded either an image (familiar face, unfamiliar face, or chair conditions) or a blank page (no image condition).

At the end of the experiment, participants were probed for suspicion. Data were removed for those who correctly guessed the purpose of the experiment.[2]

2.3.3. Stimuli

Participants in the familiar face, unfamiliar face, and chair conditions were presented with images. Six different images were used for each of these conditions. Each participant in these three conditions was shown just one of the images. We used a monitor with a viewable image size of 59.69 cm and a screen resolution of 1920 by 1080 pixels.

The individuals chosen for the familiar face condition were Kristen Stewart, Rihanna, Taylor Swift, Barack Obama, Danielle Radcliffe, and Tom Hanks. The individuals selected for the unfamiliar face condition were mostly models or celebrities from outside North America, chosen because their images were similar in style and attractiveness to those in the familiar face condition. We attempted to match the familiar and unfamiliar faces on gender, approximate age, and ethnicity. Half of the faces were male and half were female, so we were able to investigate the dependent measures according to the masculinity/femininity of the surveillance cues. Each face image had an interpupillary distance of 115 or 116 pixels. All faces were aligned so there was no head tilt.

[1] Data from all four experiments are available at the first author's website.

[2] Unfortunately, we do not know the exact number of participants whose data were removed, as these records were lost; our best estimate is 4 or 5. These data were removed before any data analysis.