THE SENSE OF BEING STARED AT: ANALYSIS OF PREVIOUS DATA AND A PILOT REPLICATION

by DEAN RADIN

ABSTRACT

Data collected in the majority of 'the sense of being stared at' experiments indicate that people tend to respond correctly when being stared at, and at chance when not being stared at. After statistical adjustment for overall response bias (the tendency to respond 'yes'), this asymmetry disappears. But such an analysis assumes that response biases are constant across all trials. The question arises as to whether response biases may differ during stare and no-stare conditions. In a small-scale replication study with feedback after each trial ($N = 625$ trials), hit rates similar to those reported by Sheldrake were obtained, and a nearly significant difference in response biases was observed between stare and no-stare trials ($p = 0.06$). This asymmetry was observed primarily in the first six trials in 20-trial runs, arguing against a subliminal cuing hypothesis.

INTRODUCTION

In an experiment on the 'sense of being stared at', for ease of exposition let us call the person performing the staring *Jack* and the person being stared at *Jill.* Jack and Jill sit within a few meters of each other, Jill with her back to Jack. Jack follows a randomized schedule which determines on each successive trial whether he should stare or not stare at the back of Jill's head. Cued with a clicking tone, Jill responds 'yes' if she believes that Jack is staring at her, or 'no' if she does not. The result of each such trial falls into into one of four categories: hit, miss, false alarm, and correct rejection, as shown in Table 1.

Table 1

Types of Trial Outcome in a Staring Experiment

		Response	
		Yes	*No*
condition	*staring*	hit	miss
	not staring	false alarm	correct rejection

Sheldrake (1998, 1999, 2000, 2001, 2003) reported a series of experiments based on this design, some involving trial-by-trial feedback under casual conditions, such as tests conducted by pairs of children in classrooms, and others involving blindfolds, no feedback, or more secure conditions such as having Jack stare at Jill through a window. From tables in these published reports, plus those in Coover (1913) and Poortman (1959), I was able to collect a total of 33,357 trials for analysis. The data included breakdowns of the raw numbers of hits, misses, correct rejections and false alarms in each experiment.

The overall success rate in this database, i.e. the combined proportion of hits and correct rejections, was 54.5%. On average, this would require only one additional hit over the chance-expected 10 hits in a typical run of 20 trials, but given the statistical power afforded by 33,357 trials, the observed success rate excludes chance as an explanation ($z = 16.3$, effect size per trial = $z/\sqrt{(N)} = 0.09$, $p << 10^{-10}$).

Sheldrake (2003) has reported a consistent pattern of responses in these tests, as shown in Figure 1. Jill is able to tell when Jack is staring (57.8% hit rate), but not when he is not staring (48.9% false alarm rate). Sheldrake suggests that this pattern makes sense because Jill could respond to the positive act of staring, but could only guess during the absence of staring, which "has no parallel in real-life conditions" (Sheldrake, 2003, p. 174).[1]

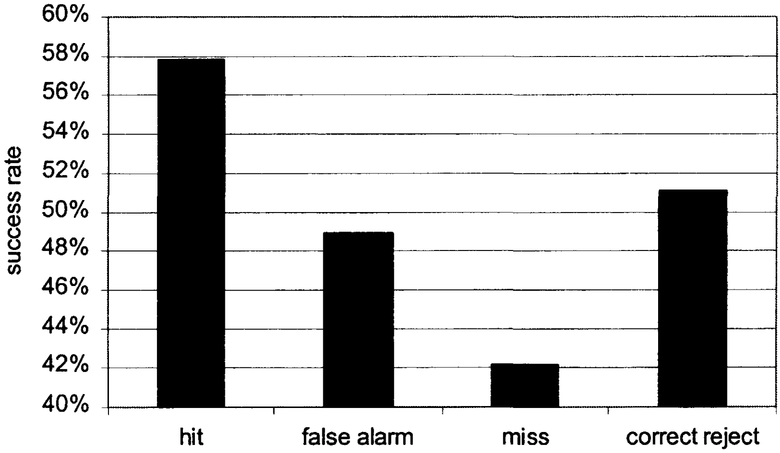

Figure 1. Observed outcome percentages in previous staring experiments, N = 33,357 trials.

In interpreting each of these four types of outcomes, they should be adjusted for Jill's response biases (Schmidt, 2001; Schmidt, Müller & Walach, 2003). That is, while stare and no-stare conditions are typically distributed uniformly at random or randomly counterbalanced to preclude Jill from inferring the next target,[2] Jill cannot be expected to respond at random. In fact, overall she said 'yes' 53.4% of the time in these experiments. That bias would inflate the statistical significance associated with the success rate for hits and decrease the success rate for correct rejections.

A simple way to assess the statistical significance of the hit rates shown in Figure 1, adjusted for response bias, is to use the formula for a z-score of a difference in proportions:–

$$z = (p_1 - p_0) / \sqrt{p_0 q_0 / N}$$

---

[1] Of course, if the act of staring is associated with differential sensory cues, e.g. subtle changes in breathing when Jack is staring at Jill, or in infra-red radiation depending on how Jack's face is oriented towards Jill, etc., then the asymmetrical pattern confounds a genuine staring effect with sensory cues.

[2] These conditions were distributed at rates of 50.1% staring vs. 49.9% no-staring within this database.

where p_1 is the observed hit rate, p_0 is the chance-expected hit rate, $q_0 = 1-p_0$, and N is the number of trials under consideration. If p_0 is taken as 0.5 and N = number of staring trials, then the overall 'raw' z-score is $z(\text{hit}) = 19.9$. But if p_0 is Jill's observed response bias of 53.4%, then the adjusted z-score is $z(\text{hit}) = 11.5$. This is still far from chance, but it is also significantly smaller than the unadjusted z-score (see Figure 2). Also, notice that the asymmetrical 'raw score' hit rates now become symmetrical.

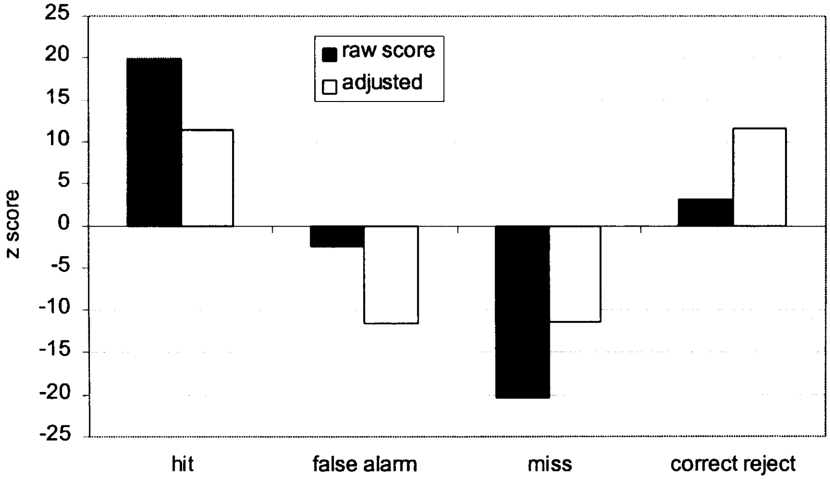

Figure 2. Raw and adjusted z-scores for the four types of outcome.

To explore the subliminal cuing hypothesis in more detail, I partitioned Sheldrake's data into sets with progressively stronger controls for sensory leakage. Of the 33,357 trials, some 21,168 were collected with the pairs in close proximity—within two or three meters—and with trial-by-trial feedback; 5,580 trials were collected in close proximity and no feedback; and 4,800 were collected with the pair separated by a window.

Figure 3 shows that study outcomes did decline as testing conditions shielded increasingly against artifacts from potential sensory cues. However, the effects did not decline to zero, nor do they differ significantly from one another, as indicated by the error bars. In addition, evidence from an independent class of studies employing rigorous shielding against sensory cues, no feedback, and unconscious physiological measurements instead of conscious responses supports the idea that there is a genuine sense of being stared at (Schmidt, Schneider, Utts & Walach, 2004).

While the above results suggest that the staring effect may be robust, recent replication attempts have not shown such dramatic results. Lobach and Bierman (2004) reported three unsuccessful attempts to replicate and concluded that results of the earlier studies may be explained as a combination of response biases and response strategies. They further suggested that failures to replicate may also involve differences in participant population, type of experimental setting, and rigor of the experimental procedures.

Another limitation of most of the previous staring studies is that the outcome of each trial was recorded manually, and this is known to be vulnerable to both unintentional and motivated mistakes. In addition, the possibility of a "psi-mediated response bias" (Lobach & Bierman, 2004) has not been explicitly explored. To study these issues, I conducted a small-scale pilot study using an automated, computer-based method.

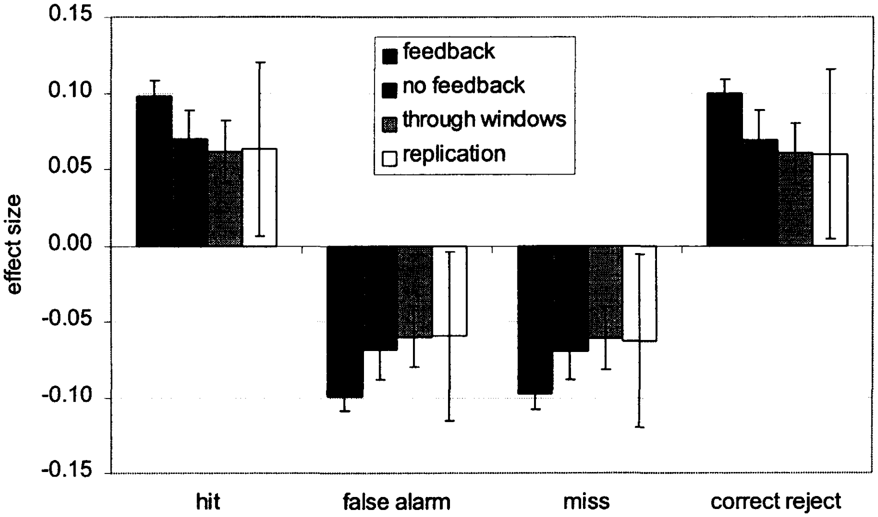

Figure 3. Adjusted effect sizes ($e = z / \sqrt{N}$) and one-standard-error bars, for three testing conditions described in previously reported studies, and in the pilot replication reported here.

METHOD

A person being stared at (Jill) and a person assigned to stare (Jack) sat two meters apart, with Jill's back to Jack, and Jack sitting in front of a laptop computer. Jill wore a blindfold to help block visual cues, and in her hands she held a gamepad (Microsoft's 'Sidewinder USB', a computer peripheral used to control video games). This gamepad provides a left- and right-hand trigger button used by the two index fingers, and several buttons on top of the gamepad designed to be used by the right thumb.

Jill initiated each trial by pressing a button on the game-pad, whereupon a computer-synthesized voice announced "Prepare for trial #", where # was the current trial number. As the same time, the words "Stare" or "No Stare" silently appeared on Jack's laptop screen. Jack followed this randomly assigned instruction either by gazing intently at the back of Jill's neck, or by closing his eyes and thinking about something else. Five seconds later, the computer sounded a click tone. This signaled Jill to respond at will by pressing her right index finger button to indicate her guess, "I'm being stared at", or her left index finger button to indicate, "I'm not being stared at". After Jill responded, the computer provided feedback by speaking one of four phrases:

"Stare, correct" if Jack was staring and Jill's response was staring; or similarly "Stare, incorrect", "No stare, correct", or "No stare, incorrect", depending on the outcome. This sequence constituted one trial, and one run consisted of 20 such trials. Jack and Jill listened to the computer's prompts over headphones to help block subliminal audio cues.

The sequence of trials was determined pseudo-randomly, with $p(\text{stare}) = p(\text{no-stare}) = 0.5$, thus the number of stare and no-stare trials in a given run would be expected to average around 10 each, but there was no guarantee of an identical distribution. The Microsoft Visual Basic 6 pseudo-random function was used to make these random selections; the algorithm was seeded by the computer system's clock time when the program began.

RESULTS

This experiment was conducted as a pilot test without a preassigned number of trials. Participants were recruited opportunistically from among visitors to the Institute of Noetic Sciences (IONS). A total of 12 pairs of people participated in the experiment. Of these, five pairs were children ranging in age from 8 to 14, and the remainder were adults over 18. All sessions were supervised by the author. To avoid problems associated with selective data reporting, all collected data are reported, including one partially-completed run of 5 trials. Those trials were contributed by a pair of children who decided to end the run prematurely.

Thirty-one sessions of 20 trials plus the abbreviated session were collected for a total of 625 trials. Figure 4 shows that the results in the four outcome categories were in agreement with Sheldrake's reported results; hits were over chance expectation (56.9%) and false alarms were near chance (50.8%). Of the 625 trials in the pilot test, 331 successes (hits plus correct rejections) were obtained, or 53%, $z = 1.48$, $p = 0.07$. The overall response bias (responding 'yes' regardless of trial type) was 53.8%, significantly above chance expectation ($p = 0.03$, one-tailed).

Figure 5 shows the z-score associated with the success rate per trial, and cumulatively across all trials. Each of the first six trials was above chance expectation, cumulating to about 3 standard errors over chance at trial 6 and then declining to a terminal z-score of 1.48. This 'serial position effect' pattern is reminiscent of decline effects often observed in forced-choice tests. The average success rate for trials 7 through 20 was close to chance expectation.

To see whether there might have been patterns in the target sequences that Jill could learn and thus bias her responses accordingly, an autocorrelation analysis was conducted through lag 20 (i.e. comparison of all pairs of adjacent targets, then pairs of targets two steps apart, three steps apart, etc.). None of the resulting correlations was statistically significant, so there were no clues available in the target sequences. A similar analysis of the sequence of Jill's responses showed a significant negative correlation for lag +1, $r = -0.14$, $z = -3.89$, $p = 0.0001$. This is due to a known sequential response bias in which participants tend to avoid giving the same response twice in a row. Lag +7 was also positive, $r = 0.10$, $z = 2.80$, $p = 0.003$, but given the 20 multiple tests this is likely to be due to chance.

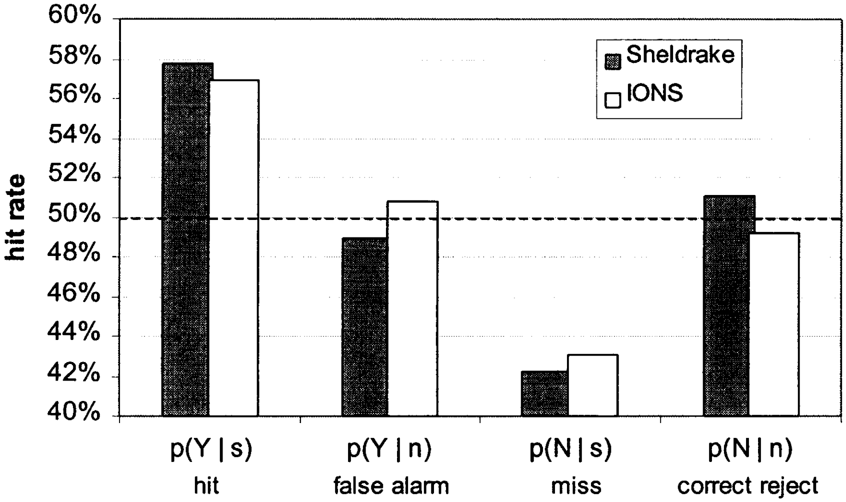

Figure 4. Hit rates in previous studies and the IONS pilot replication in the four possible outcome categories. p(Y | s) refers to the probability of saying 'yes' when 'staring' is the target.

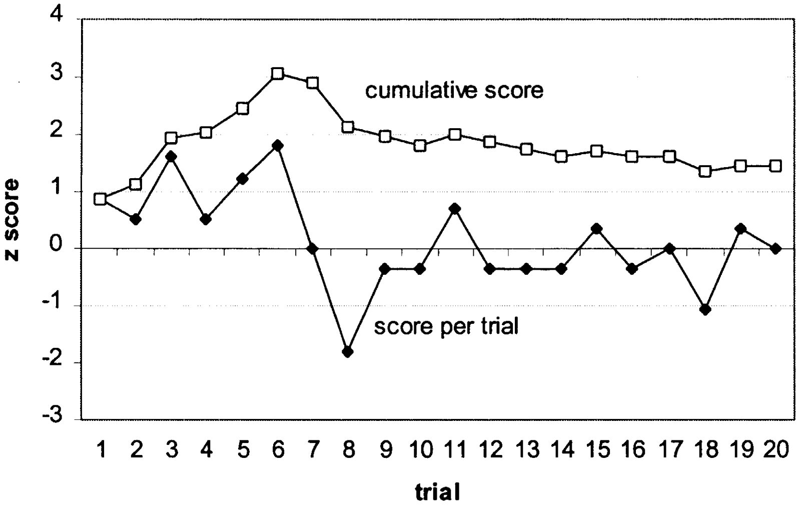

Figure 5. z-score associated with success rate per trial and cumulatively.

DISCUSSION

This replication attempt produced a similar pattern of responses and a marginally significant overall success rate consistent with the results reported by Sheldrake. The design did not absolutely exclude subliminal sensory cues, but other proposed artifacts, such as patterns in the sequence of stare/no-stare conditions and selective data reporting (e.g. Colwell et al, 2000; Marks & Colwell, 2000) are inadequate explanations.

What about the effect of response biases? The statistical adjustment for Jill's response bias, as shown in Figure 3, is based upon the assumption that her response biases were distributed equally across the stare and no-stare conditions. But note that responding 'yes' when the trial condition is stare is identical to what I've called a 'hit', and responding 'yes' when the target is no-stare is what I've called a 'false alarm'. Van Bolhuis (no date) has also noted this equivalency. Assuming that the chance response bias is 50%, then the response bias observed in Jill's stare trials was significantly above chance, $z = 2.41$, $p = 0.008$, and the response bias in no-stare trials was close to chance, $z = 0.28$, $p = 0.39$. The difference between the two proportions is associated with $z = 1.54$, $p = 0.06$, modestly supporting Sheldrake's claim that Jill behaved differently when Jack was staring at her compared with when he was not.

The present experiment lends further support to Sheldrake's results because the computer-based design obviated motivated or inadvertent recording errors, and success rates in the first few trials of the average 20-trial run were better than in later trials. This latter finding is contrary to expectations about cuing artifacts, which would presumably have resulted in progressively increasing success rates over the course of runs of 20 trials. A similar lack of evidence for improved scores was observed by Sheldrake (in press).

In future experiments, to help eliminate the potential for subliminal cues, an inexpensive approach that would retain the ecological validity of the design (i.e. line-of-sight staring from behind), but also avoids requirements for sound-proofed booths and one-way mirrors, might be to conduct experiments between offices in adjacent buildings. Jack could watch Jill through two sets of windows, perhaps using binoculars, and the experimenters could use an instant-messaging internet service between Jack's and Jill's offices to track and record each trial. Most instant-messaging software allows the conversation to be recorded, so that log could act as the official record of the results of each session.

In such an experiment, assuming a per-trial hit rate (i.e. responding 'yes' in staring trials) of 57% in unselected subjects, where 50% is expected by chance, then results significant at the $p = 0.01$ (one-tailed) level can be achieved with 90% probability with $N = 1{,}320$ total trials. This may be achieved with 66 runs of 20 trials each, although given the observation of a potential decline effect after six trials, it might be better to conduct 132 runs of 10 trials each.

Institute of Noetic Sciences
101 San Antonio Road
Petaluma
CA 94952–9524, U.S.A. DeanRadin@noetic.org

REFERENCES

Colwell, J., Schröder, S. and Sladen, D. (2000) The ability to detect unseen staring: a literature review and empirical tests. *British Journal of Psychology 91,* 71–85.

Coover, J. E. (1913) The feeling of being stared at: experimental. *American Journal of Psychology 24,* 570–575.

Lobach, E. and Bierman, D. J. (2004) The invisible gaze: three attempts to replicate Sheldrake's staring effects. *Proceedings of the Parapsychological Association Annual Convention, 2004.*

Marks, D. and Colwell, J. (2000) The psychic staring effect: an artifact of pseudo randomization. *Skeptical Inquirer (September/October),* 41–49.

Poortman, J. J. (1959) The feeling of being stared at. *JSPR 40,* 4–12.

Schmidt, S. (2001) Empirische Testung der Theorie der morphischen Resonanz: können wir entdecken wenn wir angeblickt werden? Kommentar. *Forschende Komplementärmedizin 8,* 48–50.

Schmidt, S., Müller, S. and Walach, H. (2003) Correspondence. *JP 67,* 409–411.

Schmidt, S., Schneider, R., Utts, J. and Walach, H. (2004) Distant intentionality and the feeling of being stared at: two meta-analyses. *British Journal of Psychology 95,* 235–247.

Sheldrake, R. (1998) The sense of being stared at: experiments in schools. *JSPR 62,* 311–323.

Sheldrake, R. (1999) The 'sense of being stared at' confirmed by simple experiments. *Biology Forum 92,* 53–76.

Sheldrake, R. (2000) The 'sense of being stared at' does not depend on known sensory clues. *Biology Forum 93,* 209–224.

Sheldrake, R. (2001) Experiments on the sense of being stared at: the elimination of possible artefacts. *JSPR* 65, 122–137.

Sheldrake, R. (2003) *The Sense of Being Stared At.* New York: Crown Publishers.

Sheldrake, R. (in press) The sense of being stared at: do scores improve as tests go on? Unpublished MS.

Van Bolhuis, J. (no date) An analysis of response bias in some of Rupert Sheldrake's experiments.