The feeling of being stared at: An analysis and replication

Dean Radin
Institute of Noetic Sciences
101 San Antonio Road
Petaluma, CA 94952
DeanRadin@Noetic.org

When data generated in “the feeling of being stared at” experiments are adjusted to account for response biases, hit rates associated with responding “yes” when being stared at, and “no” when not being stared at, are virtually identical. Experiments conducted where the possibility of such cues were reduced showed similar results, arguing against sensory artifacts. In a computer-based replication attempt, nearly significant results (p = 0.07) were obtained that were consistent with the estimates of previous effect sizes. A strong decline effect was also observed across runs of 20 trials.

Introduction

Consider a simple form of “the feeling of being stared at” experiment. For the sake of exposition let us call the starer(s) in such experiments “Jack” and the staree(s) “Jill.” Jack and Jill sit within a few meters of each other, Jill with her back to Jack. Jack follows a random 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,” she believes she Jack is staring at her, or “no,” he is not. The result of each such trial may be compiled into one of four categories: hit, miss, false alarm, and correct rejection, as shown in Table 1.

		response	
		yes	no
condition	staring	hit	miss
	not staring	false alarm	correct rejection

Table 1. Types of trial outcomes in a staring experiment.

Sheldrake (1998, 1999, 2000a,b, 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, and more secure conditions such having Jack stare at Jill through a window. From these published reports, plus those of Coover (1913) and Portmann (1957), I was able to retrieve a total of 33,357 trials for analysis.

The overall success rate in this database, i.e., the proportion of hits plus correct rejections, is 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 large statistical power afforded by 33,357 trials, there is little doubt that this success rate excludes chance as an explanation ($z = 16.3$, effect size per trial $= z/\sqrt{(N)} = 0.09$, $p << 10^{-10}$).

As shown in Figure 1, Sheldrake (2003) cited a consistent pattern of responses across these tests: Jill was able to tell when Jack was staring (57.8% hit rate), but not when he was not staring (51.1%). Sheldrake suggested that the existence of this pattern argues against a subliminal cuing explanation, because with a combination of subliminal cues and trial-by-trial feedback one would expect Jill to be able to learn to discriminate between staring and not staring conditions, and thus the hit rates on both staring and not staring trials should be about the same. Given that they are not, Sheldrake argued that the data suggest an anomaly consistent with a genuine “staring effect.”

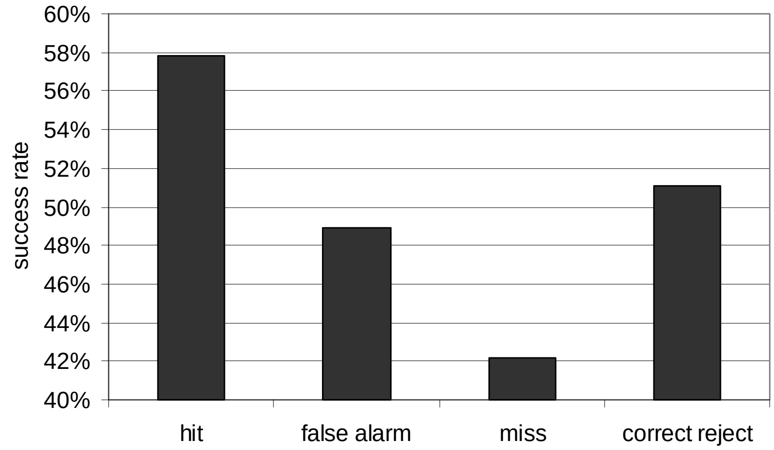

Figure 1. Observed success rates for the four types of outcomes in a staring experiment.

However, to properly interpret the statistical significance of these hit rates, they must be adjusted for Jill's response biases.[1] While the stare vs. not stare conditions were distributed randomly in these studies to preclude Jill from easily inferring the next target (50.1% vs. 49.9%, respectively), Jill cannot be expected to respond at random. In fact, in these experiments, she said "yes" 53.4% of the time and "no" 46.5%. This response bias inflates the statistical significance associated with the observed success rate for hits, and deflates the success rate for correct rejections.

To see why, consider the case where Jill is feeling paranoid and responds "yes" on every trial. Jill's response bias will not inflate her overall success rate, but her hit rate and false alarm rates will both be 100%, and her miss and correct rejection rates will be 0%. A less extreme bias would be reminiscient of the pattern noted by Sheldrake, namely a higher percentage of hits and a lower percentage of correct rejections.

A simple way to re-interpret the hit rates shown in Figure 1 is to use the formula for a z score of a difference in proportions: $z = (p_1 - p_0)/\sqrt{p_0 q_0 / N}$, where p_1 is the observed success rate, p_0 is the expected rate, q_0 = 1-p_0, and N is the number of trials under consideration. If p_0 is taken to be chance expectation of 0.5, and N = number of staring trials, then the overall "raw" z score is z(hit) = 19.9. But if p_0 is the observed response bias of 53.4% then the adjusted z score is z(hit) = 11.5. This is still far from chance, but it is also significantly smaller than the unadjusted or raw z score (see Figure 2). From this perspective it is clear that the adjusted success rates for hits and correct rejections are virtually identical, as are the rates for false alarms and misses. This is more in alignment with what one might expect if the results were due to subliminal perception.

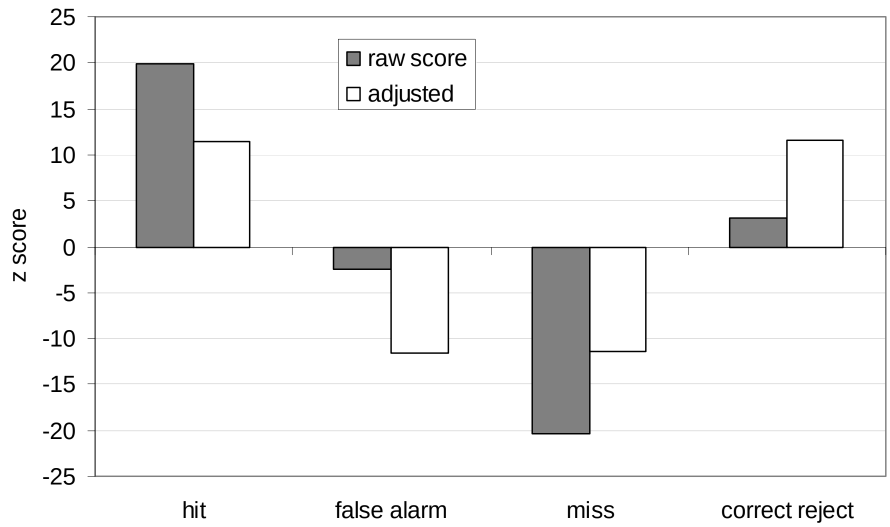

Figure 2. Raw and adjusted z scores for the four types of outcomes.

To explore the subliminal hypothesis in more detail, I partitioned Sheldrake's data into settings with progressively stronger controls for sensory cuing. Of the 33,357 trials, a total of 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 with no feedback; 4,800 were collected with the pair separated by a window.

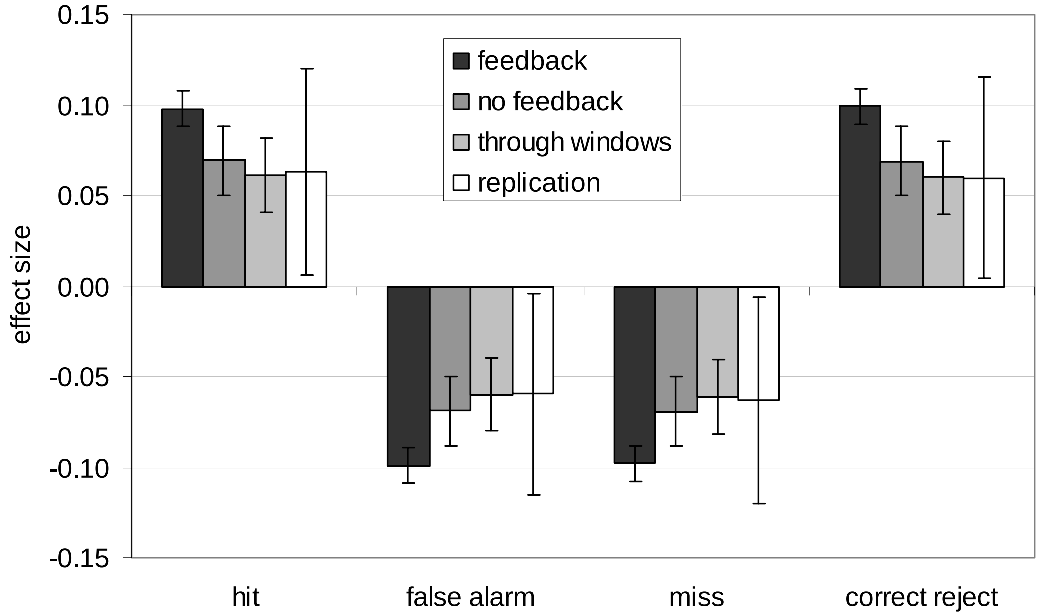

Figure 3. Adjusted effect sizes ($e = z/\sqrt{N}$), and one standard error bars, for three testing conditions and a new replication conducted as part of this study.

Figure 3 shows that study outcomes declined as testing conditions increasingly shielded against subliminal cues. However, the effects did not decline to zero, nor were the results significantly different from one another, as indicated by the error bars. This provides no support for the idea that the staring effect is solely due to sensory artifacts.

A remaining limitation of most of the previous studies is that the outcome of each trial was recorded manually, and this is known to be vulnerable to both mistakes and motivated biases. To explore what effect recording errors may have had on the overall outcome, I conducted a new experiment using an automated, computer-based design.

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 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 used by the right thumb.

Jill initiated each trial by pressing a button on the gamepad, 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 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 "I'm being stared at" or pressing 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.

Jill was required to sit directly in front of Jack, facing away from Jack at all times. Both Jack and Jill listened to the computer's prompts and sounds over headphones. The sequence of stare and not-stare trials was determined randomly, with p(stare)= p(not-stare) = 0.5. The Microsoft Visual Basic 6 pseudorandom function was used to make these random selections; the algorithm was seeded by the computer system clock time when the program began.

Results

A total of 31 sessions of 20 trials, and one session of 5 trials, were contributed by 12 pairs of participants (not including the author), or 625 trials. A total of 331 successful responses were recorded, or 53%, z = 1.48, p = 0.07. Figure 3 shows that the adjusted effect sizes across the four possible outcomes were closely in alignment with previously reported results.

Figure 4 shows the z score associated with the success rate per trial and cumulatively. Each of the first six trials were above chance expectation, accumulating to about 3 standard errors over chance at trial 6 and then declining to a terminal z score of 1.48. This pattern is reminscient of similar decline effects often observed in forced-choice tests.

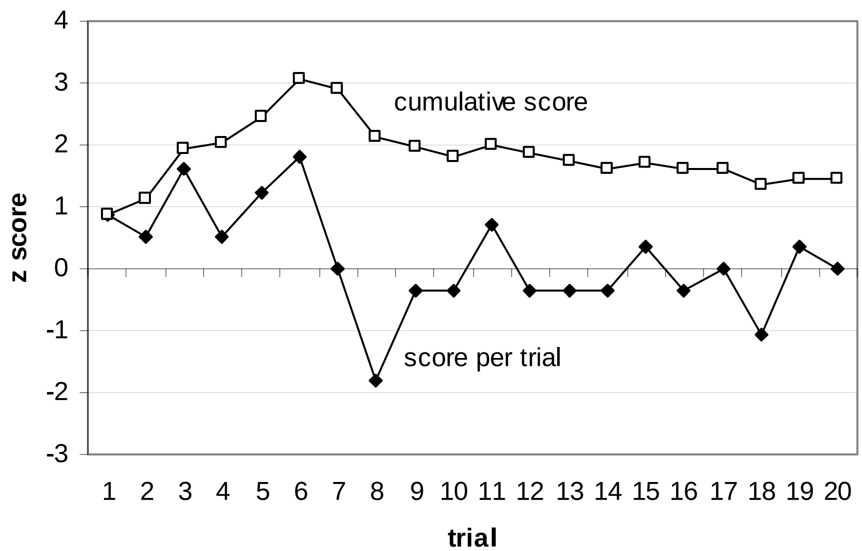

Figure 4. Z score associated with success rate per trial and cumulatively.

Examination of autocorrelations in the target sequence indicated that none of the correlations through lag 20 were statistically significant. A similar examination of the sequence of responses showed a significant negative correlation for lag +1, r = -0.14, z = -3.89, p = 0.0001. This was expected and is due to a common response bias in which subjects avoid giving the same answer or 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 result is on the cusp of significance and may not be meaningful.

Discussion

The adjusted statistical assessment of Sheldrake's data and the present replication data show a clear symmetry between hits and successful rejections, and misses and false alarms. These outcomes are consistent with both a subliminal cuing effect and a genuine staring effect. Arguing against the cuing explanation is the fact that staring without feedback and through windows resulted in effect sizes that are statistically indistinguishable from staring close-up with feedback.

In addition, two features of the present experiment suggest that Sheldrake's reported results reflect a genuine effect: (a) The computer-based recording method obviated motivated or inadvertent errors, and (b) success rates in the first few trials of the average 20-trial run were much better than in later trials. The 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.

In conclusion, the feeling of being stared at in these experiments appears to be a genuine, small magnitude effect. Assuming a modest per-trial overall success rate of 53% where 50% is expected by chance, a power analysis indicates that a $p = 0.01$ result can be achieved with power $= 0.90$ with $N = 3{,}632$ trials. This may be achieved with 181 runs of 20 trials each, although given the observation of a potential decline effect after six trials, it might be better to conduct 362 runs of 10 trials each.

If the reader is interested in using the Windows-based program I have developed to run this experiment, please contact me.

Bibliography

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

Poortman, J. J. (1959). The feeling of being stared at. Journal of the Society for Psychical Research, 40, 4-12.

Sheldrake, R. (1998) The sense of being stared at: Experiments in schools. *Journal of the Society of Psychical Research* 62: 311-323.

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

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

Sheldrake, R. (2000b). The sense of being stared at: Effects of blindfolding subjects and giving them feedback. *Journal of the Society of Psychical Research*.

Sheldrake, R. (2003). The sense of being stared at. New York: Crown Publishers.

[1] Stefan Schmidt has independently noted that such an adjustment is required.