

EXPERIMENTER EFFECTS AND THE REMOTE DETECTION OF STARING

BY RICHARD WISEMAN AND MARILYN SCHLITZ

ABSTRACT: Each of the two authors recently attempted to replicate studies in which the "receivers" were asked to psychically detect the gaze directed at them by unseen "senders." R. W.'s studies failed to find any significant effects; M. S.'s study gave positive results. The authors then agreed to carry out the joint study described in this paper, in the hope of determining why they had originally obtained such different results. The experimental design was based on each author carrying out separate experiments, but running them in the same location, using the same equipment/procedures, and drawing participants from the same subject pool. The 32 experimental sessions were divided into two sets of randomly ordered trials. Half were "stare" trials during which the experimenter directed his/her attention toward the receiver; half were "non-stare" (control) trials during which the experimenter directed his/her attention away from the receiver. The receivers' electrodermal activity (EDA) was continuously recorded throughout each session. The EDA of R. W.'s receivers was not significantly different during stare and non-stare trials. By contrast, the EDA of M. S.'s receivers was significantly higher in stare than non-stare trials. The paper discusses the likelihood of different interpretations of this effect and urges other psi proponents and skeptics to run similar joint studies.

... the experimenter effect is the most important challenge facing modern experimental parapsychology. It may be that we will not be able to make too much progress in other areas of the field until the puzzle of the experimenter effect is solved. (Palmer, 1986, pp. 220-221.)

The apparent detection of an unseen gaze (i.e., the feeling of being stared at, only to turn around and discover somebody looking directly at you) is a common type of ostensible paranormal experience, with between 68% and 94% of the population reporting having experienced the phenomenon at least once (Braud, Shafer, & Andrews, 1993a; Coover, 1913).

Some parapsychologists have attempted to assess whether this experience is based, at least in part, on genuine psi ability. Such studies use two

The authors would like to thank the following organizations for supporting the research described in this paper: The Perrott-Warrick Fund, Cambridge University, the Institute for Noetic Sciences, UltraMind, Ltd., the Hodgson Fund, Department of Psychology, Harvard University, and the University of Hertfordshire. We are also grateful to Matthew Smith and Emma Greening for their help in running this experiment and analyzing the data, John Palmer, Dorothy Pope, and the blind reviewers for their helpful comments and suggestions.

participants: a "sender" and a "receiver." These individuals are isolated from one another, but in such a way that the sender can see the receiver. Early experiments had the sender sitting behind the receiver (Coover, 1913; Poortman, 1959; Titchener, 1898); some later studies have used one-way mirrors (Peterson, 1978) or a closed-circuit television system (Braud, Shafer, & Andrews, 1993a, 1993b; Williams, 1983). The experimental session in this type of study is divided into two sets of randomly ordered "stare" and "non-stare" trials. During stare trials the sender directs his/her attention toward the receiver; during non-stare trials the sender directs his/her attention away from the receiver. Either during or after each trial a response is made by the receiver. In early studies, the receivers made verbal guesses as to whether they believed they had been stared at; later studies have measured receivers' electrodermal activity (EDA) throughout each trial. A number of studies have obtained statistically significant differences between responses to stare and non-stare trials and in a recent review of this work, Braud, Shafer, and Andrews (1993b) concluded:

We hope other investigators will attempt to replicate these studies. We recommend the design as one that is straightforward, has already yielded consistent positive results, and addresses a very familiar psi manifestation in a manner that is readily communicable and understandable to the experimental participants and to the public at large. (p. 408)

Both authors of the present paper previously attempted to replicate this staring effect. The first author (R. W.) is a skeptic regarding the claims of parapsychology who wished to discover whether he could replicate the effect in his own laboratory. The second author (M. S.) is a psi proponent who has previously carried out many parapsychological studies, frequently obtaining positive findings. The staring experiments carried out by R. W. showed no evidence of psychic functioning (Wiseman & Smith, 1994; Wiseman, Smith, Freedman, Wasserman, & Hurst, 1995). M. S.'s study, on the other hand, yielded significant results (Schlitz & LaBerge, 1997).

Such "experimenter effects" are common within parapsychology and are open to several competing interpretations (see Palmer, 1989a, 1989b). For example, M. S.'s study may have contained an experimental artifact absent from R. W.'s procedure. Alternatively, M. S. may have worked with more psychically gifted participants than R. W. had, or may have been more skilled at eliciting participants' psi ability. It is also possible that M. S. and R. W. created desired results via their own psi abilities, or fraud. Little previous research has attempted to evaluate these competing hypotheses. This is unfortunate, because it is clearly

important to establish why experimenter effects occur, both in terms of assessing past psi research and attempting to replicate studies in the future. For these reasons, the authors agreed to carry out a joint study in the hope of learning why our original studies obtained such dramatically different results.

METHOD

Design

Our joint study required M. S. and R. W. to act as separate experimenters for two different sets of trials. The two sets of trials were carried out at the same time (early October, 1995) and in the same location (R. W.'s laboratory at the University of Hertfordshire in the U.K.). In addition, the experimenters used the same equipment, drew subjects from the same subject pool, and employed exactly the same methodological procedures. The only real difference between the trials was that one set was carried out by M. S. and the other set was run by R. W. We were curious to discover if, under these conditions, we would continue to obtain significantly different results. Each study had one independent variable with two levels—stare and non-stare. The dependent variables were the receivers' EDA during the experimental session and their responses to a "belief-in-psi" questionnaire.

Participants

Thirty-two subjects (10 males and 22 females; mean age of 25.72, age range 18 – 49) acted as receivers. Thirty of these were undergraduate psychology students studying at the University of Hertfordshire. The remaining two were the authors' colleagues. M. S. and R. W. acted in a dual capacity as both experimenter and sender.

Apparatus and Materials

Layout of room. It was clearly important to minimize the possibility of any sensory leakage between sender and receiver during the experimental sessions. For this reason the receiver was located in the University's Social Observation Laboratory while the sender was located in a small room approximately 20 meters away from the laboratory (see Figure 1).

Video equipment. A Panasonic AG-450 video camera was positioned in front of the receiver and relayed an image (via a long cable connecting the two rooms) to a 14-inch JVC color TV monitor in the sender's room.

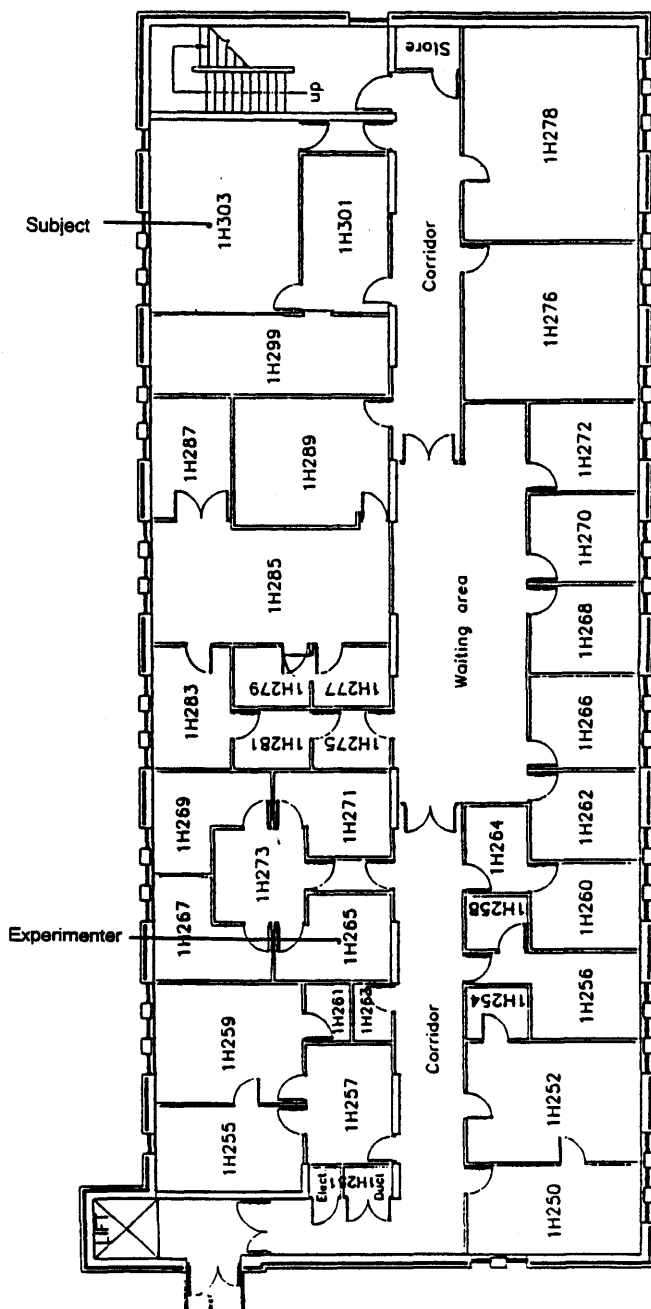


Figure 1. Locations of experimenter and subject during session.

This one-way closed circuit television system allowed the experimenter to see the subject, but not vice versa.

EDA measurement. The receivers' EDA (electrodermal activity) was recorded by the RelaxPlus system (a commercially available hardware and software package produced by UltraMind, Ltd.). This system measures skin resistance level by placing a constant current across two stainless steel electrodes and then recording the resistance encountered by that current at a rate of 10 samples per second. The system filters for possible artifacts (caused, for example, by movement) and records data to the computer's hard disk. The equipment (i.e., electrodes, input device, computer, computer monitor) was located next to the receiver throughout the experiment. The part of the program involved in storing the details of subjects and their physiological data could be accessed only via a password known only to M. S. and R. W. Data from the Relax-Plus system were then fed into a spreadsheet (Microsoft's Excel) in order to calculate the mean EDA for each 30-second trial. All statistical analyses were carried out using the Statview software package.

Belief-in-psi questionnaire. The receivers were asked three questions concerning their attitudes toward psi (see Appendix). They indicated their responses on a seven-point scale ranging from -3 to +3. A general "belief-in-psi" score was obtained by summing the receiver's responses over all three questions. Low scores on this questionnaire were taken to indicate strong belief in psi.

Trial randomization. The receivers' EDA may decline during a session for several reasons (e.g., the apparatus measuring EDA may warm up or the participants may habituate to their surroundings). This decline could lead to artifactual evidence for psi if stare trials tend to precede non-stare trials. The following randomization procedure was devised to minimize this possible artifact.

Prior to the experiment, an individual not involved in running the experiment (Matthew D. Smith) prepared a set of 32 sheets, each of which contained the order of the 32 stare or non-stare trials for one session. For 16 of these sheets the trial orders were generated in the following way: M. D. S first opened the random number table (Robson, 1983, Appendix Three), chose a number as an entry point into the table, and then threw a die twice. The numbers that came up determined how he moved from this entry point to an actual starting point. The eight consecutive numbers located in the row to the right of this starting point determined the order of the stare and non-stare trials. An even number translated into an ABBA (stare, non-stare, non-stare, stare) order while an odd number translated into a BAAB (non-stare, stare, stare, non-stare) order. The trial order for the remaining 16 sheets was determined by counterbalancing the orders of the randomized sheets just described.

Thus, a stare, non-stare, non-stare, stare on a randomized sheet became a non-stare, stare, stare, non-stare on a counterbalanced sheet. All 32 sheets were then mixed together, placed in an opaque folder, and kept in a locked drawer in R. W.'s office. M. D. S. was aware of the experimental hypotheses prior to carrying out the above randomization procedure.

Procedure

The receivers were run individually. On arriving at the laboratory, each one was met by either R. W. or M. S. Most were run by whichever of the experimenters was free to carry out the session; however, on a few occasions (e.g., when a receiver was a friend or colleague of one of the experimenters) the experimenter would be designated in advance of the trial. Thus most subjects were assigned to experimenters in an opportunistic way, rather than by one that was properly randomized (e.g., via random number tables or the output of a random number generator). The experimenter showed the subject to the receiver's room and explained the purpose of the experiment. Next, the experimenter attached electrodes to the first and third fingers of the participant's nondominant hand and made sure that the RelaxPlus system was correctly monitoring their EDA. The receivers were asked not to move their hand unnecessarily, nor to try to guess when they might be being stared at, but instead to simply remain as open as possible to any remote influence. The experimenter entered the receiver's personal data in a computerized database, initiated the recording of EDA, started a stopwatch, and left the receiver's room.

It was important that receivers were not aware of the order of the stare and non-stare trials *before* the start of the experimental session. For this reason, the list of trial orders was only selected by the experimenter only *after* he or she had left the receiver's room. The experimenter then went to R. W.'s office, retrieved the folder containing the lists of trial orders, selected any sheet he or she wanted, and proceeded to the sender's room.

Two minutes after initiating the recording of the receiver's EDA, the experimenter started to carry out the designated order of stare and non-stare trials; this order was presented to the experimenters in the form of a list. During stare trials, the experimenter quietly directed his/her attention toward the receiver; during non-stare trials the experimenter quietly directed this attention away from the receiver. Each trial lasted 30 seconds. Throughout this time the receiver completed the belief-in-psi questionnaire and then read some magazines. All of the magazines were selected to be relatively bland in content in order to minimize possible effects on the receivers' EDA.

On completion of all 32 trials, the experimenter returned to the receiver's room, thanked the participant, and told him or her that feedback of the overall results would be given within the next few weeks.

At the end of each experimental day, both experimenters copied that day's data (from their own participants as well as from the other experimenter's participants) onto their own floppy disk.

RESULTS¹

Primary Analyses

All analyses were preplanned. A Wilcoxon signed rank test was used to compare receivers' total EDA for the 16 stare trials with their total EDA during the 16 non-stare trials.² Receivers run by R. W. did not differ from chance expectation (Wilcoxon $z = -.44$, $df = 15$, $p = .64$, two-tailed). In contrast, receivers run by M. S. showed a significant effect (Wilcoxon $z = -2.02$, $df = 15$, $p = .04$, two-tailed).

A "detect score" was then calculated for each subject by subtracting the total EDA during the stare trials from the total EDA for the non-stare trials. An unpaired t test revealed that the detect scores of M. S.'s subjects were not significantly different from those of R. W.'s ($df = 30$, $t = 1.39$, $p = .17$, two-tailed).

Secondary Analyses

Table 1 contains the correlation coefficients between participants' belief-in-psi questionnaire scores and their detect scores. Spearman rank correlation coefficients revealed that none of these correlations were significant. Table 1 also contains the means (and standard deviations) of

¹This experiment was first reported at the 1996 Convention of the Parapsychological Association (Wiseman & Schlitz, 1996). While preparing the paper for journal publication, the authors reviewed the data and discovered an error in the way one subject's data had been transferred into the statistical package used for the analyses. For this reason the results reported here are slightly different from those reported in Wiseman and Schlitz (1996).

²Previous studies (e.g., Braud et al., 1993a, 1993b) have assessed their results by creating a "psi score" (the sum of EDA during stare trials divided by the sum of the total EDA) for each participant and then using a one-sample t test to determine the degree to which these scores deviate from chance expectation. This procedure obscures the question of whether an overall result is caused by a very small number of participants performing extremely well. The Wilcoxon sign rank test is more conservative than the one-sample t test because it is less influenced by the size of the deviation between participants' scores.

the questionnaire scores for R. W.'s group, M. S.'s group, and all participants.

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR
THE BELIEF IN PSI QUESTIONNAIRE
AND
CORRELATION COEFFICIENTS AND *p* VALUES BETWEEN SUBJECTS'
QUESTIONNAIRE SCORES AND DETECT SCORES

	R.W's participants	M. S.'s participants	All participants
Mean	1.94	-.81	.56
Standard deviation (<i>SD</i>)	4.22	4.12	4.33
Correlation (<i>r</i>) (Corrected for ties)	-.15	.32	.15
<i>z</i> score	-.58	1.23	.84
<i>p</i> value, two-tailed	.56	.22	.39

DISCUSSION

Subjects run by R. W. did not respond differently to stare and non-stare trials. In contrast, participants run by M. S. were significantly more activated in stare than non-stare trials. These findings can be interpreted in several ways.

First, one might argue that M. S.'s significant results were caused by some type of experimental artifact. Several steps were taken to guard against this possibility. For example, neither the receivers nor the experimenters knew the order of the stare and non-stare trials before the start of the experiment; the location of the rooms minimized the possibility of any sender-to-receiver sensory leakage; and the randomization procedure ensured that the results were unlikely to be caused by progressive errors. This, coupled with the fact that one would expect any artifact to influence the results of both studies, suggests that M. S.'s significant results are unlikely to have been caused by a methodological error.

Second, one could argue that either R. W.'s or M. S.'s results were caused by receivers' cheating. For example, subjects could have discovered the order of stare and non-stare trials before the experimental session and altered their EDA accordingly. Alternatively, participants could have altered their data files so that they coincided with the order of stare and non-stare trials. Several factors mitigate against these

possibilities. First, such cheating would have been far from straightforward. For example, the selection of trial order was carried out a few moments before the start of the experimental session and it could only have been accessed by a participant who had installed some kind of covert monitoring equipment in the sender's room. Likewise, the computer could only be accessed if a participant had discovered a password which was known only to the experimenters. Also, neither R. W.'s or M. S.'s significant results are due to one exceptional participant, and one would therefore have to hypothesize that several participants successfully cheated.

Third, the results could have been caused by experimenter fraud. Although the experiment was not designed to make such fraud impossible, its design does mean that certain types of cheating would have been extremely unlikely. For example, neither experimenter could have decided to include data only from certain subjects because the full list of all subjects was known to both experimenters. However, more sophisticated forms of cheating were *theoretically* possible. For example, one experimenter could have substituted false sets of EDA values for subjects' actual values before the data were analyzed. Although possible, this would have been far from straightforward because subjects were frequently scheduled back-to-back (thus cutting to a minimum the time available for recording a false replacement session), and each experimenter made a back-up disk of all of the day's sessions at the end of each day (thus minimizing the possibility of an experimenter's substituting data after the day they had been recorded). In addition, no evidence of any cheating was uncovered during the running of the experiment or analysis of the data.

Fourth, one could argue that M. S. was working with a more "psychically gifted" population than R. W. was. This also seems unlikely because the receivers were assigned to the two experimenters in an opportunistic fashion.

Fifth, it is possible that M. S. was more skilled at eliciting subjects' psi ability than R. W. was. Interestingly, M. S.'s subjects scored higher on the "belief-in-psi" questionnaire than R. W.'s subjects did (although this difference just failed to reach significance: unpaired t value = 1.86, $df = 30$, $p = .072$, 2-tailed). Given that participants were opportunistically assigned to experimenters, this difference might be a reflection of the different ways in which R. W. and M. S. oriented receivers at the start of the experiment. It seems quite possible that the experimenters' own level of belief/disbelief in the existence of psi caused receivers to express different levels of belief/disbelief in psi and to have different expectations about the success of the forthcoming experimental session.

Videotapes of R. W.'s and M. S.'s induction procedures are currently being analyzed to identify differences in interaction and content.

Finally, it is also possible that both R. W. and M. S. used their own psi abilities to create the results they desired. This interpretation, if genuine, supports past research which suggests that successful experimenters (i.e., those who consistently obtain significant effects in psi studies) outperform unsuccessful ones on a variety of psi tasks (see Palmer, 1986, for a review of the literature supporting this notion).

In conclusion, this study reveals the value of developing collaborative relationships between skeptics and psi proponents. Both authors view this study as an initial step in the investigation of experimenter effects in psi research. Additional experiments would further aid our understanding of such effects. For example, it would be useful to carry out an experiment in which one experimenter interacted with the receiver and the other carried out the stare and non-stare trials during the experimental session. Such a study would help discover whether our initial interactions with the receiver or our behavior during the experimental session caused the results reported in this paper. We, the authors, hope to carry out such a study in the near future, and we urge other psi proponents and skeptics to run similar studies.

REFERENCES

- BRAUD, W., SHAFER, D., & ANDREWS, S. (1993a). Reactions to an unseen gaze (remote attention): A review, with new data on autonomic staring detection. *Journal of Parapsychology*, *57*, 373-390.
- BRAUD, W., SHAFER, D., & ANDREWS, S. (1993b). Further studies of autonomic detection of remote staring: replications, new control procedures, and personality correlates. *Journal of Parapsychology*, *57*, 391-409.
- COOVER, J. E. (1913). The feeling of being stared at. *American Journal of Psychology*, *24*, 570-575.
- PALMER, J. (1986). ESP research findings: the process approach. In H. L. Edge, R. L. Morris, J. Palmer, & J. H. Rush (Eds.), *Foundations of parapsychology* (pp. 184-222). London: Routledge & Kegan Paul.
- PALMER, J. (1989a). Confronting the experimenter effect. *Parapsychology Review*, *20*, 1-4.
- PALMER, J. (1989b). Confronting the experimenter effect. Part 2. *Parapsychology Review*, *20*(5), 1-5.
- PETERSON, D. M. (1978). Through the looking glass: an investigation of extrasensory detection of being stared at. M.A. Thesis, University of Edinburgh.
- POORTMAN, J. J. (1959). The feeling of being stared at. *Journal of the Society for Psychical Research*, *40*, 4-12.

- ROBSON, C. (1983). *Experiment, design and statistics in psychology*. London: Penguin Books.
- SCHLITZ, M. J., & LABERGE, S. (1997). Covert observation increases skin conductance in subjects unaware of when they are being observed: A replication. *Journal of Parapsychology*, **61**, 185-196.
- TITCHENER, E. B. (1898). The feeling of being stared at. *Science*, **8**, 895-897.
- WILLIAMS, L. (1983). Minimal cue perception of the regard of others: The feeling of being stared at. Paper presented at the 10th Annual Conference of the Southeastern Regional Parapsychological Association, West Georgia College, Carrollton, GA. See *Journal of Parapsychology*, **47**, 59-60.
- WISEMAN, R., & SCHLITZ, M. (1996). Experimenter effects and the remote detection of staring. *Proceedings of the Parapsychological Association 39th Annual Convention*, 149-157.
- WISEMAN, R., & SMITH, M. D. (1994). A further look at the detection of unseen gaze. *Proceedings of the Parapsychological Association 37th Annual Convention*, 465-478.
- WISEMAN, R., SMITH, M. D., FREEDMAN, D., WASSERMAN, T., & HURST, C. (1995). Two further experiments concerning the remote detection of an unseen gaze. *Proceedings of the Parapsychological Association 38th Annual Convention*, 480-492.

*Dept. of Psychology
University of Hertfordshire
College Lane
Hatfield, Hertfordshire
England AL10 9AB
UK*

*Institute of Noetic Sciences
475 Gate Five Road
Suite 300
Sausalito, CA 94965*