EXPERIMENTER EFFECTS AND THE REMOTE DETECTION OF STARING: AN ATTEMPTED REPLICATION

Richard Wiseman
Perrott Warrick Research Unit
University of Hertfordshire
&
Marilyn Schlitz
Institute for Noetic Sciences

ABSTRACT

Both authors recently ran experiments to discover whether people can psychically detect when another person is staring at them. R.W. is a skeptic regarding claims of parapsychology and M.S. is a psi proponent. R.W.’s studies obtained chance results whilst M.S.’s study obtained statistical significance. The authors then carried out a joint study to help determine why their experiments had obtained different results. M.S. and R.W. acted as separate experimenters for two different sets of trials. These trials were carried out at the same time, in the same location, used the same equipment, drew participants from the same pool and employed the same procedures. The data from M.S.’s participants were statistically significant whilst the data from R.W.’s participants were not. This paper describes an attempted replication of this initial joint study. Participants were hooked up to a computer that recorded their electrodermal activity (EDA). A videocamera was placed in front of the participant and fed an image of them to a monitor located in a separate room. Each experimental session consisted of thirty two, thirty second, periods. Half of these periods were randomly allocated to a ‘stare’ condition and half to a ‘nonstare’ condition. During the stare condition the experimenter looked at the monitor and attempted to remotely influence the participant’s EDA. During the nonstare condition the experimenter looked away from the monitor. The EDA of R.W.’s participants was not significantly different between the two conditions. In contrast, the EDA of M.S.’s participants was significantly lower during the stare than nonstare periods. The paper discusses competing interpretations of these results and possible future research in this area.

INTRODUCTION

Parapsychologists have carried out several experiments designed to discover whether people can psychically detect when another person is staring at them (see Braud, Shafer & Andrews, 1993). A few years ago both authors attempted to replicate these experiments. The first author (R.W.) is a skeptic regarding claims of parapsychology whilst the second author (M.S.) is a psi proponent. The staring studies conducted by R.W. obtained chance results (Wiseman &
were statistically significant (Schlitz & LaBerge, 1997).

The authors then agreed to carry out a joint study to help determine why these experiments had obtained different results (Wiseman & Schlitz, 1997). The study involved M.S. and R.W. acting as separate experimenters for two different sets of trials. These trials were carried out at the same time, in the same location, used the same equipment, drew participants from the same pool and employed the same procedures.

Participants were hooked up to a computer that constantly recorded their electrodermal activity (EDA). A videocamera was placed in front of the participant and fed an image of them to a monitor located in a separate room. Each experimental session lasted for sixteen minutes and consisted of thirty two, thirty second, periods. Half of these periods were randomly allocated to a ‘stare’ condition and half to a ‘nonstare’ condition. During the stare condition the experimenter looked at the monitor and attempted to remotely influence the participant’s EDA. During the ‘nonstare’ condition the experimenter looked away from the monitor and directed his/her attention away from the participant.

The two sets of trials obtained different results. M.S.’s participants were significantly more activated in the stare than nonstare condition. There was no significant difference between the stare and nonstare EDA levels of R.W.’s participants.

Given the potential importance of such ‘experimenter effects’ in parapsychology (see Palmer, 1989a,b; Wiseman & Schlitz, 1997) the authors decided to attempt to replicate their joint study. This paper outlines the methods and results of the attempted replication.

**METHOD**

**Design**

M.S. and R.W. acted as separate experimenters for two different sets of trials. The trials were carried out at The Institute of Noetic Sciences (San Francisco, USA) between June and August 1998. The study had one independent variable with two levels - stare and nonstare. The dependent variables were the participants’ EDA during the experimental session and their responses to a ‘belief-in-psi’ questionnaire.

**Participants**

Participants were either employees or members of The Institute of Noetic Sciences. R.W. and M.S. each ran thirty five participants. Our previous joint study involved sixteen participants per experimenter and the data from M.S.’s participants obtained an effect size of .51. Increasing the number of participants to thirty five per experimenter resulted in 85% power with a two tailed p value <= .05 (Howell, 1992). R.W. visited the Institute in June and August 1998. R.W. completed thirty one trials during his first visit and the remaining four during his second visit. M.S. completed seven trials during R.W.’s first visit, nine between visits and nineteen during his second visit. Table 1 contains the gender and age distribution for participants. R.W. and M.S. acted as in a dual capacity as both experimenter and sender.
Table 1: Participants’ Gender and Age Distribution

<table>
<thead>
<tr>
<th>Age category</th>
<th>M.S.’s participants</th>
<th>R.W.’s participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>Female</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>18-25</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>26-35</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>36-45</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>46-55</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>56-65</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Over 65</td>
<td>3</td>
<td>9</td>
</tr>
</tbody>
</table>

Apparatus and Materials

**EDA Measurement**

Participants’ EDA was monitored and recorded using the Pro Comp system (a commercially available hardware and software package produced by Thought Technology Ltd). This system measures skin resistance level by placing a constant current across two electrodes and measuring the resistance encountered by the current at a rate of two samples per second. The system filters for possible artifacts (caused, for example, by hand movement) and records data to the computer’s hard disk.

**Trial Randomization**

The order of the stare and nonstare periods was randomly determined via a specially written computer programme. This programme used the computer’s system time to seed a Quickbasic pseudorandom number generator that provided a string of eight digits made up of 0s and 1s. Each digit determined the order of four stare/nonstare periods. A ‘0’ was translated into an ABBA (stare-nonstare-nonstare-stare) order whilst a ‘1’ resulted in a BAAB (nonstare-stare-nonstare-nonstare) order. The ABBA/BAAB method of randomization is designed to minimize the possible effect of any progressive errors, such as the warming up of equipment or participants habituating to their surroundings (see Wiseman & Schlitz, 1997). The computer monitor displayed the word ‘stare’ throughout the whole of the stare periods and the word ‘control’ during the nonstare periods. At the end of each stare/nonstare period the programme randomly inserted a gap of either one, two, three, four or five seconds before initiating the next period. This random time interval was designed to disrupt the possible effect of any periodicity in participant’s physiology (see Schlitz and LaBerge, 1997). The resulting order of stare and nonstare periods, and the number of seconds between each period, were stored on the computer’s hard disk.
Room layout

The participant and experimenter were located in two rooms separated by a corridor (see Figure 1). A videocamera (Cannon H800 Hi8) was positioned in front of the participant and relayed an image, via a long cable connecting the two rooms, to a colour monitor (15 inch Sakata SC100) in the experimenter’s room. An opaque sheet of paper was attached to the top of the monitor in such a way that allowed it to be quickly and silently flipped down to cover the monitor screen. This closed circuit television system allowed the experimenter to see the participant, but not vice versa. The participant’s room contained a comfortable chair, the videocamera, electrodes for EDA measurement, a selection of magazines, a tape recorder (AIWA CA-W80), audio tape of white noise and headphones (Optimize LV-20). The computer running the EDA measurement system was located directly outside the participant’s room. The experimenter’s room contained the monitor, and the computer involved in assigning the order of stare and nonstare periods.

![Diagram of room layout](image)

Figure 1: Schematic diagram of room layout (measurements approx.)

Belief-in-psi questionnaire

This questionnaire presented participants with the following three questions regarding their belief in psychic phenomena: ‘Do you believe that some people have some form of psychic ability?’, ‘Do you believe that you have some form of psychic ability?’, ‘Do you believe that you might be able to demonstrate some form of psychic ability in this experiment?’. Participants were asked to rate their level of agreement/disagreement with each question on a 1 (Definitely No) to 7
(Definitely Yes) scale. A general belief-in-psi score was obtained by summing their responses over all three questions. Low scores on this questionnaire indicate low belief in psi.

**PROCEDURE**

All participants were run individually. On arriving at the experimental area, each participant was met by either R.W. or M.S. Most participants were run by whichever of the experimenters was free to carry out the session, however, on a few occasions (e.g., when the participant was a friend of M.S.) the experimenter would be designated in advance of the session. Thus participants were assigned in an opportunistic way, rather than via a more formal method of randomization.

The experimenter explained the purpose of the study and asked participants to complete the belief-in-psi questionnaire. Participants were then shown into the participant’s room and asked to sit in the comfortable chair. The experimenter placed the headphones on the participant, attached electrodes to the first and third fingers of their nondominant hand, and ensured that the Pro Comp system was properly monitoring their EDA.

The experimenter asked the participant to minimize any unnecessary movement of their nondominant hand, not to try to guess when (they might be started at, but instead remain as open as possible to any remote staring effect. The experimenter encouraged the participant to either relax or look at the magazines during the experimental session. Finally, the experiment turned on an audio tape containing white noise and left the participant’s room, closing the door behind them.

The experimenter initiated the recording of the participant’s EDA and entered the experimenter’s room, closing the door behind them. The experimenter initiated the programme that assigned the order of stare and nonstare periods, and started to carry out the designated order. During the stare periods the experimenter looked at the image of the receiver on the monitor and quietly directed his/her attention toward the participant. During the nonstare periods the experimenter silently covered the screen with a opaque sheet of paper and directed his/her attention away from the participant. Each period lasted thirty seconds.

On completion of all thirty two periods, the experimenter returned to the participant’s room, thanked the participant, and told him or her that feedback of the overall results would be given a few weeks after the study had been completed.

At the end of each experimental day on which he was present R.W. copied files containing participants’ EDA and stare/nonstare orders onto a floppy disk. Data collected on days when R.W. was absent were copied onto the disk whenever he next returned to the laboratory.

**RESULTS**

Each participant’s results were collated using a specially written computer programme that combined data from the two computer files containing their EDA, and order of stare and nonstare periods. The programme first discarded the EDA during the average time taken between the experimenter initiating the EDA recording and the start of the first stare/nonstare period, and then summed the participant’s total EDA over the sixteen stare and sixteen nonstare periods. Data
from three participants were then analyzed by hand to ensure that these scores were accurate. All statistical analyses were preplanned and carried out using the Statview software package.

**Primary Analyses**

A Wilcoxon signed rank test was used to compare participants’ total EDA for the sixteen stare periods with their total EDA during the sixteen nonstare periods. Participants run by R.W. did not differ from chance expectation (z [corrected for ties]=-.39, p=.69 [two tailed], effect size=-0.07). Participants run by M.S. showed a statistically significant effect, with participants being significantly less activated during the stare than nonstare periods (z [corrected for ties]=-1.93, p=.05 [two tailed], effect size=-0.33).

A ‘detect’ score was then calculated for each participant by subtracting the total EDA during the stare periods from their total EDA for the nonstare periods (see Wiseman and Schlitz [1997] for a detailed discussion of this measure). An unpaired t test revealed that the detect scores of M.S.’s participants were not significantly different from those of R.W.’s participants (df=68, t=-.77, p=.44 [two tailed], effect size=0.09).

**Secondary Analyses**

Table 2 contains the means and standard deviations of participants’ belief-in-psi scores, and the Spearman rank correlation coefficients between these scores and their detect scores. None of the correlations were significant.

<table>
<thead>
<tr>
<th></th>
<th>R.W.’s Participants (N=35)</th>
<th>M.S.’s Participants (N=34)</th>
<th>All Participants (N=69)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean and standard</td>
<td>5.26 (1.12)</td>
<td>5.64 (.78)</td>
<td>5.44 (.98)</td>
</tr>
<tr>
<td>deviation (in brackets)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spearman rank</td>
<td>-.07</td>
<td>.18</td>
<td>.05</td>
</tr>
<tr>
<td>correlation coefficients (corrected for ties)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>z-score</td>
<td>-.43</td>
<td>1.02</td>
<td>.39</td>
</tr>
<tr>
<td>p-value, two tailed</td>
<td>.66</td>
<td>.31</td>
<td>.69</td>
</tr>
</tbody>
</table>

N.B. Due to time constraints one of M.S.’s participants did not complete the belief-in-psi questionnaire.

**DISCUSSION**

This study replicated the findings of our initial joint study. The EDA of R.W.’s participants was not significantly different during stare and nonstare periods. In contrast, the EDA of M.S.’s participants was significantly higher in the nonstare than stare periods. Interestingly, the effect in M.S.’s data was opposite to that found in our previous joint study, wherein her
participants were significantly lower in the nonstare than stare periods. As noted in Wiseman & Schlitz (1997), the differences in results between experimenters could be interpreted in many different ways.

First, the result could have been caused by an experimental artifact in the randomization procedure that affected M.S.’s, but not R.W.’s, data. However, such an explanation seems unlikely given that the experiment was designed to exclude obvious artifacts. For example, both experimenter and participant were unaware of the random order of stare and nonstare periods prior to the start of each session. Also, the method of stare and nonstare assignment minimized the potential effect of any progressive errors or periodicity within participants’ physiology.

Second, the results could have been due to some form of undetected sensory leakage between experimenter and participant. The degree of shielding against sensory leakage in this study was not as great as in our initial experiment. In our initial experiment the participant and experimenter were approximately 20 meters apart. In this study, space constraints resulted in the participant and experimenter being only a few meters apart. However, the participant’s room and experimenter’s rooms were separated by a corridor, the doors to both rooms were closed throughout all sessions, the experimenters made very little noise during the stare and nonstare periods, and during the experimental session the participant’s wore headphones that carried white noise. As such, the authors believe that there existed little potential for visual or auditory leakage.

Third, it is possible that the results are due to participant cheating. Again, such a possibility seems unlikely for several reasons. First, the random orders of stare and nonstare periods were generated during the experimental session and so could not have been discovered beforehand. Second, it is unlikely that participants could have easily altered their EDA data, or the order of stare and nonstare periods, without detection as all data files were regularly copied to floppy disk. Finally, as in our initial study, neither R.W.’s or M.S.’s results are due to one exceptional participant, and one would therefore have to hypothesize that several participants cheated.

Fourth, the results could have been caused by experimenter fraud. As noted in Wiseman & Schlitz (1997), the experiment was not designed to make such fraud impossible, but that certain types of cheating would have been unlikely to have been successful. However, there was one major difference between this experiment and our previous joint study. In our previous study both experimenters were present when all of the experimental sessions were being conducted. In this study, time constraints forced M.S. to run nine of her sessions in between R.W.’s visits to the Institute. However, there were no significant difference between the detect scores of these nine sessions and the twenty six sessions conducted by M.S. whilst R.W. was present (Mann Whitney U prime = -11, z[corrected for ties]=-11, p[two tailed]=.91). In addition, no evidence of any cheating was uncovered during the running of the experiment or analysis of the data.

Fifth, it is possible that M.S. was working with more psychically gifted participants than R.W. This seems unlikely given the opportunistic way in which participants were assigned to experimenters. However, in future studies the authors intend to examine this notion by using a more formal randomization procedure (e.g., random number tables or a RNG) to assign participants.
Sixth, M.S. may have been more skilled at eliciting participants’ psi abilities than R.W. In our initial study M.S.’s participants scored higher on the belief-in-psi scale than R.W.’s participants, and it seemed possible that this difference might reflect the different ways in which M.S. and R.W. oriented participants at the start of the experiment. Interestingly, this trend also existed in the present study (see Table 2), although the difference between the belief levels of the two groups was not statistically significant (df=67, t=-1.63, p=.1 [2 tailed]).

Finally, as discussed in Wiseman & Schlitz (1997), some psi proponents might argue that R.W. and/or M.S. used their own psi abilities to create the results that they desired. For example, the randomisation sequence was initiated by the experimenters pressing the computer space bar at the start of each session. Thus, it was possible that they psychically chose to initiate the sequence in such a way as to ensure that the stare periods matched especially high EDA values.

In conclusion, this experiment replicated the findings of our initial joint study and suggests that different experimenters may obtain different results when carrying out the same study. An understanding of such experimenter effects are vital within parapsychology, both for evaluating past psi research and attempting to replicate studies in the future. Having twice obtained such an effect the authors now intend to carry out additional joint studies to help tease apart the competing interpretations outlined above, and would strongly encourage other skeptic and proponent pairs to carry out similar work.

REFERENCES


FOOTNOTES

1) The authors would like to thank The Society for Psychical Research, The Institute of Noetic Sciences, The Perrott Warrick Fund and The University of Hertfordshire for funding the research described in this paper. We are also grateful to Dean Radin for writing the computer programmes used to randomly assign the order of stare and nonstare periods, and to collate participants' total EDA scores. In addition, we would like to thank Jerry Solfin for supplying and setting-up the Pro Comp system, and Adrienne Smucker for her invaluable help in recruiting and scheduling participants.

2) At the end of six of the sessions (two run by R.W. and four by M.S.) the experimenter discovered that participant's EDA had not been properly recorded due to equipment failure. These sessions could not be included in the database and the authors ran additional participants until they had successfully completed thirty five sessions each.