EXPERIMENTER EFFECTS WITH A REMOTE FACILITATION OF ATTENTION FOCUSING TASK: A STUDY WITH MULTIPLE BELIEVER AND DISBELIEVER EXPERIMENTERS

BY CAROLINE WATT AND PETER RAMAKERS

ABSTRACT: This article reports the 4th study in a series investigating experimenter effects with a remote facilitation of attention focusing psi task. The “helper” focuses attention on a candle and presses a button whenever he or she feels distracted. Simultaneously, the remote “helper” follows a randomized counterbalanced schedule of “help” and “control” periods. It was predicted that the helper would have fewer distractions during the help periods compared with the control periods. Nine psi believers and 5 disbelievers were trained to conduct a psi session and then conducted 36 psi trials in total. It was predicted that participants tested by believer experimenters would show greater remote facilitation of focusing than those tested by disbelievers. Questionnaires measured participants’ paranormal belief, expected and perceived success at the psi task, experimenter ratings, and experimenters’ personality and cognitive ability. Overall, there were significantly fewer help presses than control presses, indicating an effect of remote facilitation on the focusing task. Participants tested by believer experimenters had higher scores on the psi task than those tested by disbeliever experimenters, indicating an experimenter effect. There were no differences between participants or experimenters on the questionnaire measures.

The experimenter effect, in which certain experimenters seem consistently to obtain positive psi results while others do not, is one of parapsychology’s most vexing problems. It is perhaps a pivotal question for the issue of replicability of psi effects. Many researchers over the years have suggested different factors that may contribute to experimenter effects in parapsychology, and these factors may not be mutually exclusive. The experimenter’s own psi may be the source of the results in their experiments (see Stanford, 1981, 1990, for discussion of this thorny issue). In addition, successful experimenters may create a psi-conducive atmosphere or psi-conducive expectancies in their participants (e.g., Honorton, Ramsey, & Cabibbo, 1975; Schneider, Binder, & Walach, 2000; Watt & Baker, 2002).

We are grateful to the Perron-Warrick Fund and the Koestler Parapsychology Unit for providing financial support for this study. An earlier version of this article was presented at the 2002 Parapsychological Association Convention, Paris, France. We thank the referees for their helpful comments.
A more direct approach to the question of the experimenter's psi belief comes from a preliminary study by Parker (1975) in which the experimenters were ESP believers and ESP disbelievers. Each experimenter group was given a talk intended to bolster their prior expectancies (whether positive or negative), and the disbelievers were told to expect scoring around chance level, whereas the believers were told to expect scores above chance level. Despite the small sample size, the believer experimenters obtained significantly more hits than the disbelievers, a result that may indicate the utility of this method of investigating the effect of experimenter belief on psi outcome.

The present study follows up Parker's (1975) promising results by using a similar "multieXperimenter" design. It is the latest in a series of studies by Caroline Watt into experimenter effects (Watt & Baker, 2002; Watt & Brady, 2002). For the psi measure in these studies, one individual, the "helper," attempts to help another remote individual, the "helpee," have fewer distractions on an attention-focusing task. This task was virtually identical to that used in two previous studies that found significantly fewer distractions during help periods compared with control periods (Brady & Morris, 1997; Braud, Shaffer, McNell, & Guerra, 1995) and conceptually follows the DMILS (direct mental interaction with living system) paradigm. The helper follows a random schedule of "help" and "control" (i.e., no help) periods while at the same time in a separate sensorily isolated room the helpee attempts to focus his or her attention on a candle. The helpee is asked to press a button whenever he or she becomes distracted from the focus, and the dependent variable is the number of button presses (i.e., self-reported distractions) per epoch. A remote facilitation of attention focusing effect is indicated by the helpee having significantly fewer button presses during help epochs compared with control epochs.

Both the present study and the previous studies in this series (Watt & Baker, 2002; Watt & Brady, 2002) used the same number, duration, and counterbalanced random scheduling of epochs as were used in the studies by Braud et al. (1995) and Brady and Morris (1997).

In the present study, we selected a number of novice experimenters on the basis of their psi belief or disbelief scores, trained them to run a psi session, and then had them conduct one or more psi sessions. As with the previous studies in the series, all participants were paid in appreciation of their involvement. In the present study, an additional financial reward was given to experimenters whose participants scored at a predetermined level in line with the experimenters' belief. The aim was to encourage a divergence in scoring between believer and disbeliever experimenters. We
administered the same session questionnaire measures as in the earlier studies in the series, to measure participants’ belief in the paranormal, expectations for success, and experimenters’ evaluations. Additionally, for exploratory purposes we took measures of the experimenters’ personality and cognitive ability, to investigate whether these factors correlated with the experimenters’ paranormal belief. The cognitive measures were chosen to be the same as those in several previous studies of cognitive factors in paranormal belief (e.g., Irwin, 1991; Smith, Foster, & Stood, 1998; Watt & Wiseman, 2002), in order to contribute to that literature. Two kinds of cognitive measures would be taken. The first, following Irwin’s work, is a syllogistic reasoning task. Several studies have been done using this task, primarily to test skeptics’ claims that believers in the paranormal may have inferior reasoning ability. The second cognitive measure, Raven’s Matrices, has been used in the past because it is a more general measure of intellectual ability that is reliable and quickly administered.

There were two main hypotheses. Hypothesis 1 predicted that there would be an overall remote facilitation of focusing effect, as indicated by the participant having significantly fewer distractions during help epochs compared with control epochs. Hypothesis 2 predicted that an experimenter effect would be found; that is, those participants tested by believer experimenters would score more highly on the psi task than those tested by disbeliever experimenters.

**METHOD**

**Participant Recruitment**

**Experimenters.** The experimenters had previously taken part in Watt’s studies with Brady and Baker and were selected on the basis of their belief in the paranormal scores obtained during these studies. The 25% with the highest belief scores and the 25% with the lowest belief scores were contacted by letter and phone call and invited to participate for a payment of £5 per session. The experimenters were mostly members of the general public. Experimenters were aware that they had been recruited on the basis of their paranormal belief and knew that the overall purpose of the study was to investigate experimenter effects in parapsychological research.

**Participants.** The participants were recruited mostly by e-mail and through appeals at lectures. They were mostly undergraduate psychology students, who were paid £5 each for participating. They were asked to attend their testing session in pairs and to bring in a friend or family member for this purpose. Participants were aware that it was a parapsychology study.

**Investigators.** Caroline Watt (CW) designed and supervised the study. Peter Ramakers (PR) conducted the study as part of the requirement for his master’s psychology degree while he visited the Koestler Parapsychology Unit, University of Edinburgh, Scotland, as a student coworker. PR recruited, scheduled, and trained participants and was in attendance for every session. PR also acted as experimenter for 3 sessions in which the scheduled experimenter did not turn up.

**Materials**

1. **Administered to Participants**

   **Belief in paranormal questionnaire.** This questionnaire was extracted from the Koestler Parapsychology Unit Participant Information Form. It contained 1 item on participants’ belief that they could demonstrate psi in a controlled experiment and 11 items on belief in and experience of telepathy, clairvoyance, precognition, and psychokinesis. These terms were all defined for the participants. Responses were made by ticking a box on a 7-point scale. High scores indicated higher belief. Scores could potentially range from 12 to 84.

   **Luck questionnaire.** As part of a separate line of investigation, participants were asked to complete this brief two-item questionnaire about their luck beliefs. Results of this are not reported here.

   **Helper’s session questionnaire.** This was intended to obtain qualitative information about the helper’s response to the session and also to check whether the experimenters were perceived as psi-believers or psi-disbelievers. There were two presession questions:

   - How comfortable are you with the idea of your friend remotely helping you to focus? (5-point scale ranging from not at all comfortable to very comfortable).
   - How confident are you of having fewer distractions when your friend is attempting to help you to focus? (5-point scale ranging from not at all confident to very confident).

   There were two postsession questions:

   - To what extent do you feel that you were being helped to focus by your friend? (5-point scale ranging from not at all to completely).
   - How would you rate your experimenter? (a) Warm – Cold; (b) Professional – Unprofessional; (c) Instilling confidence for task – Not instilling confidence for task; (d) Believing in psi – Not believing in psi (each on a 5-point scale, reverse scored).

   There were two additional questions about the helper’s confidence in the hypothesised scenario of having the experimenter be the helper. These are not reported on further here.

   **Helper’s session questionnaire.** This was intended to obtain qualitative information about the helper’s response to the session and consisted of two presession questions:

   - How comfortable are you with the idea of remotely helping your friend to focus? (5-point scale ranging from not at all comfortable to very comfortable).
• How confident are you of being able to help your friend to focus during the help periods? (5-point scale ranging from not at all confident to very confident).

There were two postsession questions:

• To what extent do you feel that you were helping your friend to focus during the help periods? (5-point scale ranging from not at all to completely).

• How would you rate the quality of your own focusing, during the help periods? (5-point scale ranging from completely focused to completely distracted, reverse scored).

2. Administered to Experimenters

Belief in paranormal questionnaire. Experimenters had been administered this questionnaire in a previous study. The questionnaire was identical to that administered to participants.

Personality questionnaire. The 48-item Eysenck Personality Questionnaire (EPQ-R) was used because it is short and reliable. This gave measures of Extraversion, Psychoticism, and Neuroticism. Scores could range from 0 to 12 for each factor.

Raven’s Advanced Progressive Matrices test. Set I was used to assess nonverbal reasoning ability (Raven, Court, & Raven, 1985). This was a 12-item task. For each item, participants were required to indicate which of eight possible symbols correctly completed a sequence of symbols. Participants were given 5 min to complete this task. Scores could range from 0 to 12.

Syllogisms test. This test of reasoning skills contained 24 items and was an abbreviated version of the 48-item test developed by Irwin (1991), omitting the 24 parapsychological items. Each item contained a pair of statements, followed by a conclusion. The participant’s task was to indicate whether the conclusion was valid or invalid, as derived logically from the statements. Participants were given 5 min to complete this task. Scores could range from 0 to 24.

Trainee experimenter guide. This was a three-page step-by-step guide to help the trainee experimenter conduct the session.

Rooms and Apparatus

The experiment took place in four rooms in the Department of Psychology, University of Edinburgh. In the parapsychology experimental suite, participants met the experimenter and completed questionnaires in a reception room. Also in this suite was the helper’s room—a windowless sound-attenuated room with double sound-insulated doors, a reclining chair, a table on which was placed a candle in a blue glass holder and a handheld response button that was connected to a computer in the adjacent control room (where the experimenter was stationed). The computer controlled the helper’s influence schedule, monitored and recorded the helper’s button presses, and recorded the results of each session. The helper’s room was located about 20 m away up a short flight of stairs and along a corridor. This room, which was not sound-insulated, had skylight windows, a reclining chair, headphones, a computer monitor on which the session instructions were displayed, and a similar candle holder to the helper’s.

During the psi task, the monitor in the control room displayed “session in progress,” whereas the monitor in the helper’s room either displayed “HELP” or “CONTROL.” The influence schedule was arranged in four pairs of help-control epochs and four pairs of control-help epochs, with each epoch lasting 60 s, giving a total of 16 min. When the epoch changed, a tone was played over the helper’s headphones, so that the helper could check the monitor for possibly changed instructions. The eight pairs of epochs were presented in a fresh randomised schedule for each session, with counterbalancing within each pair. The randomisation was done by the computer once the program was initiated for each session. The program recorded session details (influence schedule, button presses) and summary statistics consisting of the sum of help presses, the sum of control presses, and the percentage influence score (PIS, which is a ratio calculated from these two sums such that a score of 0.5 represents an equal number of help and control presses, >0.5 represents fewer help than control presses, and <0.5 indicates fewer control than help presses). This PIS measure, as used originally by Braud and Schiltz (1991), can be used as a single outcome psi measure for each session. However, the PIS score has limitations (as discussed in Schmidt, Schneider, Binder, Bärkle, & Walach, 2001; Watt & Brady, 2002), and for present purposes we use it only to enable comparisons with earlier research.

Procedure

Overview. The experiment was in two parts: training sessions and psi sessions. First, PR trained the experimenter how to conduct a psi session with two naive participants. Second, PR scheduled the experimenter alongside the pair of participants, and the experimenter conducted the psi session as trained. For the psi session, participant pairs swapped roles so that each was helper once and helpee once, thus giving two psi trials per session.

Training sessions. PR met trainee experimenters individually. After explaining the purpose of the study, he administered the personality questionnaire and the cognitive questionnaires. With the help of the printed guide, PR then led the trainee experimenter through the procedure of a psi session (all trainee experimenters had previously done such a session as a bona fide participant, so they already had some familiarity with the procedure). Trainee experimenters were encouraged to act naturally during the session and not to deliberately attempt to encourage or discourage psi scoring from their participants. The trainee experimenters then role-played conducting a session with PR. Believer and disbeliever experimenters were informed that they would receive an additional £5 for each one of their participants who
obtained a PIS score of ≥0.6 or ≤0.4, respectively. Trainees were given the printed guide to take home and study prior to the PSI session. Finally, they were paid £5 and PR scheduled a time for them to return to conduct the PSI session.

**PSI sessions.** Experimenters could conduct more than one PSI session, but most conducted only one (see Results section for more details). Having scheduled the experimenter, PR then scheduled the first available pair of participants for the session. During the session, PR remained in the control room to monitor the session and provide guidance if needed. The experimenter first conversed with participants about previous parapsychology experiments and possible PSI experiences they might have had. He or she summarised the procedure and then tossed a coin to see who would be helper first. The experimenter then gave a more detailed description of the roles of helper and helper. The helper was told that when distracted from focusing on the candle, he or she was to press the button to “register” that distraction and then to return attention to the candle. The helper was reminded that at times the friend would be trying to help him or her to focus and was asked to try to feel open to being helped, but not to guess when he or she was being helped as this might be distracting in itself. The helper was told that when the monitor said “Help,” he or she was to focus on the candle and at the same time maintain the mental intention to help his or her friend focus. When the monitor said “Control,” the helper was told to take a break from focusing and to let his or her mind wander.

The experimenter then administered the Luck and Belief questionnaires and the appropriate presession questionnaire. Both participants were then taken to the helper’s room, the helper was seated, and the candle was lit and positioned on a table about 2 m in front of the helper. The helper was given final instructions and was asked to begin the focusing task as soon as he or she was alone, then the lights were extinguished. The helper was then taken to his or her room and was seated, the candle was lit, and the helper was given final instructions. PR then initiated the computer program for the PSI task. Note that the helper was unaware of the precise time when the PSI task began or ended. PR remained in the control room, while the experimenter waited in the reception room for the first PSI session to end and then asked the helper and helper to complete the postsession questionnaires in their respective rooms. The experimenter then brought both participants back to the reception room to swap roles, recap the procedure, and complete the appropriate presession questionnaire. The participants were seated in their respective rooms as before and the PSI task proceeded as before. After the second PSI task ended, the helper and helper completed the postsession questionnaires in their rooms, then returned to the reception room for a brief chat about the session. Finally, the experimenter took the participants to the control room where PR revealed the session results and paid the participants and experimenter, who were then free to leave.

In a small number of sessions, PR acted as experimenter when the scheduled experimenter failed to show up. He also completed the belief, cognitive, and personality questionnaires.

**Results.**

Analyses

Hypothesis 1 (that there would be an overall remote facilitation of focusing effect) was tested by calculating a related $t$ test comparing the sum of help presses with the sum of control presses in each session. We made the one-tailed prediction that there would be fewer button presses in the help epochs compared with the control epochs. PSI scores were compared with MCE as a secondary measure.

Hypothesis 2 (that those participants tested by believer experimenters would score more highly on the PSI task than those tested by disbelief experimenters—again, a one-tailed prediction) was tested by first calculating two related $t$ tests comparing sum of help presses with sum of control presses for the believer condition and for the disbelief condition, and then comparing the Fisher $Z$ corresponding to the associated effect sizes (Cohen’s $r = (Z_1^2 + Z_2^2)^{1/2}$) found for each $t$ test (as described in Rosenthal & Rosnow, 1991, p. 495). This latter comparison produces a $Z$ score and an associated $p$ value. As an exploratory measure, and to enable comparison with previous research, we conducted a second analysis comparing the PSI scores for each condition using an unrelated $t$ test.

Exploratory analyses compared participants’ belief and session questionnaire responses for the two conditions using Mann-Whitney tests. Likewise, exploratory analyses compared the personality and cognitive scores of believer and disbelief experimenters.

As a result of logistical difficulties, progress was slow and the experiment terminated when a total of 18 sessions (36 psi trials) by 14 experimenters had been conducted. The study was thus low-powered and smaller than planned. However, we do not consider optional stopping to apply in this case because the end-point of the study was determined by CW while she was still blind to the results of the testing and because PR’s departure date had been preplanned.

**Experimenters.**

Believer experimenters. There were nine believer experimenters (mean paranormal belief score = 70, $SD = 4.24$), consisting of three males and six females (mean age = 36.9 years, $SD = 11.4$, range = 28–63 years). Seven experimenters conducted 1 session, one conducted 2, and one (PR) conducted 3 sessions, thus giving a total of 12 sessions (24 psi trials). Three
experimenters received one bonus payment because their participant scored at or above PIS = 0.6, and three received two bonus payments because two of their participants scored at this level. There were five disbeliever experimenters (mean paranormal belief score = 28.6, SD = 6.11), consisting of three males and two females (mean age = 44.8 years, SD = 19.02, range = 21-64 years). Four experimenters conducted one session and one conducted two, giving a total of six sessions (12 psi trials). One experimenter received one bonus payment because the participant scored at or below PIS = 0.4, and one received two bonus payments because two participants scored at this level.

As expected, there was a significant difference in the believer and disbeliever experimenters' belief scores (Mann-Whitney Z = 3.01, p < .001, one-tailed, n = 9), thus confirming that our attempt to select trainee experimenters with widely divergent paranormal belief scores had been successful.

Table 1 presents descriptive statistics for the experimenters' cognitive and personality measures. Not surprisingly given the low statistical power, there are no significant differences between the believer and disbeliever experimenters, though there seems to be a trend for the disbeliever experimenters to have higher scores on both cognitive tasks.

<table>
<thead>
<tr>
<th>Measure</th>
<th>Believer experimenter (n = 9)</th>
<th>Disbeliever experimenter (n = 5)</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raven's Advanced Progressive Matrices</td>
<td>7.6 (2.2)</td>
<td>9.8 (2.2)</td>
<td>-1.550</td>
<td>.12</td>
</tr>
<tr>
<td>Syllogisms task</td>
<td>12.2 (2.5)</td>
<td>15.6 (5.3)</td>
<td>-1.142</td>
<td>.25</td>
</tr>
<tr>
<td>Extradimensional</td>
<td>8.8 (3.8)</td>
<td>7.4 (2.6)</td>
<td>1.011</td>
<td>.31</td>
</tr>
<tr>
<td>Psychoticism</td>
<td>3.8 (1.6)</td>
<td>5.4 (1.7)</td>
<td>1.440</td>
<td>.15</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>3.4 (2.7)</td>
<td>5.4 (1.7)</td>
<td>0.416</td>
<td>.19</td>
</tr>
</tbody>
</table>

Table 1: Mean (and Standard Deviation) of Scores for Cognitive and Personality Measures

Note: Scores are based on Mann-Whitney comparisons; p values are two-tailed.

Participants

There was a total of 96 participants, consisting of 21 females and 15 males (mean age = 29.7 years, SD = 5.9, range = 18-43 years). The mean paranormal belief score for participants in the believer experimenter condition was 52.3 (SD = 12.6), which was nonsignificantly different from the mean of 58.1 (SD = 10.7) for participants in the disbeliever experimenter condition. When asked to rate their experimenter's psi belief, participants in the two conditions hardly differed in their ratings (believer experimenters' mean belief rating = 5.4, SD = 0.7; disbeliever experimenters' mean belief rating = 3.6, SD = 1.0; Mann-Whitney Z = -3.08, p = .001, two-tailed, N = 36). In other words, although the experimenters differed significantly on their psi belief, this difference did not seem to be apparent to participants, whose mean ratings of their experimenters' belief fell near the midpoint of the scale.

Experimenter Effects in the Psi Task

Hypothesis 1 predicted an overall remote facilitation of focusing effect. Overall, the mean number of help presses (12.03, SD = 11.34) was significantly lower than the mean number of control presses (13.47, SD = 11.32); related t(35) = 2.088, p = .02, one-tailed, ES = 0.37. This indicates an effect of remote facilitation on the focusing task, with participants showing significantly fewer distractions during the epochs when they were being remotely helped compared with the control epochs. The secondary measure, mean PIS score, was 0.546 (SD = 0.13). This differed significantly from chance, t(35) = 2.153, p = .04, two-tailed.

Table 2: Mean (and Standard Deviation) Button Presses and PIS Scores, and a Comparison of Help (H) and Control (C) Presses

<table>
<thead>
<tr>
<th>Condition</th>
<th>H presses</th>
<th>C presses</th>
<th>t</th>
<th>p</th>
<th>ES</th>
<th>PIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Believer experimenter (n = 24)</td>
<td>12.25</td>
<td>14.54</td>
<td>2.737</td>
<td>.005</td>
<td>.50</td>
<td>.58</td>
</tr>
<tr>
<td>Disbeliever experimenter (n = 12)</td>
<td>11.58</td>
<td>13.33</td>
<td>-0.223</td>
<td>.807</td>
<td>.07</td>
<td>.49</td>
</tr>
</tbody>
</table>

Note: p values are one-tailed. PIS = percentage influence score.

Hypothesis 2 predicted an experimenter effect: Those participants with believer experimenters would have higher psi scores than those with disbeliever experimenters. Table 2 gives the descriptive statistics for the two conditions, together with the results of related t tests comparing mean help and

---

3 In three cases the participants' ages were not obtained, so these were based on estimates made by their friends and classmates.
control presses within each condition, and mean PIS scores. The table shows that the overall significant psi effect in this study is entirely due to those participants with believer experimenters, who have independently significant psi scoring ($ES = 0.50$). Those in the disbeliever experimenter condition scored at chance ($ES = 0.07$). These effect sizes differ in the predicted direction but not to a significant degree ($Z = 1.202, p = .12$, one-tailed). A secondary analysis, which compares PIS scores for each condition using an unrelated test (mean scores are in Table 2), finds a statistically significant difference in the predicted direction: $t(32) = 2.024, p = .05$, two-tailed, $ES = 0.33$. Thus this study finds suggestive evidence of an experimenter effect on the psi task.

**Session Questionnaire Measures**

Table 3 presents descriptive statistics and Mann-Whitney Z scores comparing participants’ responses on the session questionnaire measures, by condition. In almost every case, the participants with believer experimenters gave nonsignificantly lower ratings of prior comfort and confidence, feelings of helping, ratings of the experimenter, and quality of the helper’s focusing. Thus, although an experimenter effect was found for the psi task, there is no evidence of an experimenter effect on the session questionnaire measures.

**Table 3**

**Session Questionnaire Results (Mean scores, Standard Deviation in Parentheses), Comparing Participants With Believer Experimenters With Those With Disbeliever Experimenters**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Believer experimenter</th>
<th>Disbeliever experimenter</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Helper questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Prior comfort</td>
<td>4.4 (0.9)</td>
<td>4.7 (0.5)</td>
<td>-0.592</td>
<td>.58</td>
</tr>
<tr>
<td>2. Prior confidence</td>
<td>2.8 (1.2)</td>
<td>3.5 (0.5)</td>
<td>-1.585</td>
<td>.11</td>
</tr>
<tr>
<td>3. Feeling helped</td>
<td>2.9 (1.0)</td>
<td>3.1 (0.8)</td>
<td>-0.615</td>
<td>.54</td>
</tr>
<tr>
<td>4. Experimenter ratings</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Warmth</td>
<td>3.8 (0.9)</td>
<td>4.2 (0.6)</td>
<td>-1.146</td>
<td>.16</td>
</tr>
<tr>
<td>Professionalism</td>
<td>3.4 (1.0)</td>
<td>4.0 (1.0)</td>
<td>-1.706</td>
<td>.09</td>
</tr>
<tr>
<td>Instilling confidence</td>
<td>3.3 (1.1)</td>
<td>3.8 (0.9)</td>
<td>-1.377</td>
<td>.17</td>
</tr>
<tr>
<td><strong>Helper questionnaire</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Prior comfort</td>
<td>4.1 (1.0)</td>
<td>4.7 (0.5)</td>
<td>-1.553</td>
<td>.12</td>
</tr>
<tr>
<td>2. Prior confidence</td>
<td>2.9 (1.3)</td>
<td>3.4 (0.9)</td>
<td>-1.354</td>
<td>.18</td>
</tr>
<tr>
<td>3. Feeling helped</td>
<td>3.2 (0.9)</td>
<td>3.0 (0.7)</td>
<td>0.645</td>
<td>.52</td>
</tr>
<tr>
<td>4. Quality focusing</td>
<td>3.0 (1.0)</td>
<td>3.2 (0.8)</td>
<td>-0.443</td>
<td>.66</td>
</tr>
</tbody>
</table>

*Note. Z scores are based on Mann-Whitney comparisons; p values are two-tailed.*

**Discussion**

This study forms the latest in a series (Watt & Baker, 2002; Watt & Brady, 2002) that examined experimenter effects using as a psi task the remote facilitation of attention focusing. We investigated the effect of the experimenter’s psi belief on the outcome of the psi task by pre-selecting multiple experimenters on the basis of their belief or disbelief in the paranormal. We found an overall remote facilitation of focusing effect, with $ES = 0.33$. This replicates the main findings of two previous studies using the same psi task (Brady & Morris, 1997, $ES = 0.27$; Bradt et al., 1995, $ES = 0.35$), though another two previous studies using this task (Watt & Baker, 2002, $ES = 0.12$; Watt & Brady, 2002, $ES = -0.11$) did not find a psi effect. On closer inspection it could be seen that the effect in the present study was entirely due to participants who had been tested by believer experimenters. This group obtained independently significant scoring on the psi task, whereas those tested by disbeliever experimenters scored at chance. This suggests that the “multiple experimenter” method demonstrated an experimenter effect on the psi task, though, replication with greater numbers of experimenters in each condition is recommended. Taken together with Parker’s (1975) findings of significant differences in psi scoring between believer and disbeliever experimenters, these results indicate the potential utility of this kind of multi-experimenter design for the study of experimenter effects.

Participants’ responses on the questionnaire measures indicated little difference in their perceptions of the experimental task and of the experimenter, regardless of whether they had a believer or disbeliever experimenter. Likewise, the experimenters differed little in the personality and cognitive measures (though this comparison suffers from low power). Watt’s previous focusing studies (with Brady and Baker) found no psi effect, but participants’ questionnaire measures seemed to indicate that participants were responding differently to the experimenters. Taken together with these earlier findings then, the present experimenter effect could suggest an experimenter psi effect, because the psychological measures seem to show no difference between conditions.

However, it would be premature to conclude on the basis of these three studies that the present experimenter effect is due to experimenter psi rather than to an experimenter interaction effect. The experimental system is extremely complex, and further research may reveal more subtle interactions. For instance, perhaps the session questionnaire measures were insufficiently sensitive instruments to detect differences in how the participants were responding to experimenters. Certainly, with only 12 participants in one condition and 24 in the other, any individual differences would have to be quite gross to be detected. In addition, it is quite possible that the believer experimenters are encouraging their participants to score positively on the psi task in a way that is not being detected by the present
session questionnaire measures. So, our effect may not be due to experimenter psi, but to participant psi being elicited more effectively by the believer experimenters. Finally, needless to say, experimenter psi and participant psi need not be mutually exclusive factors.

The cognitive measures taken in this and much previous research into experimenter effects and paranormal belief have focused on reasoning skills and general intellectual ability. However, this is in part a historical hangover from studies investigating skeptics' claims that paranormal belief is associated with inferior reasoning skills. For future research, we would recommend a different starting point, where we consider measures that might tap into the quality of the interaction between experimenter and participant. Intellectual ability may still be part of it. However, there is growing interest, in both theoretical and applied psychology, in the concept of emotional intelligence. Although there is still considerable academic debate about the validity of this construct (e.g., Carrozza, Chan, & Caputi, 2000a, 2000b; Davies, Stanek, & Roberts, 1998; Petrides & Furnham, 2000), it is claimed to indicate the ability to perceive, understand, and manage emotions. It may predict social competence and appears to be unrelated to intellectual ability (Bar-On & Parker, 2000). We would suggest that these skills might be of considerable importance in the experimenter-participant interaction. For future research, then, we would recommend the inclusion of a measure of emotional intelligence.

The most important advantage of this multi-experimenter design over previous studies (e.g., Watt & Baker, 2002; Watt & Brady, 2002; West & Fisk, 1953; Wiseman & Schütz, 1997, 1998) is that it allows for a large number of experimenters to interact with their participants in a natural manner in a single study. This design also enables researchers to measure the experimenters' individual differences on characteristics such as personality and cognitive ability, measures that would be uninformative when dealing with only one or two experimenters.

We also hoped that the present study, by selecting participants on the basis of their relatively strong belief or disbelief in the paranormal, might create a more realistic research scenario than a single role-playing experimenter (as used in Watt & Baker, 2002). However, one could argue that despite this, our trainee experimenters had little intrinsic motivation to obtain a particular pattern of results, compared with career researchers who might have a stronger motivation to find either positive or null psi results. We attempted to increase this motivation by offering a financial incentive for experimenters obtaining results in line with their belief, in the hope that this would widen the difference in scoring between believer and disbeliever experimenters. Although several experimenters obtained this financial incentive, anecdotally, it was repeatedly in the hope that this would widen the difference in scoring between believer and disbeliever experimenters. Although several experimenters obtained this financial incentive, anecdotally, they often commented to PR that they had forgotten about it so their motivation may not have been strongly affected by the financial incentive. However, they were prepared to make two separate visits to the research unit and to give 3 hr or more of their time for the visits to the research unit and to give 3 hr or more of their time for the study, suggesting some intrinsic motivation to take part. There has been very little research on the effect of offering a financial incentive, but the only study of which we are aware (Palmer & Miller, 1972) found no difference in ESP scoring between sheep and goat participants with prior knowledge of a financial prize for the highest scorer, whereas sheep and goat participants who were unaware of the prize did show a difference in ESP scoring in the expected direction. Thus, the monetary incentive appeared to inhibit the sheep-goat effect. In the case of the present study, then, given that we did find a sheep-goat effect (in terms of the different scoring obtained by those tested by believer experimenters compared with disbeliever experimenters) with a monetary incentive, there remains the possibility that our effect might have been even stronger without this incentive.

Further research using a similar design but a larger number of believer and disbeliever experimenters may help parapsychologists to understand what factors may contribute to parapsychology's experimenter effect. One would also want to move towards multifactorial analysis, rather than the relatively simple statistical approach used in this area. However, we found that, despite being able to pay participants in appreciation of their involvement, there were considerable practical difficulties in recruiting, training, and scheduling experimenters alongside pairs of participants. Indeed it was these difficulties that necessitated the premature termination of the study. This kind of design is therefore quite a challenge to parapsychologists' limited resources. Alternatively, a study using multiple genuine believer and disbeliever career researchers would be desirable, but again would face many logistical difficulties.

References


Department of Psychology
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ
Scotland, UK
cwatt@afbl.ssc.ed.ac.uk