The Journal of Parapsychology, Vol. 66, December 2002 (pp.)

EXPERIMENTER DIFFERENCES IN COGNITIVE CORRELATES OF PARANORMAL BELIEF AND IN PSI

By Caroline Watt and Richard Wiseman

ABSTRACT: It has been claimed that experimenter effects may account for inconsistent findings in the study of cognitive correlates of paranormal belief and in psi research. The present study investigates these 2 strands by having 2 experimenters each administer to 30 participants a paranormal belief questionnaire, 2 tests of cognitive ability (a syllogistic reasoning task and Raven's Progressive Matrices), and an ESP task. For all 60 participants, a significant negative correlation was found between paranormal belief and syllogisms performance. This correlation was attributable to just 1 of the experimenters, and the experimenters' belief–cognitive ability correlations significantly differed, thus demonstrating an experimenter effect for this measure. Additional post hoc analyses were conducted to clarify the mechanism underlying the belief–cognitive ability correlation. There was no evidence of an experimenter effect for the ESP task.

Experimenter effects, that is, experimenters' influence on their research participants and subsequently on the outcome of their research, have a long history in psychology (e.g., Harris & Rosenthal, 1985; Rosenthal, 1967, 1976, 1990; Rosenthal & Rubin, 1978). Many different factors may play a role in experimenter effects, such as expectancy and belief (e.g., Luborsky et al., 1999) and personality (e.g., Cooper & Hazelrigg, 1988; Hazelrigg, Cooper, & Strathman, 1991).

The present study builds on this work by exploring the possible existence of experimenter effects within two important areas: (a) the psychology of belief in the paranormal and (b) psi research. The experiment is made up of two strands, and the background to each aspect of the study is described in turn.

Cognitive Correlates of Belief in the Paranormal

Some psychologists have examined whether belief in the paranormal correlates with performance on various cognitive tasks, including critical

We are grateful to the Fundaçao Bial for funding this study. We thank Emma Greening and Ciarán O'Keefe for assisting with participants in the study. Thanks also to Professor Harvey Irwin and Professor Ian Deary for their helpful comments on an earlier draft of this article. We also thank our referees, whose comments have helped us to revise the article.

thinking, reasoning skills, and IQ (see reviews by French, 1992; Irwin, 1991, 1993, 1999). Although findings are mixed, the overall trend is that people who disbelieve in the paranormal tend to outperform believers on these tasks (Blackmore & Troscianko, 1985; Roberts & Seager, 1999; Wiezerbiki, 1985), and Irwin named this the *cognitive deficits* hypothesis. However, many of these studies have been conducted by experimenters who are skeptical about the existence of the paranormal (e.g., Alcock & Otis, 1980), and some experimenters who are more open to the notion of ESP have failed to replicate these results (e.g., Irwin, 1991; Roe, 1999). This has led some researchers (e.g., Irwin, 1991; Smith, Foster, & Stovin, 1998) to hypothesise that these inconsistent results may be due, at least in part, to the experimental context influencing participants' reports concerning their level of belief in the paranormal and/or their performance on cognitive tasks.

Research investigating how the belief–cognitive ability relationship may be influenced by experimental context has had mixed results. Merla-Ramos (2000) found that paranormal believers tended to have poorer syllogistic reasoning ability than disbelievers, but only for syllogisms that included a paranormal or religious content. No belief–reasoning correlation was found for syllogisms that had a neutral content. Irwin (1991) used a similar syllogistic reasoning task, including neutral, proparanormal, and antiparanormal syllogisms. Overall he found a nonsignificant syllogisms–belief correlation, and the content of the syllogisms did not affect the correlation. Smith et al. (1998) varied the context in which participants completed a paranormal belief questionnaire and did an intelligence test (Raven's Advanced Progressive Matrices), with some participants being given an ESP-supportive context and some an ESP-unsupportive context. This study found that participants expressed greater paranormal belief in the ESP-supportive condition. Overall a negative correlation was found between belief and performance on the matrices task. However, contrary to prediction, context did not appear to affect this correlation.

What are we to make of this inconsistent body of findings? Many of these experiments may differ in the kinds of participant population they are drawing upon, in the kinds of procedures they use, and in the experimental setting. In addition, all of these studies used a single experimenter, and it is possible that differences in the experimenter's characteristics, from one study to another, may be a factor in producing inconsistent results. For instance, experimenter–participant dynamics may vary depending on the gender of the experimenter, his or her personality, his or her beliefs about the paranormal, or a host of other possible factors. With such a complex web of factors, it is difficult to begin to untangle them to answer the question of which may be important in eliciting certain patterns of results.

The present study attempts to take a first step towards untangling the web by asking a very simple question: When one keeps constant the participant population, experimental setting, and procedures but uses two different experimenters, is it possible for the two experimenters to obtain different patterns of results? Stated another way, the present study investigates the proposition that inconsistent findings in the study of cognitive correlates of paranormal belief may be attributed to (as yet unidentified) differences between experimenters.

In the present study, two different experimenters (CW and RW) carry out a joint experiment examining whether belief in the paranormal correlates with performance on two cognitive tasks. To build on the previous literature reviewed earlier, the two tasks selected to assess cognitive ability are a syllogistic reasoning task (as used by Irwin, 1991) and Raven's Progressive Matrices (as used by Smith et al., 1998). To facilitate comparisons with these earlier studies, we also use the same belief questionnaire as used by Irwin and Smith et al., and we use correlational analyses as they did. Hypothesis 1 is therefore that the experimenters will differ significantly in the correlation each finds between cognitive ability and paranormal belief.

Experimenter Effects in Psi Research

The second strand of the present study concerns the question of experimenter effects in psi research. Certain parapsychologists may be "psi-facilitators" because they consistently carry out studies that obtain evidence for psi, whereas others may be "psi-inhibitors" because they consistently obtain chance results (Schmeidler, 1997). Researchers have suggested a variety of different ways in which experimenter effects may operate (see reviews by Kennedy & Taddonio, 1976; Palmer, 1989a, 1989b, 1997).

Some have suggested that psi-facilitators may be carrying out their experiments with psychic participants or are using experimental procedures that enhance participants' psychic ability. For example, a study by Honorton, Ramsey, and Cabbibo (1975) used experimenters who either interacted with participants in a friendly, casual, and supportive way or interacted in an abrupt, formal, and unfriendly way. Participants who had experienced the positively-toned interactions had significantly higher scores on an ESP test than those who had experienced negatively-toned interactions. A similar method was used more recently by Schneider, Binder, and Walach (2000) in an EDA-DMILS (electrodermal activity–direct mental interactions with living systems) experiment. Participants either experienced a *personal* condition, in which the experimenter tried to create a psi-conducive atmosphere, or a *neutral* condition, in which there was minimal interaction with an experimenter and participants received most of their instructions from a computer. Somewhat contrary to Honorton et al.'s findings, Schneider et al.'s study found that the effect size for the neutral condition was three times larger than that for the personal condition.

Experimenter effects have also been found in the area of remote staring detection. In her research into remote staring detection, Marilyn Schlitz (MS) has obtained evidence of psi (Schlitz & LaBerge, 1994). In contrast, Richard Wiseman (RW) has obtained no evidence of psi in his remote staring detection studies (Wiseman & Smith, 1994; Wiseman, Smith, Freedman, Wasserman, & Hurst, 1995). To help determine why they had obtained such different results, MS and RW then carried out two joint studies, using the same participant pool, location, equipment, and procedure. Results revealed evidence of an experimenter effect, with MS's participants showing a significant difference in EDA during stare periods compared with no-stare periods, whereas RW's trials were at chance (Wiseman & Schlitz, 1997, 1999).

The findings of the joint Wiseman and Schlitz studies suggest that their pattern of results may not be due to differences in participant population or experimental procedures. Whether these effects are due to experimenter psi or are related to the interaction between experimenter and participant has yet to be determined. Studies using multiple psi-believer and psi-disbeliever experimenters (Parker, 1975; Watt & Ramakers, in press) seem to suggest that experimenter psi belief may be an important factor, and it is well known that MS and RW differ in psi belief. However, the two differ on many other factors in addition to their belief.

There is still a long way to go before parapsychologists can identify which personal characteristics of the experimenter, such as gender, personality, and belief in the paranormal, may be important in eliciting psi effects. A questionnaire study by Smith and Gordon (2002) found quite a wide range of opinions about which parapsychologists are psi-conducive or psi-inhibitory. In addition, perceived psi-conduciveness was found to be unrelated to experimenters' self-ratings of personality, belief that ESP is possible, belief in their own ESP, personal psi experiences, and practise of a mental discipline. The literature on experimenter effects in parapsychology therefore paints a confusing picture.

The present study uses two experimenters who differ in their attitudes towards psi and in their psi research track record. RW is skeptical about the existence of psi (e.g., Wiseman, 1997), and his experimental research into psi has obtained null results (see, e.g., the remote staring detection studies referenced earlier). CW is more positive about the existence of psi and has conducted a number of studies that have obtained positive psi results (e.g., Watt, 1996; Watt & Morris, 1995; Watt & Ramakers, in press; Watt, Ravenscroft, & McDermott, 1999). The present study may be viewed as a conceptual replication of the Wiseman/Schlitz studies, as it is comparing two researchers who appear to differ in psi-conduciveness. Hypothesis 2 is therefore that the experimenters will obtain a significant difference in the outcome of the psi test. However, there are many differences between CW and MS. For instance, although CW has published positive psi results several times, she does not seem to have the rate of success that MS has, and the present study also uses a completely different psi task from the Wiseman/Schlitz studies. Therefore the present study outcome is not designed to comment directly on the Wiseman/Schlitz findings.

In sum, both strands of the present study are examining a similar fundamental question: Can one observe experimenter differences in scoring either in the cognitive correlates of paranormal belief (Hypothesis 1) or in performance at a psi task (Hypothesis 2) when using the same participant pool and experimental procedures?

Method

Participants

The volunteer participants were mostly undergraduate students at the University of Hertfordshire, Hertfordshire, U.K., who received either £5 or course credit for taking part. Thirty participants were tested by RW and 30 were tested by CW. There were 19 male and 40 female participants (sex data were not available for 1 participant); mean age was 24.9 years (SD = 8.2; age data were not available for 3 participants).

Materials

Belief in the paranormal questionnaire. This was Tobacyk's Revised Paranormal Belief Scale (Tobacyk, 1988), a 26-item questionnaire containing items measuring a variety of paranormal beliefs. Participants respond to each item on a 7-point scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Scores could range from a minimum of 26 to a maximum of 182. The overall scale consists of seven subscales. However, there has been debate over the construct validity and the factorial structure of this belief scale (Lange, Irwin, & Houran, 2000; Lawrence, 1995; Tobacyk & Thomas, 1997), therefore we only report overall paranormal belief scores, not subscale scores.

Raven's Matrices test. Set 1 of the Raven's Advanced Progressive Matrices 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 a minimum of 0 to a maximum of 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). Because of time constraints, the 24 parapsychological items in the original test were omitted. 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 a minimum of 0 to a maximum of 24.

ESP targets. A single target pool of five video clips, each approximately 2 min in duration, was used for the ESP task. Each clip was on a separate video cassette labelled A to E. Each clip had a different instrumental musical soundtrack.

Testing room. The experiment took place in a laboratory in the Department of Psychology at the University of Hertfordshire. The room contained comfortable seating and three unobtrusive remotely controlled cameras, two of which were used to film the experimenter and the participant during the session. This filming was monitored and controlled by a technician in an adjacent sound-attenuated control room.

Procedure

ESP target randomisation procedure. Prior to the study beginning, a person otherwise unconnected with the study used published random number tables to designate a target clip for each participant (Rand Corporation, 1955). A calculator RND function was used three times to select (a) a page of random numbers (1 of 10); (b) a column of numbers; and (c) a row, to give an entry point to the table. The randomiser read from the selected entry point until a digit from 1 to 5 was encountered. This digit was then transformed to the letters A to E according to a rotating code, such that even if a single digit unexpectedly occurred more frequently than others throughout the study, this would translate to different target identities. The written records of designated targets for each session were individually sealed in double envelopes to keep experimenters and participants blind to the target identity. The duplicate record of designated targets was sealed by the randomiser in a plastic tamper-evident security bag. To avoid leakage of information about targets, the randomiser had no contact with the experimenters or anyone else involved in the experiment after the randomisation had taken place.

Session procedure. Participants were allocated to experimenters in a counterbalanced fashion, by signing themselves up against different time slots on a poster on a noticeboard. The time slots gave no indication of who the experimenter would be for any session. There could be a maximum of 3 participants per session. At the appointed time, an assistant brought the participants to the experimenter in the testing room. The experimenter introduced himself or herself and briefly explained the procedure to participants and had them sign a consent form.

Seated so that their responses remained private, participants next completed the belief questionnaire, followed by either the syllogisms task or the matrices task, in a counterbalanced fashion. Following completion of the questionnaire measures, 1 participant remained with the experimenter in the testing room while an assistant took the remaining 2 participants to a distant waiting area. The experimenter dimmed the lights and seated the participant in a comfortable chair for the ESP task.

ESP task. The ESP task was discreetly filmed, with the participants' permission. The participant sat facing a video screen, and the experimenter sat adjacent to the participant. The experimenter presented the task as a precognitive one, that is, that the participant would in a few minutes be viewing the randomly selected target clip. The participant was asked to give his or her impressions of the clip that would shortly be viewed. This was done without any prior relaxation or altered state induction. Instead, the target mentation was recorded in the form of a brief interview between the experimenter and the participant. The experimenter asked questions to elicit additional information from the participant and wrote down the participant's responses. After about 5 min of interview, the experimenter judged the mentation items against the five target possibilities and ranked the targets in order of similarity to the mentations. The experimenter's ranking choice was written on the session record sheet and was stated aloud so that it was recorded on video. Then the experimenter retrieved and opened the envelope containing the target identity (letter A to E). Video recording ended once the experimenter had revealed the target identity and noted it on the session record sheet. The experimenter then selected and played the appropriate video cassette containing the target clip. The ESP task took about 10 min, including viewing of the target clip. The participant never saw any of the other target possibilities. The experimenter then debriefed the participant and asked the participant not to discuss the target clip with other potential participants, and the participant was then free to leave. The experimenter then brought the 2nd participant to the testing room. The ESP task took place as already described. Finally, the experimenter brought the 3rd participant to the testing room for the ESP task. The session ended after this participant had completed the ESP task.

The distribution of participant groups was equal for each experimenter: Each did seven sessions with 3 participants, four sessions with 2 participants, and one session with 1 participant. There were no dropouts from the study.

Results

The experiment terminated when, as planned, each experimenter had tested 30 participants.

Missing Data

One participant did not complete the syllogisms task, owing to language difficulties, therefore $N = 59$ for this measure. One participant misunderstood the instructions for doing the matrices task, so $N = 59$ for this measure. On double-checking whether the correct target had been played during each session, we found that in two sessions the wrong target had been played because CW had misread the handwritten target identification letter. We decided to exclude the data from these two sessions from the psi analyses.

Individual Differences Measures

Table 1 presents mean scores for the belief in the paranormal questionnaire, overall and broken down by experimenter. There was no difference between the experimenters' belief scores. The table also presents data for the matrices and syllogisms tasks. Though RW's participants tended to have higher scores on the cognitive tasks than CW's, this difference was not statistically significant.

Table 1
Mean (and Standard Deviation) Paranormal Belief, Matrices, and Syllogisms Scores

Measure	Overall	CW	RW	*t*	*df*
Paranormal belief	97.7 (24.1)	97.6 (24.3)	97.9 (24.4)	0.04 (.97)	58
Matrices	8.9 (1.8)	8.6 (2.1)	9.1 (1.3)	0.90 (.37)	57
Syllogisms	12.0 (3.5)	11.4 (3.8)	12.8 (3.0)	1.54 (.13)	57

Note. The fourth column gives *t* tests of the difference in scores for CW and RW, with two-tailed *p*-values in parentheses.

Experimenter Differences for Cognitive Tasks and Belief in the Paranormal

To explore the question of whether the two experimenters would obtain different patterns of correlation between participants' paranormal belief and performance on the cognitive tasks, we calculated Pearson correlation coefficients (see Table 2). Overall, there was no correlation between performance on the Raven's Matrices and belief in the paranormal, neither was there any correlation for the experimenters individually. This does not confirm the previous finding by Smith et al. (1998) of a negative correlation between paranormal belief and the Raven's Matrices scores. As Table 2 shows, overall, there was a statistically significant negative correlation between belief in the paranormal and performance on the syllogisms task. That is, those with greater belief in the paranormal tended to have lower scores on the syllogisms task. The breakdown of correlations by experimenter reveals that this correlation was almost entirely due to CW's participants. RW's participants scored close to chance, whereas CW's had an independently significant negative correlation between belief and syllogisms scores. The difference between the experimenters' correlations (calculated using Fisher *z* transformation) was statistically significant ($p = .03$, two-tailed) for the syllogisms task. This suggests an experimenter effect for the correlation between paranormal belief and performance on the syllogisms task, and therefore confirms Hypothesis 1.

Table 2
Correlation Between Paranormal Belief and Cognitive Task Performance

Belief Correlated With	Overall	CW	RW
Matrices	–.04 (.74)	–.06 (.74)	–.02 (.94)
Syllogisms	–.28 (.03)	–.45 (.01)	–.08 (.70)

Note. The figures represent Pearson correlation coefficients, with two-tailed *p*-values in parentheses.

ESP Task

Table 3 shows the distribution of target ranks for the 58 psi trials (with Rank 1 indicating the target judged to be most similar to the participant's mentation). The planned analysis for the ESP task was sum-of-ranks, chosen because it was considered to be a more sensitive measure of psi than a direct hits analysis. This was calculated for the data in Table 3 using the formula in Solfvin, Kelly, and Burdick (1978, p. 99). The overall mean target rank was 2.9 (SD = 1.54), and the sum-of-ranks was 165 (compared with MCE sum-of-ranks of 174), which gives a near-chance z = .07 (adopting the convention of giving a positive sign to the z when the data are in the ESP-hitting direction). Therefore there is no evidence that the target video clip was identified more often than chance.

Table 3
Number of Times Rank Allotted to Target

Experimenter	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
CW (N = 28)	6	6	4	6	6
RW (N = 30)	11	4	5	3	7
Total	17	10	9	9	13
MCE	11.6	11.6	11.6	11.6	11.6

Note. The total number of psi trials is 58, not 60, because two trials were excluded as a result of a procedural error.

Experimenter Differences for ESP Task

Table 3 shows the distribution of target ranks separately for the two experimenters. The mean target rank for CW was 3.0 (SD = 1.5), compared with a mean of 2.7 (SD = 1.6) for RW. This difference was not statistically significant on a Mann–Whitney test (z = –.90, p = .37, two-tailed). Therefore there is no indication of an experimenter effect on ESP results and no support for Hypothesis 2.

Post Hoc Analyses: Investigating Possible Mechanisms for Experimenter Differences

How are we to interpret the observed pattern of a negative correlation between belief and syllogisms performance for CW but not for RW? Perhaps the cognitively skilled participants are inhibiting their admission of paranormal belief when tested by different experimenters (Irwin, 1991). However, CW's and RW's participants did not differ in their paranormal belief scores, which does not support the hypothesis of shifting belief scores.

Perhaps participants are shifting their syllogisms scores in response to their experimenter. RW's participants had a mean syllogisms score of 12.8, which is nonsignificantly higher than the mean of 11.4 for CW's participants, $t(57)$ = 1.545, p = .13, two-tailed. This may suggest that some participants have slightly shifted their syllogisms performance, perhaps due to increased or decreased motivation to do well. Therefore in terms of overall mechanisms underlying the belief–cognitive performance correlation, our data may indicate that participants are varying their performance on the syllogisms task rather than varying their responses on the belief questionnaire.

To investigate this question further, we divided all 60 participants into believers and disbelievers on a median split. Analyses verified that there was no indication that participants had shifted their belief scores for the different experimenters because there was no significant difference between the belief scores of RW's believers and CW's believers: RW mean belief = 114.6, SD = 15.5; CW mean belief = 118.2, SD = 14.2; $t(28)$ = –0.67, p = .51, two-tailed. Similarly, there was no significant difference between the belief scores of RW's disbelievers and CW's disbelievers: RW mean belief = 78.8, SD = 17.9; CW mean belief = 79.6, SD = 14.8; $t(28)$ = –0.13, p = .90, two-tailed.

To identify whether believers or disbelievers were shifting their syllogisms performance, we compared the syllogisms scores of CW and RW's believers and disbelievers. If the paranormal believers shifted their syllogisms scores, then we would expect CW's believers to significantly differ from RW's believers in their syllogisms scores. We would also expect to see no significant difference between the syllogisms scores of CW's disbelievers and RW's disbelievers. As Table 4 shows, this is exactly what was found.

Conducting the same post hoc analysis for believers' and disbelievers' scores on the Raven's Matrices, we found the same pattern of scoring, suggesting a shift in believers' performance on the matrices task (see Table 4). CW's paranormal believers had lower matrices scores than did RW's believers. There was little difference between the matrices scores of the disbelievers. This post hoc analysis gives some indication of an experimenter effect even for the Raven's Matrices task, although the planned correlational analyses did not show any evidence of an experimenter effect or of a belief–matrices performance correlation. Perhaps, with a small potential range of scoring possible for the matrices task, relative to the syllogisms task, the planned correlational analysis was insufficiently sensitive to detect any belief-matrices relationship. The post hoc analysis on the other hand seems to show a tendency for the same pattern of scoring as was found for the syllogisms task.

TABLE 4
MATRICES AND SYLLOGISMS SCORES FOR BELIEVERS AND DISBELIEVERS

Measure	CW	RW	*t*	*df*
Believers syllogisms	10.3 *n* = 14, *SD* = 2.2	12.6 *n* = 16, *SD* = 3.4	2.16 (.04)	28
Disbelievers syllogisms	12.3 *n* = 16, *SD* = 4.7	13.0 *n* = 13, *SD* = 2.6	0.47 (.64)	27
Believers matrices	7.9 *n* = 14, *SD* = 2.3	9.2 *n* = 16, *SD* = 1.0	1.97 (.06)	28
Disbelievers matrices	9.3 *n* = 15, *SD* = 1.8	8.9 *n* = 14, *SD* = 1.5	–0.65 (.52)	27

Note. *p*-values (in parentheses) are two-tailed.

DISCUSSION

This study explored whether, when working with the same participant pool and experimental procedures, two experimenters can obtain different patterns of scoring for cognitive correlates of paranormal belief and for performance on a psi task. Overall, a significant negative correlation was found between paranormal belief and syllogisms performance. This correlation was attributable to CW's and not RW's participants, and the two experimenters' correlations differed significantly, thus supporting Hypothesis 1. This suggests that some of the inconsistencies in previous literature on the cognitive correlates of paranormal belief may be due to as yet unidentified experimenter differences rather than to differences in participant pool or experimental procedures. Post hoc analyses found some evidence of a similar experimenter effect on the matrices task. Our data indicate that it is the believers rather than the disbelievers who are responding differentially to the two experimenters. However, we cannot at this stage say if it is RW's or CW's believers who are changing their cognitive performance, or if both groups might have shifted to some degree.

Hypothesis 2, of a difference between the experimenters in the outcome of the psi task, was not supported. The experimenters in this study clearly differ on some factors that are likely to be important in eliciting psi, such as psi research track record, gender, and personality. However, perhaps they do not sufficiently differ on other potentially relevant factors such as paranormal belief and social skills. Any study with only two experimenters will find it difficult to go beyond a basic demonstration of experimenter differences in scoring. To begin to draw inferences about what it is that makes an experimenter psi-conducive, one will need more complex designs, for instance using multiple experimenters (e.g., Watt & Ramakers, in press). On a methodological note, we had to discard the data from two psi sessions because of an unanticipated problem with reading the randomiser's handwritten records. Losing these two sessions had no effect on the study's conclusions. However, it is a reminder that despite careful planning, pitfalls can remain.

Many commentators have noted the inconsistent findings of the literature researching the link between belief in the paranormal and cognitive ability. Some have suggested that these inconsistencies may be due to experimenter effects. Our study is the first to provide evidence of an experimenter effect in investigating the belief–cognitive ability relationship. Future research will be needed to uncover the mechanisms underlying this effect. The mechanisms are likely to be complex. For instance, it has been suggested that critical thinking ability may be differentially applied in different domains and that attitudes and personality traits may be important factors in moderating the generalizability of critical thinking skills (Royalty, 1995). Research into interpersonal expectancy effects suggests that the experimenter's motivation for control and the participant's need for social approval may both be important moderating factors in the operation of such effects (Hazelrigg et al., 1991). In the meantime, it is vital that parapsychologists and psychologists researching this area realize that their participants' performance may be affected by the experimental context. The experimenter's own beliefs and idiosyncracies in interacting with participants may affect participants so as to elicit, or to obscure, a relationship between paranormal belief and cognitive ability.

References

Alcock, J. E., & Otis, L. P. (1980). Critical thinking and belief in the paranormal. *Psychological Reports*, **46**, 479–482.

Blackmore, S., & Troscianko, T. (1985). Belief in the paranormal: Probability judgements, illusory control, and the "chance baseline shift." *British Journal of Psychology*, **76**, 459–468.

Cooper, H., & Hazelrigg, P. (1988). Personality moderators of interpersonal expectancy effects: An integrative research review. *Journal of Personality and Social Psychology*, **55**, 937–949.

French, C. C. (1992). Factors underlying belief in the paranormal: Do sheep and goats think differently? *The Psychologist: Bulletin of the British Psychological Society*, **5**, 295–299.

Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, **97**, 363–386.

Hazelrigg, P., Cooper, H., & Strathman, A. J. (1991). Personality moderators of the experimenter expectancy effect: A reexamination of five hypotheses. *Personality and Social Psychology Bulletin*, **17**, 569–579.

Honorton, C., Ramsey, M., & Cabibbo, C. (1975). Experimenter effects in extrasensory perception. *Journal of the American Society for Psychical Research*, **69**, 135–139.

Irwin, H. (1991). Reasoning skills of paranormal believers. *Journal of Parapsychology*, **55**, 281–300.

Irwin, H. J. (1993). Belief in the paranormal: A review of the empirical literature. *Journal of the American Society for Psychical Research*, **87**, 1–39.

Irwin, H. (1999). *An introduction to parapsychology* (3rd ed.). Jefferson, NC: McFarland.

Kennedy, J. E., & Taddonio, J. L. (1976). Experimenter effects in parapsychological research. *Journal of Parapsychology*, **40**, 1–33.

Lange, R., Irwin, H. J., & Houran, J. (2000). Top-down purification of Tobacyk's Revised Paranormal Belief Scale. *Personality and Individual Differences*, **29**, 131–156.

Lawrence, T. R. (1995). How many factors of paranormal belief are there? A critique of the Paranormal Belief Scale. *Journal of Parapsychology*, **59**, 3–25.

Luborsky, L., Diguer, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., et al. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, **6**, 95–106.

Merla-Ramos, M. (2000). Belief and reasoning: The effects of beliefs on syllogistic reasoning. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, **61**, 558.

Palmer, J. (1989a). Confronting the experimenter effect: Part 1. *Parapsychology Review*, **20**(4), 1–4.

Palmer, J. (1989b). Confronting the experimenter effect: Part 2. *Parapsychology Review*, **20**(5), 1–5.

Palmer, J. (1997). The challenge of experimenter psi. *European Journal of Parapsychology*, **13**, 110–125.

Parker, A. (1975). A pilot study of the influence of experimenter expectancy on ESP scores [Abstract]. In J. D. Morris, W. G. Roll, & R. L. Morris (Eds.), *Research in parapsychology 1974* (pp. 42–44). Metuchen, NJ: Scarecrow Press.

Rand Corporation. (1955). *A million random digits and 100,000 normal deviates.* New York: Free Press.

Raven, J. C., Court, J. H., & Raven, J. (1985). *A manual for Raven's Progressive Matrices and Vocabulary Scales.* London: H. K. Lewis.

Roberts, M., & Seager, P. B. (1999). Predicting belief in paranormal phenomena: A comparison of conditional and probabilistic reasoning. *Applied Cognitive Psychology*, **13**, 443–450.

Roe, C. A. (1999). Critical thinking and belief in the paranormal: A re-evaluation. *British Journal of Psychology*, **90**, 85–98.

Rosenthal, R. (1967). Covert communication in the psychological experiment. *Psychological Bulletin*, **67**, 356–367.

Rosenthal, R. (1976). *Experimenter effects in behavioral research* (Enlarged ed.). New York: Irvington Press.

Rosenthal, R (1990). Replication in behavioural research. *Journal of Social Behaviour and Personality*, **5**, 1–30.

Rosenthal, R., & Rubin, D .B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, **3**, 377–415.

Royalty, J. (1995). The generalizability of critical thinking: Paranormal beliefs versus statistical reasoning. *Journal of Genetic Psychology*, **156**, 477–488.

Schlitz, M. J., & LaBerge, S. (1994). Autonomic detection of remote observation: Two conceptual replications. *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention,* 352–360.

Schmeidler, G. R. (1997). Psi-conducive experimenters and psi-permissive ones. *European Journal of Parapsychology*, **13**, 83–94.

Schneider, R., Binder, M., & Walach, H. (2000). Examining the role of neutral versus personal experimenter–participant interactions: An EDA-DMILS experiment. *Journal of Parapsychology*, **64**, 181–194.

Smith, M. D., Foster, C. L., & Stovin, G. (1998). Intelligence and paranormal belief: Examining the role of context. *Journal of Parapsychology*, **62**, 65–78.

Smith, M. D., & Gordon, M. S. (2002). The psychology of the "psi-conducive" experimenter. *Proceedings of Presented Papers: The Parapsychological Association 45th Annual Convention*, 247–255.

Solfvin, G. F., Kelly, E. F., & Burdick, D. S. (1978). Some new methods of analysis for preferential-ranking data. *Journal of the American Society for Psychical Research*, **72**, 93–110.

Tobacyk, J. J. (1988). *A Revised Paranormal Belief Scale*. Unpublished manuscript, Louisiana Tech University, Rushton, LA.

Tobacyk, J., & Thomas, A. (1997). How the big orthogonal seven is really the oblique seven. *Journal of Parapsychology*, **61**, 337–342.

Watt, C. A. (1996). Knowing the unknown: Participants' insight in three forced-choice ESP studies. *Journal of the American Society for Psychical Research*, **90**, 97–114.

Watt, C. A., & Morris, R. L. (1995). The relationships among performance on a prototype indicator of perceptual defence/vigilance, personality, and extrasensory perception. *Personality and Individual Differences*, **19**, 635–648.

Watt, C. A., & Ramakers, P. (in press). Experimenter effects with a remote facilitation of attention focusing task: A study with multiple believer and disbeliever experimenters. *Journal of Parapsychology.*

Watt, C. A., Ravenscroft, J., & McDermott, Z. (1999). Exploring the limits of direct mental influence: Two studies comparing "blocking" and "co-operating" strategies. *Journal of Scientific Exploration*, **13**, 515–535.

Wiezerbiki, M. (1985). Reasoning errors and belief in the paranormal. *Journal of Social Psychology*, **125**, 489–494.

Wiseman, R. (1997). *Deception and self-deception: Investigating psychics.* Amherst, MA: Prometheus.

Wiseman, R., & Schlitz, M. (1997). Experimenter effects and the remote detection of staring. *Journal of Parapsychology*, **61**, 197–208.

Wiseman, R., & Schlitz, M. (1999). Experimenter effects and the remote detection of staring: An attempted replication. *Proceedings of Presented Papers: The Parapsychological Association 42nd Annual Convention*, 471–479.

Wiseman, R., & Smith, M. (1994). A further look at the detection of unseen gaze. *Proceedings of Presented Paper: The Parapsychological Association 37th Annual Convention*, 465–478.

Wiseman, R., Smith, M., Freedman, D., Wasserman, T., & Hurst, C. (1995). Two further experiments concerning the remote detection of an unseen gaze. *Proceedings of Presented Papers: The Parapsychological Association 38th Annual Convention*, 480–490.

Department of Psychology
University of Edinburgh
7 George Square
Edinburgh EH8 9JZ
United Kingdom
cwatt@afb1.ssc.ed.ac.uk

Psychology Department
University of Hertfordshire
College Lane
Hatfield, Hertfordshire AL10 9AB
United Kingdom
R.Wiseman@herts.ac.uk