Experimenter as Participant:

What Can We learn from the Experimenter Effect?

by CHRIS ROE, University of Northampton

In the last issue of *Mindfield* I noted that even where an effect is real and robust we should still expect variations in the effect sizes captured in any particular study as a consequence of sampling error, and that the likelihood of replicating an earlier finding is dependent not on the capricious nature of psi (and whether we have performed the correct propitiatory rites to the Fates), but more soberly on the prevailing effect size and the power of the current study. I illustrated that the pattern of outcomes we see from ganzfeld replications are a reasonable fit with this form of stochastic replication (and the recent meta-analysis of "feeling the future" studies also gives outcomes that are a remarkable approximation of statistical expectation — see the funnel plot on p. 10 of Bem, Tressoldi, Rabeyron, & Duggan, 2016). However, it is also true that some of the variance in outcomes can be attributed to other sources; in particular, we frequently see that some researchers or laboratories seem to consistently get positive results while others consistently score at or below chance.

This differential performance is known as the experimenter effect, and has been recognised as an important characteristic of parapsychological findings. Smith (2003a, p. 70), for example, asks "why do some investigators seem to be consistently unable to obtain evidence for psi, whilst others continue to obtain psi effects in their experiments? What is it about these apparently 'successful' experimenters that distinguish them from 'unsuccessful' experimenters? … [T]hese questions lie at the heart of the issue of replication in parapsychology". Palmer and Millar (2015) similarly assert "The identity of the principal investigator (PI) … is the best predictor we have of the outcome of a psi experiment" (p. 293). And Palmer (1986, pp. 220–221) has earlier claimed, "the experimenter effect is the most important challenge facing modern parapsychology. It may be that we will not be able to make too much progress in other areas of the field until the puzzle of the experimenter effect is solved". In this article I would like to look at the experimenter effect

more closely to see what we can learn from its occurrence.

Examples from Parapsychology

Perhaps the most striking early demonstration of an experimenter effect in parapsychology was reported by West and Fisk (1953). The collaborators had had contrasting fortunes in previous research, with Fisk having reported a number of successful studies while West had only produced chance outcomes. They collaborated on a forced choice ESP experiment in which 20 participants were asked to predict the times shown on a series of cards that bore a simple clock face showing a hand pointing to one of the numbers 1-12. Half of the packs were prepared by West and half by Fisk. Participants performed in line with their experimenter's previous track record, with West's scoring at chance and Fisk's significantly above chance. Importantly for our later consideration of explanations for the experimenter effect, the researchers had very little contact with participants in this postal study. Intriguingly, when Fisk did all the mailing in a second study but produced only half the target sets while the others were prepared by West using exactly the same random number table, participants still showed differential performance, scoring above chance for targets that Fisk had prepared but with null scores for those produced by West.

Perhaps the most striking early demonstration of an experimenter effect in parapsychology was reported by West and Fisk (1953).

Being psi conducive is not necessarily all-or-nothing. Sometimes researchers can find that certain approaches or populations consistently give rise to null results despite their ability to produce above chance scoring in other circumstances. Gaither Pratt, for example, had reported high ESP scores in studies that involved participants who were of a similar age or disposition to himself (e.g., Pratt, 1973), but when he collaborated with Margaret Price on a study involving children in which Pratt tested the girls and Price the boys, the former scored close to chance (deviation +81 after nearly 54,000 trials) whereas the latter scored above chance (+609 in just over 29,000 trials). In a second series, both tested equal numbers of girls and boys, with each acting as the other's research assistant, so that they could monitor the procedure and ensure that it stayed uniform. Now the difference between girls' and boys' scores was negligible, but Price's participants again had significantly high ESP scores, and Pratt's did not. Schmeidler (1997, p. 85) attributes this to differences in their nature: Pratt was "a quiet, careful, methodical young man," while Price was "a charming, friendly, outgoing young woman" who also was experienced at running experiments with schoolchildren; indeed, their paper describes Price's attempts to encourage Pratt to engage the participants in conversation and have them "open up," and notes the positive effects this had on their scoring.

The most salient recent test of experimenter effects in parapsychology is a collaboration between proponent Marilyn Schlitz, who had a track record of producing psi effects (e.g., Braud & Schlitz, 1991), and Richard Wiseman, a skeptic who had a track record of failing to produce them (e.g. Wiseman & Greening, 2002). They had both previously conducted experiments looking at staring detection, with Schlitz again reporting evidence of psi (Schlitz & LaBerge, 1994) and Wiseman again reporting no evidence of it (Wiseman & Smith, 1994). Interestingly they agreed to conduct a joint study in which they would both use the same experimental set-up and draw participants from the same pool to see whether they could replicate their differential performance. The basic design of their staring detection studies involved either Wiseman or Schlitz acting as the influencer

The first experiment (Wiseman & Schlitz, 1998) was conducted at Wiseman's Hertfordshire laboratory and involved 16 participants each. Conditions were identical except for the identity of the experimenter/starer.

("starer"), observing the participant ("staree") intermittently via cctv. During staring trials they would attend to the starees with the intention of physiologically arousing them, while during no stare trials they would look elsewhere and focus their attention on something else. The sequence of stare and no stare trials was randomly determined by the experimental program. Meanwhile the starees simply relaxed as much as possible while their electrodermal activity (EDA) was continuously monitored. A remote staring effect would be evidenced by a significant difference in EDA between the stare and no stare periods.

The first experiment (Wiseman & Schlitz, 1998) was conducted at Wiseman's Hertfordshire laboratory and involved 16 participants each. Conditions were identical except for the identity of the experimenter/starer. The EDA of Schlitz's participants was significantly higher in stare than in no-stare trials, as predicted, whereas the EDA of Wiseman's participants showed no effect, thus confirming their earlier, independent, findings. A second joint project (Wiseman & Schlitz, 1999) took place at Schlitz's laboratory at the Institute of Noetic Sciences (IONS) and again both experimenters employed the same procedures, equipment, and participant pool ($N = 35$ for each experimenter). Schlitz's participants again showed a small but statistically significant effect, whereas Wiseman's did not; however, in this case Schlitz's participants were significantly *less* activated during the stare than nonstare periods, contrary to study one. In a third study, the roles of experimenter and starer were separated in a 2x2 design so that some participants were met and briefed by the person who would act as their starer but others were briefed by one researcher while the other acted as their starer. At the time of the briefing session the researchers did not know which would be the starer. The study was again run at IONS and the 100 participants were staff members or local volunteers. The mean effect was somewhat larger when Schlitz was greeter rather than Wiseman, and when Wiseman was starer rather than Schlitz, but none of the effects deviated meaningfully from chance. This is a shame given that the study promised to tease apart two important roles, but with no evidence of psi there was no possibility to explore experimenter effects. The authors noted that they both experienced a diminished motivation and interest in the experiment, but this could not be evidenced (for example as a decline effect), and the disappearance of the effect could equally be attributed to methodological refinements.

Space does not permit me to give details of other examples of experimenter effects in parapsychology, but to illustrate the ubiquitous nature of this phenomenon, I should like to note that Bem (2012) has accounted for the Ritchie, Wiseman, and French failures to replicate his anomalous anticipation effects by commenting, "I believe that some major variables determining the success or failure of replications are likely to be the experimenters' expectations about, and attitudes toward, the experimental hypothesis" (p. 353), and Millar (in Palmer & Millar, 2015) attributes the failure of the large European replication of the PEAR laboratory's micro psychokinesis work to experimenter effects.

Accounting for the Experimenter Effect

So if the case is made that the experimenter can affect the outcome of a parapsychology experiment, how is that brought about? Smith (2003a) offers four categories of explanation: experimenter fraud,

experimenter error, experimenter-participant interaction, and experimenter psi. I have considered experimenter fraud in a previous *Mindfield* article and do not regard it as a sufficiently pervasive factor to account for the effect. Although unwitting recording errors are more likely to be consistent with expectation, the variety of circumstances under which experimenter effects have been identified (e.g., in data that are automatically recorded, such as in staring detection) suggests that it is not a primary explanation. (Schmeidler [1997, p. 83] considered both of these explanations "frivolous.")

Instead I will focus on the two explanations that Palmer and Millar (2015) identify: the experimenter psi hypothesis and the experimenter behaviour hypothesis. The first of these proposes that the outcome from a parapsychology experiment reflects action of the experimenter's own psi; the second that it reflects the experimenter's ability to set a positive expectation for the trial and to put participants sufficiently at ease to express their own psi abilities. Schmeidler (1997) refers to such experimenters as psi *conducive* and psi *permissive* respectively, a distinction that I think is very helpful.

I will consider the psi conducive experimenter first, since explanations in terms of experimenter psi are given more emphasis by Palmer and Millar (2015) in their recent overview — the section in their chapter devoted to behavioural causes amounts to 1 1/2 pages while experimenter psi is given three times that much space. The grounds for proposing the experimenter psi hypothesis seem primarily to derive from a perception that psi may be boundless and goal-oriented, and in such circumstances we should acknowledge that the person whose needs are most likely to be satisfied by the statistical outcome of an experiment will probably be the experimenter rather than the participant (particularly where the latter has little connection with the study and its outcomes once debriefing has been completed). Indeed, some experimental data — such as from the Fisk and West study described earlier —are admittedly difficult to explain except in terms of experimenter psi, given that there is virtually no contact between experimenter and participant to enable the right ambiance to be set for participants to exhibit their own psi.

In support of this interpretation, Smith (2003b) surveyed active experimental parapsychologists and among 40 responses found that psi conduciveness correlated with beliefs about one's own ESP and PK abilities (although of course this could be circular, with experimental success boosting one's conviction of personal ability). Millar (in Palmer & Millar, 2015) characterises psi conducive experimenters as virtuosi and estimates that although there might be as few as half a dozen such

The grounds for proposing the experimenter psi hypothesis seem primarily to derive from a perception that psi may be boundless and goal-oriented.

persons who are research active in parapsychology at any one time, they "dominate parapsychology," which seems plausible to me — there would probably be good agreement in the community as to the names of the likely candidates among the current crop of researchers. But while this might account for impressive initial programmes of study from particular laboratories (e.g., Bem, 2011; Bem & Honorton, 1997; Radin, 1997), it struggles to account for subsequent successful replication attempts involving a wide range of other laboratories (e.g., Bem et al., 2016; Storm, Tressoldi, & Di Risio, 2010; Mossbridge, Tressoldi, Utts, Ives, Radin, & Jonas, 2014). Most intriguingly, it is difficult to see how experimenter effects could be responsible for evidence of psi that derives from studies that were conducted for non-parapsychological purposes by researchers with no interest in capturing psi (see Bierman, 2000).

Of course, it could be argued that since we do not know the limits of psi then it is theoretically possible for the original experimenters — unconsciously, in a goal-directed, need-based fashion — to affect the outcome of ostensibly independent replications. And here's the rub. Allowing such an unbounded mechanism for the psi conducive experimenter effect amounts, I think, to an unfalsifiable hypothesis. Although this does not in itself negate the hypothesis, it would severely constrain our capacity to make systematic progress in testing it or studying the phenomena that are affected by it. It seems to me sensible to see how far we can go in explaining experimenter effects in terms of social and cognitive factors before we start to invoke such inscrutable and potentially omnipotent mechanisms.

Conventional Experimenter Effects

We need to consider, then, what kinds of experimenter effect are possible by more conventional psychosocial means so as to assess whether any or all of the effects we find in parapsychology might be explained in these ways. The classic work on experimenter effects by Robert Rosenthal and colleagues (see Rosenthal & Rosnow, 2009) illustrates that powerful changes in study outcome can be achieved by manipulating experimenter characteristics such as expectancy and interactive style. Rosenthal and Jacobson's (1968) study of the Pygmalion effect, for example, illustrates how expectancy can become self-fulfilling. All the children at an elementary school completed a nonverbal test of intelligence. Teachers were told that the assessment could detect those children who were about to undergo a cognitive blooming, showing rapid improvement, with approximately 20% of the children being identified as "bloomers." In fact, these were randomly chosen from each of the 18 classes across 6 grade levels. At the end of the school year, some 8 months later, all the children were administered an IQ test and it was found that those labelled as bloomers had improved 4 IQ points relative to their classmates (2 points on verbal, and a whopping 7 points on reasoning). Pygmalion effects have subsequently been demonstrated not just in classrooms but in courtrooms, nursing homes, management settings, and even swimming pools (Rosenthal, 1994, p. 178).

Effects are not restricted to human participants, who might be sensitised to verbal and nonverbal signals from a person with authority over them. Animals can also be subject to experimenter effects, as illustrated in studies purporting to investigate maze learning in rats (Rosenthal & Rosnow, 2009, pp. 423-439). Participants were told that interbreeding of rats that had done well when learning a maze would produce successive generations of rats that would do considerably better than unselected rats; similarly, rats bred from poor performers would produce offspring that were worse than average. Each of 12 students was assigned 5 rats, all of which were either from the Maze-Bright or Maze-Dull strain. Actually, the rats were randomly grouped, with the proviso that groups were matched for mean age, since maturation could affect performance. Each rat was run in the maze 10 times a day for 5 days, with the experimenter recording whether each trial was a success or failure. Performance of the Maze-Bright rats was significantly better than for Maze-Dull on days 1, 4, and 5, as well as overall. Maze-Bright rats conformed more closely to the expected 'learning curve' across the 5 days, while Maze-Dull rats showed some early improvement that subsided. This suggests that the differential effect was quite immediate rather than having to be built up more subtly through regular contact. On 20% of all trials the rats failed by making no movement at all, with 17 of these involving Bright rats and 43 Dull. If these trials are excluded as a possible confound, Bright rats were still significantly quicker than Dull ones on successful trials. Although experi-

Since these quite striking early results, interpersonal expectancy effects have been confirmed in a wide range of studies.

menter/participant interactions with their animals were not formally monitored, the experimenter group included two confederates of the researchers who were able to give feedback on research practices. They did not observe any incorrect entries or data fudging, but did see some instances of experimenters prompting the rats to move by prodding them, though this occurred slightly more often with Dull rats than Bright ones (3 versus 2 occasions) so that was not a contributor to group differences. More importantly, experimenters working with Bright rats handled them more often and more gently, and this could have been a mechanism that led to behavioural differences when maze running.

Since these quite striking early results, interpersonal expectancy effects have been confirmed in a wide range of studies. Rosenthal (1994, p. 176) refers to 464 studies with an overall *d* of .63 (*r* = .30) that demonstrates the effect in a variety of contexts, including studies of reaction time, interpretation of inkblots, animal learning, person perception, and skill learning. Harris and Rosenthal (1985) provide a meta-analysis of 135 studies that focus on 16 behaviours hypothesised to mediate the effect, including warm interpersonal climate, experimenter expectancy, focused attention, and feedback (Harris & Rosenthal, 1985; Rosenthal, 1994).

Psi Permissive Experimenter Qualities

It is interesting to note that similar factors have been highlighted by parapsychologists interested in experimenter effects. Schmeidler and Maher (1981) video recorded researchers as they gave talks and answered questions at an academic conference. Independent judges viewed footage of five "psi-conducive" and five "psi-inhibitory" researchers matched for relative age and physical features, and rated them along a number of dimensions. Psi-conducive experimenters were rated as more flexible, enthusiastic, friendly, likeable, and warm and less tense, irritable, and cold. Of course, it is possible that the way that researchers present themselves at a parapsychological conference is affected by the success or otherwise of the research they are describing, and their sense of how they might be perceived by their peers, but it is encouraging that dimensions that seem likely to put participants at ease and enable them to behave openly and naturalistically are the ones that predict success at a psi task.

Some researchers have attempted to manipulate these factors. For example, Honorton, Ramsay, and Cabibbo (1975) had two experimenters who interacted either in a positive manner (friendly, casual, supportive) during the time taken to establish rapport with the participant, or in a negative manner (abrupt, formal, and unfriendly) while they went quickly into the task. The positive treatment gave significantly higher scores than the negative one. However, when Schneider, Binder, and Walach (2000) manipulated the experimenter's interactional style analogous to warmth (personal versus neutral), they found no difference between the conditions. Similarly, Parker (1975) manipulated expectancy among six experimenters (student data collectors) so that they were presented as strong believers or strong disbelievers. Although overall scoring was null, there was a significant difference in performance between the experimenter groups, even though they, in fact, were testing the same sender-receiver pairs. In practice it is very difficult to manipulate interpersonal style without coming off as inauthentic. A preferred method is to appoint a number of experimenters who might vary

naturally on interpersonal dimensions. Watt and Ramakers (2003) found that participants working with genuine believer-experimenters on a remote facilitation of attention task scored significantly better than those working with disbeliever-experimenters (who performed at chance levels).

A further difficulty with such studies is in ensuring that the intended manipulation (e.g., of participant belief or warmth) has the expected impact on the participant. Schmeidler (1997) has noted that in practice this can vary from case to case depending on the preferences of the participant — what puts one person at ease may set another person on edge. For example, intimacy and closeness may establish rapport with some participants but be seen as intrusive and cloying by others. This can be addressed by having experimenters interact as they ordinarily would and then asking participants to retrospectively rate the interaction. I have been involved in studies of ganzfeld ESP and of PK that adopted this method, with participants completing questionnaires after the interaction but before the psi task (with the assurance that their responses will be scored by an independent judge and the experimenter would never have access to them). In one study (Sherwood, Roe, Holt, & Wilson, 2005) we found that ganzfeld success was associated with more relaxed, more optimistic, and more confident senders and experimenters and more confident receivers. When looking at a PK task (Roe, Davey, & Stevens, 2006) we found performance was associated with positive mood, positivity toward the task, and relaxation, but especially experimenter confidence of success. This strongly suggests to me that we can be sensitive to aspects of the interaction that have a direct bearing on the participants' ability to demonstrate psi.

In this context, it is worth reflecting on Caroline Watt's valuable insight into the intentions and interactive style of psi permissive and psi inhibitory experimenters Schlitz and Wiseman (Watt, Wiseman, & Schlitz, 1998). Watt conducted interviews after two of the studies described earlier had been completed, to see if they could identify differences in their research practice that might have contributed to the differences in outcome. From these it is clear that Schlitz had a preparatory ritual that focused her on the trial and the participant, and set a clear intention for the session. On greeting the participant she sought to personalise the interaction and develop rapport, as well as normalise the phenomenon and set expectation of success based on previous experiences and research findings. She reflects (p. 23), "I would try to give [participants] every reason to feel optimistic that this particular thing we're doing together could actually produce something."

It is worth reflecting on Caroline Watt's valuable insight into the intentions and interactive style of psi permissive and psi inhibitory experimenters Schlitz and Wiseman.

During debriefing she would provide closure regarding the experience by discussing and interpreting the data. In contrast, Wiseman described how his preparation was restricted to practical matters such as ensuring equipment and materials were organised. His briefing of participants was more "matter of fact," and although he did not declare his scepticism it seems likely it would have been recognised, "sometimes I'd say, 'I know this sounds like quite a weird experiment, but let's try it'. If the participant was sceptical I'd say, 'Well, let's give it a go.'" Indeed, Wiseman suspected that his scepticism would be apparent to participants, "there are probably big differences in that initial chat with them, just in terms of how much they walked away thinking that yes indeed this was a procedure that was going to work" (p 23). (Smith [2003a, p. 75] notes that Schlitz's participants reported stronger belief in psi and speculates that the

Wiseman-Schlitz effect may have been due to the communication of experimenters' expectations to the participants.) Wiseman would give a clear explanation of what the study entailed but did not attempt to develop rapport, "for the most part, no, there's not a great deal of other chat. I would describe it as more businesslike, but not unfriendly" (p. 21) and this did not extend to acknowledging any spontaneous experiences as a means of setting positive intention, "if some of them started to talk through their experiences with me, I found that quite difficult to relate to because I'm quite sceptical about these things. You don't want to be confrontational with people and so you end up nodding and going 'oh, that's interesting. Anyway, back to the experiment" (p 21). From these descriptions it seems obvious who we would expect to be psi permissive and who psi inhibitory.

Conclusion

In this article I have tried to show that replication in parapsychology might be less exact than in the natural sciences because of the wide range of interpersonal factors that can affect performance. These have been bundled together under the heading of *experimenter effects* to show how subtly the demeanour and expectations of the experimenter can shape the outcome of the study. Understanding these relations might make more predictable the occurrence of the phenomena we wish to study. It hinges, in my view, on acknowledging that the social sciences are not like the natural sciences in that the researcher cannot remain separate from the system that he or she is studying, such that the practice of research is as much an art that involves refining one's awareness of how to enable participants to fulfill their potential as it is a mastery of particular techniques or methods.

I have tried to show that replication in parapsychology might be less exact than in the natural sciences because of the wide range of interpersonal factors that can affect performance.

Of course, many in the social sciences are willing to reject this rather messy (and potentially idiosyncratic) understanding of the experimenter-participant dyad, preferring instead a simpler model in which participants think and do only what is required of them by the experimental manipulation. Despite initial keen interest in Rosenthal's work and the accumulation of a persuasive evidence base for the occurrence of experimenter effects, the topic is barely covered in mainstream psychology texts today. As Tart (2016, p. 54) observes, "the topic pretty much disappeared from the psychology literature. Not that the issue had been dealt with and the problem could now be dismissed as solved, people just stopped writing about it — my guess was that such a threat to the necessary objectivity of psychological studies was just too threatening and nobody wanted to think about it. ... it had only been a hundred years since they let us out of the philosophy department and we were afraid our claims to be scientists would prove unfounded and they would send us back to the philosophy department — understandable human behaviour, poor science." Quite.

References

Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*, 407-425.

Bem, D. J. (2012). Longstanding misunderstandings about replication. *The Psychologist*, *25*(5), 352-353.

Bem, D. J. & Honorton, C. (1994). Does psi exist? *Psychological Bulletin*, *115*, 4-18.

Bem, D., Tressoldi, P., Rabeyron, T., & Duggan, M. (2016). Feeling the future: A meta-analysis of 90 experiments on the anomalous anticipation of random future events. *F1000Research 2016*, *4*, 1188 (doi:10.12688/f1000research.7177.2)

Bierman, D. J. (2000). Anomalous baseline effects in mainstream emotion

research using psychophysiological variables *Proceedings of presented papers: The Parapsychological Association 43rd annual convention*, 34-47.

Braud, W. G., & Schlitz, M. J. (1991) Consciousness interactions with remote biological systems: Anomalous intentionality effects. *Subtle Energies*, *2*, 1-46.

Harris, M. J., & Rosenthal, R. (1985). Mediation of interpersonal expectancy effects: 31 meta-analyses. *Psychological Bulletin*, *97*(3), 363-386.

Honorton, C., Ramsey, M., & Cabibbo, C. (1975). Experimenter effects in extrasensory perception. *Journal of the American Society for Psychical Research*, *69*, 135-149.

Mossbridge, J. A., Tressoldi, P., Utts, J., Ives, J. A., Radin, D., & Jonas, W.B. (2014). Predicting the unpredictable: Critical analysis and practical implications of predictive anticipatory activity. *Frontiers in Neuroscience*, *8*(146), 1-10.

Palmer, J. (1986). ESP research findings: The process approach. In H.L. Edge, R.L. Morris, J. Palmer, & J.H. Rush (Eds.), *Foundations of parapsychology: Exploring the boundaries of human capability*. London: Routledge & Kegan Paul.

Palmer, J., & Millar, B. (2015). Experimenter effects in parapsychological research. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.). *Parapsychology: A handbook for the 21st century*. Jefferson, NC: McFarland & Co. (293-300).

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

Pratt, J. G. (1973). A decade of research with a selected subject: An overview and reappraisal of the work with Pavel Stepanek. *Proceedings of the American Society for Psychical Research*, *30*, 1-78.

Radin, D. I. (1997). Unconscious perception of future emotions: An experiment in presentiment. *Journal of Scientific Exploration*, *11*, 163.180.

Roe, C. A., Davey, R., & Stevens, P. (2006). Experimenter effects in laboratory tests of ESP and PK using a common protocol. *Journal of Scientific Exploration*, *20*, 239-253.

Rosenthal, R. (1994). Interpersonal expectancy effects: A 30-year perspective. *Current Directions in Psychological Science*, *3*(6), 176-179.

Rosenthal, R., & Rosnow, R.L. (2009). *Artifacts in behavioral research*. Oxford: Oxford University Press.

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.

Schlitz, M., Wiseman, R., Watt, C., & Radin, D. (2006). Of two minds: Skeptic-proponent collaboration within parapsychology. *British Journal of Psychology*, *97*, 313-322.

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.

Sherwood, S. J., Roe, C. A., Holt, N., & Wilson, S. (2005). Interpersonal psi: Exploring the role of the experimenter and the experimental climate in a ganzfeld telepathy task. *European Journal of Parapsychology*, *20*, 150-172.

Smith, M. (2003a). The role of the experimenter in parapsychological research. In J. Alcock, J. Burns, & A. Freeman (Eds.), *Psi wars: Getting to grips with the paranormal*. Charlottesville, VA: Imprint Academic. (69-84)

Smith, M. (2003b). The psychology of the psi-conducive experimenter: Personality, attitudes toward psi and personal psi experience. *Journal of Parapsychology*, *67*, 117-128.

Storm, L., Tressoldi, P. E., & Di Risio, L. (2010). Meta-analysis of free-response studies, 1992–2008: Assessing the noise reduction model in parapsychology. *Psychological Bulletin*, *136*(4), 471-485.

Tart, C.T. (2016). On attempting to fool subjects [letter]. *Journal of the Society for Psychical Research*, *80*, 52-55.

Watt, C., & Ramakers, P. (2003). Experimenter effects with a remote facilitation of attention focusing task. *Journal of Parapsychology*, *67*, 99-116.

Watt, C., Wiseman, R., & Schlitz, M. (1998). Tacit information in remote staring research: The Wiseman-Schlitz interviews. *The Paranormal Review*, *24*, 18-25.

West, D. J., & Fisk, G. W. (1953). A dual ESP experiment with clock cards. *Journal of the Society for Psychical Research*, *57*, 185-189.

Wiseman, R. & Greening, E. (2002). The mind machine: A mass participation experiment into the possible existence of extrasensory perception. *British Journal of Psychology*, *93*, 487-499.

Wiseman, R. & Schlitz, M. (1998). Experimenter effects and the remote detection of staring. *Journal of Parapsychology*, *61*(3), 197-208.

Wiseman, R., & Schlitz, M. (1999). Replication of experimenter effect and the remote detection of staring. *Proceedings of presented papers: The Parapsychological Association 42nd annual convention*, 471–479.

Wiseman, R., & Smith, M.D. (1994). A further look at the detection of unseen gaze. *Proceedings of presented papers: The Parapsychological Association 37th annual convention*, 465–478.