

# Der Versuchsleiter als Versuchsteilnehmer Was können wir vom Versuchsleiter-Effekt lernen?¹

CHRIS A. ROE²

**Zusammenfassung** – Im Bereich der experimentellen parapsychologischen Forschung sind manche Forscher besonders erfolgreich, während es anderen kaum gelingt, Psi in ihren Experimenten nachzuweisen. Dieses Phänomen ist als Versuchsleitereffekt bekannt und gilt als wichtiges Charakteristikum parapsychologischer Befunde. Mit diesem Artikel wird der Frage nachgegangen, was erfolgreiche und weniger erfolgreiche Versuchsleiter unterscheidet und womit diese Unterschiede erklärt werden können. Wissenschaftliche Befunde zur Rolle der Erwartungshaltung und zum Interaktionsstil mit Versuchsteilnehmern werden vorgestellt. Diese konventionellen Versuchsleitereffekte werden als Alternative oder Ergänzung zur Versuchsleiter-Psi-Hypothese diskutiert, nach der vorhandene Psi-Fähigkeiten des Versuchsleiters selbst für das Auftreten von Psi-Effekten in den Experimenten verantwortlich gemacht werden.

**Schlüsselbegriffe:** Versuchsleitereffekt – Replikationsproblem – Interaktionsstil – Erwartungshaltung des Versuchsleiters

## Experimenter as Participant What can we learn from the experimenter effect?

**Abstract** – In the field of experimental parapsychological research, some researchers are particularly successful, while others hardly succeed in proving Psi in their experiments. This dif-

- 
- 1 Dies ist eine Übersetzung des Artikels “Experimenter as Participant”, der im englischen Original in der Ausgabe 3/2016 (S. 89–97) der Zeitschrift *Mindfield* erschienen ist. Die Redaktion bedankt sich bei den Herausgebern von *Mindfield* und bei Chris Roe für die Druckgenehmigung der deutschen Übersetzung.
  - 2 Chris A. Roe ist Professor für Psychologie an der University of Northampton und Direktor des dort ansässigen Centre for Psychology & Social Sciences. Roe ist Perrott-Warrick Senior Researcher (Trinity College, Cambridge) und Gastprofessor an der Bucks New University. Er ist Präsident der *Society for Psychical Research*, ehemaliger Vorsitzender der *British Psychological Society Transpersonal Psychology Section*, ehemaliger Präsident der *Parapsychological Association* und internationaler Partner der *Parapsychology Foundation* für England. Seine Forschungsinteressen liegen im Verständnis der Natur anomaler Erfahrungen und umfassen Forschungen zur Phänomenologie paranormaler Erfahrungen, insbesondere im Hinblick auf das Wohlbefinden, die Psychologie des paranormalen Glaubens und der Täuschung, sowie experimentelle Ansätze zur Untersuchung außersinnlicher Wahrnehmung und Psychokinese, insbesondere im Zusammenhang mit relevanten psychologischen Faktoren.

ferential performance is known as the experimenter effect, and has been recognized as an important characteristic of parapsychological findings. This article explores the question of what distinguishes successful and less successful investigators and how these differences can be explained. Scientific findings on the role of expectations and the style of interaction with participants are presented. These conventional experimenter effects are discussed as an alternative or supplement to the experimenter psi hypothesis, according to which existing psi abilities of the experimenter are held responsible for the occurrence of psi effects in the experiments.

*Keywords:* experimenter effect – replication problem – interaction style – experimenter's expectation

In der vergangenen Ausgabe von *Mindfield*<sup>3</sup> habe ich angemerkt, dass wir selbst in den Fällen, in denen ein Effekt real und stabil ist, Schwankungen in den Effektstärken zu erwarten haben, die in jeder Einzelstudie als Folge des Stichprobenfehlers auftreten; und außerdem, dass die Wahrscheinlichkeit, einen früheren Befund zu replizieren, nicht von der Unberechenbarkeit von Psi-Effekten abhängt (und auch nicht davon, ob wir die korrekten, gnädig stimmenden Rituale für die Schicksalsgöttinnen durchgeführt haben), sondern, viel schlichter, von der vorliegenden Effektstärke und der statistischen Power der jeweiligen Studie.<sup>4</sup> Ich habe gezeigt, dass die Ergebnismuster aus Ganzfeld-Replikationen gut mit dieser Form der stochastischen Replikation zusammenpassen (und auch die kürzlich erschienene Meta-Analyse der „feeling the future“-Studien zeigt Ergebnisse, die eine bemerkenswerte Annäherung an den statistischen Erwartungswert darstellen – siehe den *funnel plot* auf S. 10 von Bem, Tressoldi, Rabeyron & Duggan, 2016). Trotzdem können die Schwankungen bei den Ergebnissen in einigen Fällen auch mit anderen Faktoren erklärt werden; es lässt sich erkennen, dass einige Forscher oder Forschungslabors häufig positive Ergebnisse erzielen, während andere durchweg Ergebnisse vorweisen, die der Zufallserwartung entsprechen oder darunter liegen.

Dieses Phänomen ist als Versuchsleitereffekt bekannt und gilt als wichtiges Charakteristikum parapsychologischer Befunde. Smith (2003a: 70) beispielsweise stellt die Frage: „Warum sind einige Forscher außerstande, Beweise für Psi zu erbringen, während andere in ihren Experimenten immer wieder Psi-Effekte erzielen? Was ist es, das diese offenbar ‚erfolgreichen‘ Experimentatoren von den ‚erfolglosen‘ Experimentatoren unterscheidet? ... [D]iese Fragen stellen den Kernpunkt des Replikationsproblems in der Parapsychologie dar.“ Auch Palmer und Miller (2015) stellen fest: „Die Persönlichkeit des Hauptuntersuchungsleiters (...) ist der beste Prädiktor, den wir

---

3 Roe, C. A. (2016). „Is inconsistency our only consistent outcome?“ *Mindfield*, 8(2): 68–73. – Anm. d. Red.

4 Mit statistischer Power wird die Wahrscheinlichkeit bezeichnet, dass ein Effekt entdeckt wird, wenn er auch tatsächlich existiert. Die statistische Power einer Untersuchung hängt vor allem von der Stichprobengröße, der Effektstärke und dem festgelegten Signifikanzniveau ( $\alpha$ -Fehler) ab (Anm. d. Übers.).

bezüglich des Ergebnisses eines Psi-Experimentes haben“ (S. 293). Palmer (1986: 220–221) hatte schon zuvor behauptet, dass „der Versuchsleitereffekt die wichtigste Herausforderung ist, mit der die modernen Parapsychologie konfrontiert ist. Es könnte durchaus sein, dass wir in anderen Teilgebieten des Feldes keine besonderen Fortschritte erzielen, solange das Rätsel des Versuchsleitereffektes nicht gelöst wurde.“ In diesem Artikel werde ich einen genaueren Blick auf den Versuchsleitereffekt werfen, um herauszufinden, was wir von seinem Auftreten lernen können.

### ***Beispiele aus der Parapsychologie***

Die wohl eindrucksvollste Veranschaulichung des Versuchsleitereffektes in der Frühzeit der Parapsychologie wurde von West und Fisk (1953) berichtet. Die beiden hatten in vorhergehenden Untersuchungen ein recht unterschiedliches Geschick: Während Fisk von einer ganzen Reihe von erfolgreichen Studien berichtete, hatte West bloß Ergebnisse erzielt, die der Zufallserwartung entsprechen. Sie arbeiteten gemeinsam an einem *Forced-Choice*-ASW-Experiment,<sup>5</sup> bei dem 20 Teilnehmer die Zeit vorhersagen sollten, die auf einer Reihe von Karten mit einem einfachen Ziffernblatt abgebildet war. Die Karten zeigten eine Hand, die auf eine der Zahlen zwischen 1 und 12 hindeutete. Die eine Hälfte der Kartensets wurde von West vorbereitet, die andere Hälfte von Fisk. Die Leistung der Teilnehmer entsprach der Erfolgsgeschichte ihres jeweiligen Versuchsleiters: West erzielte Ergebnisse, die der Zufallserwartung entsprachen, während Fisks Ergebnisse signifikant darüber lagen. Wesentlich für unsere nachfolgenden Überlegungen zur Erklärung des Versuchsleitereffektes ist die Tatsache, dass die Versuchsleiter in dieser rein auf Schriftverkehr beruhenden Studie [postal study]<sup>6</sup> nur sehr wenig Kontakt zu den Teilnehmern hatten. Erstaunlicherweise zeigten die Teilnehmer in einer zweiten Studie auch dann noch eine unterschiedliche Leistung, als Fisk den gesamten Postversand übernahm, aber lediglich die Hälfte der Kartensets vorbereitete, während die andere Hälfte von West unter Verwendung genau der gleichen Zufallszahlentabelle vorbereitet wurde: Mit den Sets, die Fisk vorbereitet hatte, erreichten die Teilnehmer Ergebnisse außerhalb der Zufallserwartung, während diejenigen, die mit den Sets von West arbeiteten, unauffällig blieben.

Ob eine Person psi-förderlich ist oder nicht, folgt nicht notwendigerweise dem Alles-oder-nichts-Prinzip. Manchmal können Forscher beobachten, dass bestimmte Herangehensweisen

---

5 Bei der *Forced-Choice*-Methode sind die möglichen Lösungen, unter denen man zu wählen hat, vorgegeben („erzwungene Wahl“). Im Unterschied dazu sind bei der *Free-Response*-Methode freie Antwortmöglichkeiten gegeben (Anm. d. Übers.).

6 Im Folgenden werden die Original-Formulierungen in eckigen Klammern angemerkt, wenn es zum Abgleich der Übersetzung sinnvoll erschien. Eckige Klammern innerhalb von direkten Zitaten wurden so aus dem Original übernommen und beziehen sich, wie üblich, auf eine Anpassung des ursprünglichen Textes (Anm. d. Übers.).

oder bestimmte Stichprobengruppen statistisch gesehen durchweg unauffällig bleiben, obwohl sie in anderen Umständen durchaus Ergebnisse erzielen könnten, die über der Zufallserwartung liegen. Gaither Pratt beispielsweise verzeichnete eine hohe Anzahl an ASW-Treffern in solchen Studien, in denen die Teilnehmer in einem ähnlichen Alter oder in einer ähnlichen Verfassung wie er selbst waren (z. B. Pratt, 1973). Als er jedoch mit Margaret Price für eine Studie zusammenarbeitete, in der die Teilnehmer Kinder waren, erzielte Pratt, der die Mädchen untersucht hatte, ein Ergebnis nahe der Zufallserwartung (Abweichung +81 nach fast 54.000 Durchgängen [trials]), während Price, die die Jungen untersucht hatte, einen Wert erzielte, der darüber lag (+609 bei nur etwa 29.000 Durchgängen). In einem zweiten Durchlauf dieses Experimentes untersuchten die beiden die gleiche Anzahl an Mädchen und Jungen und fungierten zudem als Assistent für den jeweils anderen, sodass sie die Prozedur beobachten und sicherstellen konnten, dass sie einheitlich verläuft. Jetzt war der Unterschied zwischen den Ergebnissen der Mädchen und denen der Jungen nur geringfügig. Die Teilnehmer allerdings, die Price untersucht hatte, hatten abermals signifikant nach oben abweichende ASW-Ergebnisse. Bei Pratts Teilnehmern war dies nicht der Fall. Schmeidler (1997: 85) erklärt dies mit ihrem unterschiedlichen Charakter: Pratt war „ein ruhiger, vorsichtiger, systematisch vorgehender junger Mann“, während Price „eine charmante, freundliche, kontaktfreudige junge Frau“ war, die zudem Erfahrungen damit hatte, Experimente mit Schulkindern durchzuführen; ihre Publikation erwähnt sogar die Versuche von Price, Pratt zu Gesprächen mit den Teilnehmern zu ermutigen und sie dazu zu bringen, sich „zu öffnen“, und vermerkt die positiven Effekte, die dies auf ihre Ergebnisse hatte.

Die bemerkenswerteste, aktuelle Überprüfung von Versuchsleitereffekten bestand in einer Zusammenarbeit zwischen der von Psi überzeugten Marilyn Schlitz, die gute Erfolge damit hatte, Psi-Effekte hervorzurufen (z. B. Braud & Schlitz, 1991), und dem Skeptiker Richard Wiseman, der sehr erfolgreich darin war, die Effekte nicht hervorzurufen zu können (z. B. Wiseman & Greening, 2002). Beide hatten zuvor Experimente durchgeführt, die sich mit der außersinnlichen Blickwahrnehmung [staring detection] beschäftigten. Schlitz sprach von Beweisen für Psi (Schlitz & LaBerge, 1994) und Wiseman sprach davon, dass solche nicht vorlägen (Wiseman & Smith, 1994). Erstaunlicherweise einigten sie sich darauf, ein gemeinsames Experiment durchzuführen, in der sie beide einen identischen Versuchsaufbau verwenden und ihre Teilnehmer aus dem selben Versuchspersonen-Pool ziehen sollten, um herauszufinden, ob sie ihre unterschiedlichen Ergebnisse replizieren können. Der grundlegende Aufbau ihrer Studie bestand darin, dass entweder Wiseman oder Schlitz als ein Beeinflusser [influencer] (also als anschauende Person [starrer]) fungierten und die Versuchsteilnehmer (also die Angeschauten [staree]) mittels einer Überwachungskamera zeitweilig beobachteten. Während des Beobachtungsintervalls [staring trials] ging es darum, sich den Angeschauten mit der Absicht zuzuwenden, bei ihnen eine physiologische Reaktion hervorzurufen. Bei den „Nicht-Anschauen“-Durchgängen

[no stare trials] ging es darum, woanders hinzuschauen und ihre Aufmerksamkeit auf etwas anderes zu fokussieren. Die Abfolge von „Anschauen“- und „Nicht-Anschauen“-Sequenzen wurde von einer entsprechenden Software randomisiert vorgegeben. Währenddessen war es die Aufgabe der Angeschauten, sich bestmöglich zu entspannen. Ihre elektrodermale Aktivität (EDA) wurde fortlaufend aufgezeichnet. Ein etwaiger ASW-Effekt sollte sich in einem signifikanten Unterschied in der EDA zwischen „Anschauen“- und „Nicht-Anschauen“-Sequenzen abzeichnen.

Das erste Experiment (Wiseman & Schlitz, 1998) wurde in Wisemans Labor in Hertfordshire mit 16 Teilnehmern pro Versuchsleiter durchgeführt. Abgesehen davon, dass es zwei verschiedene Versuchsleiter bzw. anschauende Personen gab, waren die Versuchsbedingungen identisch. Wenn die Teilnehmer von Schlitz angeschaut wurden, waren ihre EDA-Werte, wie vermutet, signifikant erhöht, wohingegen die EDA-Werte von Wisemans Teilnehmern keine signifikanten Unterschiede aufwiesen. Dies bekräftigte also ihre bereits zuvor unabhängig voneinander erbrachten Ergebnisse. Ein zweites gemeinsames Forschungsprojekt (Wiseman & Schlitz, 1999) fand im Labor von Schlitz am *Institute of Noetic Sciences* (IONS) statt. Abermals machten beide Versuchsleiter von derselben Prozedur, derselben Ausrüstung und demselben Teilnehmer-Pool (jeweils  $N = 35$  für jeden Versuchsleiter) Gebrauch. Die Teilnehmer von Schlitz erbrachten abermals einen kleinen, aber statistisch signifikanten Effekt zustande, wohingegen dies bei Wisemans Teilnehmern nicht der Fall war; im Gegensatz zur ersten Studie waren die Teilnehmer von Schlitz allerdings in den Momenten, in denen sie angeschaut wurden, physiologisch signifikant *weniger* aktiviert als in den Momenten, in denen sie nicht angeschaut wurden. In einer dritten Studie wurden die Rollen des Versuchsleiters und der anschauenden Person mittels eines 2x2-Designs aufgeteilt. Einige der Teilnehmer wurden von derselben Person begrüßt und instruiert, von der sie später auch angeschaut wurden, die anderen Teilnehmer hingegen wurden von einem der beiden Versuchsleiter instruiert und später dann von dem jeweils anderen angeschaut. Während der Instruktion der Teilnehmer wussten die Versuchsleiter nicht, welcher Versuchsleiter später die Funktion der anschauenden Person übernehmen würde. Die Studie wurde wieder am IONS durchgeführt und die 100 Teilnehmer waren Mitarbeiter des Instituts oder Freiwillige aus der Umgebung. Der durchschnittliche Effekt fiel etwas stärker aus, wenn Schlitz die Teilnehmer begrüßte und ebenso, wenn Wiseman die Funktion der anschauenden Person übernahm. Doch keiner der Effekte wich signifikant von der Zufallserwartung ab. Dies ist bedauerlich, da die Studie versprach, zwischen zwei wichtigen Rollen zu differenzieren, doch ohne einen Nachweis von Psi konnte man auch keine Versuchsleitereffekte erforschen. Die Autoren vermerkten, dass sie beide eine verminderte Motivation und ein vermindertes Interesse am Experiment spürten. Dies konnte jedoch nicht als wissenschaftliche Begründung herangezogen werden (beispielsweise als Ausdruck eines *Decline*-Effektes) und das Verschwinden des Effektes könnte ebenso auf methodische Verfeinerungen zurückgeführt werden.

Es würde den Rahmen sprengen, auf Details weiterer Beispiele von Versuchsleitereffekten in der Parapsychologie einzugehen. Doch um die Allgegenwärtigkeit dieses Phänomens zu illustrieren, möchte ich gerne noch anführen, was Bem (2012) in Bezug auf Ritchie, Wiseman und French gesagt hat, welche vergeblich versucht hatten, seine „Retroactive Priming“-Effekte<sup>7</sup> zu replizieren: „Ich denke, dass die Erwartungshaltungen und die Einstellungen der Versuchsleiter in Bezug auf die Forschungsthese bedeutende Einflussfaktoren sind, die den Erfolg oder den Misserfolg von Replikationsversuchen bestimmen“ (S. 353). Auch Millar (Palmer & Millar, 2015) sieht den Versuchsleitereffekt als Ursache für den Misserfolg bei der groß angelegten, europäischen Replikationsstudie der Mikro-Psychokinese-Forschung des PEAR Forschungslabors.

### ***Erklärungsmodelle für den Versuchsleitereffekt***

Wenn nun also die These aufgestellt wird, dass ein Versuchsleiter das Ergebnis eines parapsychologischen Experimentes beeinflussen kann, wie lässt sich dies erklären? Smith (2003a) bietet vier Kategorien von Erklärungen an: (1) Betrug durch den Versuchsleiter, (2) fehlerhafte Durchführung des Experimentes durch den Versuchsleiter, (3) eine Interaktion zwischen Versuchsleiter und Teilnehmer sowie (4) durch den Versuchsleiter hervorgerufene Psi-Phänomene. Über die These des Betrugs als Erklärung habe ich bereits in einem früheren *Mindfield*-Artikel Überlegungen angestellt. Ich denke nicht, dass dies als hinreichender Faktor zur Erklärung des Effektes in Frage kommt. Obwohl unabsichtliche Fehler bei der Datenaufzeichnung mit größerer Wahrscheinlichkeit in Richtung der erwarteten Ergebnisse gehen, ist nicht davon auszugehen, dass dies die primäre Erklärung ist, da Versuchsleitereffekte in sehr vielfältigen Bedingungen (z. B. in Daten, die durch automatische Aufzeichnungen generiert sind, wie etwa bei den Blickwahrnehmungsexperimenten) beobachtet worden sind. (Schmeidler [1997: 83] hat beide Erklärungsansätze als „unseriös“ bezeichnet.)

Stattdessen werde ich mich auf die beiden Erklärungen fokussieren, die von Palmer und Millar (2015) genauer bestimmt wurden: Die Versuchsleiter-Psi-Hypothese sowie die Versuchsleiter-Verhaltens-Hypothese. Die erste Hypothese beinhaltet die Annahme, dass die Resultate von parapsychologischen Experimenten eigene Psi-Fähigkeiten des Versuchsleiters widerspiegeln. Die zweite Hypothese beinhaltet die Annahme, dass die Resultate von parapsychologischen Experimenten mit der jeweiligen Fähigkeit des Versuchsleiters zusammenhängen, eine positive Erwartungshaltung bezüglich des Versuchs herzustellen und die Versuchsteilnehmer in einen Zustand der Ungezwungenheit zu versetzen, sodass sie ihre eigenen Psi-Fähigkeiten wirken lassen können. Schmeidler (1997) sprach von solchen Versuchsleitern entweder als *psi-förderlich*

---

<sup>7</sup> Teilweise auch als „anomalous anticipation effects“ bezeichnet (Anm. d. Übers.).

oder als *psi-permissiv* (oder *psi-erlaubend*),<sup>8</sup> eine Unterscheidung, die ich als sehr nützlich erachte.

Da das Erklärungsmodell von Psi bei Versuchsleitern im aktuellen Überblick von Palmer und Millar (2015) eine größere Bedeutung zugewiesen wird – der Teil in dem Kapitel zu verhaltensbezogenen Ursachen macht eineinhalb Seiten aus, während der These von Psi bei Versuchsleitern dreimal so viel Platz gewidmet ist –, werde ich zunächst Überlegungen zu psi-förderlichen Versuchsleitern anstellen. Die Grundlage der These von Psi bei Versuchsleitern scheint die Vorstellung zu sein, dass Psi grenzenlos und zielorientiert funktioniert. Unter diesen Umständen sollten wir einräumen, dass diejenige Person, deren Wünsche am ehesten durch die statistischen Endergebnisse befriedigt werden kann, wohl eher der Versuchsleiter ist und nicht der Versuchsteilnehmer (insbesondere in den Fällen, in denen die Teilnehmer kaum noch in Verbindung mit der Studie oder den Forschungsergebnissen stehen, sobald das Abschlussgespräch geführt wurde). In der Tat scheinen mir einige durch Experimente gewonnene Daten – beispielsweise die aus der bereits beschriebenen Studie von West und Fisk – ohne die Annahme der These von Psi bei Versuchsleitern recht schwer erklärbar zu sein, da es quasi keinen Kontakt zwischen Versuchsleiter und Teilnehmer gab, der für die passende Atmosphäre gesorgt hätte, die wiederum es den Teilnehmern ermöglicht hätte, ihr eigenes Psi wirken zu lassen.

Zur Unterstützung dieser Überlegung führte Smith (2003b) eine Befragung von noch tätigen, experimentell arbeitenden Parapsychologen durch. Er entdeckte, basierend auf 40 Rückantworten, eine Korrelation zwischen Psi-Förderlichkeit und dem Glauben an die eigenen ASW- und PK-Fähigkeiten (dies könnte natürlich durchaus als ein Zirkelschluss interpretiert werden: gute Erfolge in den Experimenten erhöhen die Überzeugung, dass man selbst außergewöhnliche Fähigkeiten besitzt). Millar (in Palmer & Millar, 2015) beschreibt psi-förderliche Versuchsleiter als Virtuosen und schätzt, dass solche Personen „die Parapsychologie dominieren“, obwohl es womöglich zu jedem Zeitpunkt bloß etwa ein halbes Dutzend solcher Individuen in der aktiven, parapsychologischen Forschung gibt. Dies erscheint mir plausibel – und in der wissenschaftlichen Gemeinschaft gäbe es vermutlich eine große Übereinstimmung, wenn es darum ginge, die Namen der voraussichtlichen Kandidaten unter den aktuell experimentierenden Wissenschaftlern auszuwählen. Dies könnte zwar eine Erklärung für die eindrucksvollen Anfangsstudien aus ganz spezifischen Labors bringen (z. B. Bem, 2011; Bem & Honorton, 1997; Radin, 1997), jedoch nicht für die nachfolgenden, erfolgreichen Replikationsversuche, die an zahlreichen anderen Laboren stattfanden (z. B. Bem et al., 2016; Storm, Tressoldi & Di Risio, 2010; Mossbridge, Tressoldi, Utts, Ives, Radin & Jonas, 2014). Am verblüffendsten ist jedoch die Tatsache, dass Versuchsleitereffekte auch in solchen Fällen zu einem Nachweis von Psi-Effekten

---

8 Die englischen Begrifflichkeiten aus dem Original lauten hier „*psi conducive*“ sowie „*psi permissive*“ (Anm. d. Übers.).

führen könnten, in denen die zugrunde liegenden Studien von Forschern ohne Interesse an der Erfassung von Psi für nicht-parapsychologische Zwecke durchgeführt wurden (Bierman, 2000).

Natürlich könnte man auch der Ansicht sein, dass wir die Grenzen von Psi nicht kennen und es daher theoretisch möglich sein könnte, dass die ursprünglichen Versuchsleiter das Ergebnis scheinbar unabhängiger Replikationsversuche beeinflussen – auf unbewusste, zielgerichtete und bedürfnisorientierte Weise. Und hier ist der Haken. Diesen schrankenlosen Mechanismus in Bezug auf den psi-förderlichen Versuchsleitereffekt anzunehmen käme, so denke ich, dem Akzeptieren einer nicht falsifizierbaren Hypothese gleich. Obwohl dies die Hypothese an sich zwar nicht aufhebt, so würde es doch unsere Möglichkeit, systematische Fortschritte bei ihrer Überprüfung oder der Untersuchung der von ihr beeinflussten Phänomene zu erzielen, massiv einschränken. Mir scheint es vernünftiger, zunächst herauszufinden, wie weit wir bei der Erklärung von Versuchsleitereffekten kommen, wenn wir soziale und kognitive Einflussfaktoren berücksichtigen, bevor wir uns auf solche unerforschlichen und potenziell allmächtigen Mechanismen berufen.

### *Konventionelle Versuchsleitereffekte*

Wir müssen also darüber nachdenken, welche Arten von Versuchsleitereffekten auf üblicherem psychosozialen Wege hervorgerufen werden können, um zu beurteilen, ob einige oder gar alle Effekte, die wir in parapsychologischer Forschung entdecken, auch auf diese konventionelle Weise erklärt werden können. Die klassische Arbeit zu Versuchsleitereffekten von Robert Rosenthal und Kollegen (Rosenthal & Rosnow, 2009) zeigt, dass eine drastische Veränderung der Forschungsergebnisse durch eine Beeinflussung von bestimmten Merkmalen der Forscher, wie etwa die Erwartungshaltung und der Interaktionsstil, hervorgerufen werden kann. Rosenthals und Jacobsons (1968) Studie zum Pygmalion-Effekt beispielsweise zeigt, wie eine Erwartung selbst bewahrheitend werden kann. Alle Schüler einer Grundschule absolvierten einen non-verbalen Intelligenztest. Den Lehrern wurde gesagt, dass der Test diejenigen Kinder ermitteln könne, bei denen ein kognitives Aufblühen bevorstehe und die bald rasche Fortschritte zeigen würden. Etwa 20% der Kinder wurden als „Aufblühende“ gekennzeichnet. Tatsächlich aber wurden diese Kinder zufällig aus 18 Schulklassen, die über sechs Jahrgänge verteilt waren, ausgewählt. Am Ende des Schuljahres, etwa acht Monate später, wurde ein IQ-Test mit den Kindern durchgeführt und man stellte fest, dass diejenigen, die als Aufblühende bezeichnet wurden, sich im Vergleich zu ihren Mitschülern um vier IQ-Punkte verbessert hatten (zwei Punkte im Bereich des Verbalen sowie sage und schreibe sieben Punkte im Bereich des logischen Denkens). Pygmalion-Effekte wurden daraufhin nicht nur in Klassenzimmern, sondern auch in Gerichtssälen, in Altenpflegeheimen, im Bereich der Unternehmensführung und sogar in Schwimmbädern nachgewiesen (Rosenthal, 1997: 178).



Die Effekte sind nicht auf menschliche Teilnehmer beschränkt, welche ja für verbale und nonverbale Signale einer Autoritätsperson sensibilisiert sein könnten. Auch Tiere sind dem Versuchsleitereffekt unterworfen, wie in Studien gezeigt wurde, die das Verhalten von Ratten bei Lernexperimenten in einem Labyrinth [maze learning in rats] untersuchten (Rosenthal & Rosnow, 2009: 423–439). Den Teilnehmern wurde gesagt, dass das Kreuzen von Ratten, die gute Erfolge beim Erlernen des Labyrinthes hatten, nachfolgende Generationen von Ratten produzieren würde, die deutlich besser abschnitten als solche, die nicht selektiert wurden. Gleichermaßen würden die Ratten, die von weniger erfolgreichen Individuen hervorgebracht wurden, Nachkommen zeugen, die schlechter als der Durchschnitt abschnitten. Jedem der 12 Studenten wurden fünf Ratten zugewiesen, von denen alle entweder der im Labyrinth erfolgreichen oder der weniger erfolgreichen Abstammungslinie zugehörig waren. Tatsächlich wurden die Ratten allerdings randomisiert in Gruppen eingeteilt. Hierbei wurde gewährleistet, dass die Gruppen jeweils das gleiche Durchschnittsalter haben, da der Reifungsgrad die Leistungsfähigkeit der Ratten beeinflussen könnte. Jede Ratte wurde über eine Zeitdauer von fünf Tagen zehnmal täglich in das Labyrinth gesetzt. Der Versuchsleiter zeichnete auf, ob der Versuchsdurchgang [trial] erfolgreich oder nicht erfolgreich war. Die Ratten aus der erfolgreichen Abstammungslinie schnitten an den Tagen 1, 4 und 5 sowie insgesamt signifikant besser ab als die Ratten aus der weniger erfolgreichen Abstammungslinie. Die Ratten aus der erfolgreichen Abstammungslinie entsprachen über die fünf Tage eher der erwarteten „Lernkurve“, während die Ratten aus der erfolglosen Abstammungslinie zu Beginn eine Verbesserung zeigten, die dann abflaute. Dies deutet darauf hin, dass der differenzielle Effekt recht unmittelbar auftrat und nicht erst nach und nach durch regelmäßigen Kontakt wirksam wurde. 20% aller Versuchsläufe schlugen dadurch fehl, dass die Ratten sich überhaupt gar nicht bewegten. 17 dieser Fälle kamen durch Ratten der erfolgreichen Abstammungslinie zustande und 43 durch die der weniger erfolgreichen. Selbst wenn diese Fälle als möglicher Störfaktor aus der Statistik ausgeschlossen werden, so waren die Ratten aus der erfolgreichen Abstammungslinie in erfolgreichen Versuchsläufen immer noch signifikant schneller als die aus der weniger erfolgreichen Abstammungslinie. Zwar wurden die Interaktionen zwischen den Versuchsleitern bzw. den Teilnehmern mit ihren Versuchstieren nicht formal erfasst, doch waren in der Gruppe der Versuchsleiter auch zwei eingeweihte Helfer der Studienleiter, die Rückmeldung zur Untersuchungspraxis geben konnten. Sie konnten keinerlei falsche Eintragungen oder Verfälschung der Daten feststellen, wohl aber einige Male beobachten, dass die Versuchsleiter die Ratten zur Bewegung drängten, indem sie sie anstupsten. Jedoch geschah dies bei den Ratten aus der weniger erfolgreichen Abstammungslinie lediglich geringfügig häufiger als bei den Ratten aus der erfolgreichen Abstammungslinie (3 Fälle vs. 2 Fälle), sodass sich dies nicht auf die unterschiedliche Leistung der beiden Gruppen ausgewirkt haben wird. Bedeutender ist die Tatsache, dass die Versuchsleiter, die die Ratten aus der erfolgreichen Abstammungslinie untersucht haben, sowohl einen häufigeren als auch

einen behutsameren Umgang mit diesen pflegten. Dies könnte ein Mechanismus sein, der zum unterschiedlichen Verhalten der beiden Gruppen beim Laufen durch das Labyrinth geführt haben könnte.

Nach diesen ziemlich bemerkenswerten frühen Ergebnissen wurden interpersonelle Erwartungs-Effekte in einer großen Anzahl von Studien bestätigt. Rosenthal (1994: 176) spricht von 464 Studien mit einem Gesamt- $d$  von .63 ( $r = .30$ ),<sup>9</sup> welches den Effekt in einer Vielzahl von Kontexten illustriert. Dazu gehören Studien zur Reaktionszeit, zur Ausdeutung von Tintenkleckschen, zum Lernverhalten von Tieren, zur Wahrnehmung von Personen sowie zum Erlernen von Fertigkeiten. Harris und Rosenthal (1985) lieferten eine Meta-Analyse von 135 Studien, die auf 16 Verhaltensweisen fokussieren, bei denen man einen Zusammenhang mit dem Effekt vermutete. Dazu gehören ein warmes, zwischenmenschliches Klima, die Erwartungshaltung der Versuchsleiter, die Fokussierung der Aufmerksamkeit sowie Rückmeldungen[feedback] (Harris & Rosenthal, 1985; Rosenthal, 1994).

### *Eigenschaften psi-permissiver Versuchsleiter*

Interessanterweise haben Parapsychologen, die sich für Versuchsleitereffekte interessieren, ähnliche Einflussfaktoren betont. Schmeidler und Maher (1981) haben Forscher auf Video aufgezeichnet, während sie auf einer wissenschaftlichen Konferenz einen Vortrag hielten und Fragen beantworteten. Unabhängige Begutachter sichteten das Material von fünf „psi-förderlichen“ und fünf „psi-hemmenden“ [psi-inhibitory] Forschern, deren Auswahl zuvor anhand des Alters und physischer Merkmale abgestimmt worden war. Die Sachverständigen bewerteten sie dann anhand einiger Kategorien. Psi-förderliche Forscher wurden als flexibler, enthusiastischer, freundlicher, sympathischer, wärmer und als weniger angespannt, schwerer reizbar und weniger kalt beurteilt. Natürlich ist es möglich, dass die Art und Weise, wie sich Forscher auf einer parapsychologischen Konferenz präsentieren, abhängig ist vom Erfolg oder anderen Aspekten der Forschung, die sie präsentieren, sowie von ihrem Gefühl dafür, wie sie von ihren Fachkollegen wahrgenommen werden könnten. Es ist aber ermutigend, dass die Faktoren, die die Teilnehmer zu beruhigen scheinen und es ihnen erlauben, sich offen und natürlich zu verhalten, diejenigen sind, die Erfolg bei einer psi-bezogenen Aufgabe vorhersagen.

Einige Forscher haben versucht, diese Faktoren experimentell zu beeinflussen. Honorton, Ramsay und Cabibbo (1975) beispielsweise führten eine Untersuchung mit zwei Versuchsleitern durch, die während des Kontaktaufbaus mit dem jeweiligen Teilnehmer entweder auf positive Weise (freundlich, locker, unterstützend) oder auf negative Weise (abrupt, formell und

---

9 Cohens  $d$  bezeichnet das Differenzmaß,  $r$  den Korrelationskoeffizient. Ab  $r = .30$  wird von einem mittleren Zusammenhang gesprochen (Anm. d. Übers.).

unfreundlich) agierten. Im letzteren Fall ging der Versuchsleiter schnell zur Versuchsaufgabe über. Die positive Umgangsweise erbrachte signifikant höhere Ergebnisse als die negative. Als jedoch Schneider, Binder und Walach (2000) den Interaktionsstil des Versuchsleiters hinsichtlich seiner Herzlichkeit bzw. des Grades der zwischenmenschlichen Wärme (persönlich vs. neutral) veränderten, fanden sie keinen Unterschied zwischen den beiden Versuchsbedingungen. Auf ähnliche Weise hat Parker (1975) die Erwartungshaltung von sechs Versuchsleitern (Studenten) beeinflusst, sodass sie sich entweder als stark gläubig oder als stark ungläubig zeigten. Obwohl das Gesamtergebnis nicht signifikant war, gab es einen signifikanten Unterschied in den Ergebnissen der beiden Gruppen von Versuchsleitern. Und das trotz der Tatsache, dass sie die gleichen Sender-Empfänger-Paare testeten. In der Praxis ist es ein sehr schwieriges Unterfangen, den Interaktionsstil zu variieren, ohne unauthentisch zu wirken. Eine bessere Methode wäre es, eine Reihe von Versuchsleitern einzusetzen, die sich auf natürliche Weise in ihrem zwischenmenschlichen Verhalten unterscheiden. Watt und Ramakers (2003) fanden heraus, dass Teilnehmer, die an einem Versuch teilnahmen, in dem es um Fernunterstützung zur Aufmerksamkeitsregulierung [remote facilitation of attention task] ging, signifikant besser abschnitten, wenn sie von einem Versuchsleiter betreut wurden, der selbst tatsächlich an eine Wirkung glaubte. Teilnehmer, die von einem nicht-gläubigen Versuchsleiter betreut wurden, erzielten Ergebnisse, die nicht von der Zufallserwartung abwichen.

Eine weitere Schwierigkeit, die sich bei solchen Studien ergibt, liegt darin, dass man sicherstellen möchte, dass die beabsichtigte Beeinflussung (z.B. der Glaubensvorstellungen der Teilnehmer oder der Herzlichkeit) auch die erwartete Wirkung auf die Teilnehmer hat. Schmeidler (1997) hat bemerkt, dass sich dies in der Praxis von Fall zu Fall je nach Vorlieben der Teilnehmer unterscheiden kann – was für den einen beruhigend sein kann, bringt den anderen vielleicht an die Grenze. Vertraulichkeit und Nähe beispielsweise mag bei einigen Teilnehmern zum gewünschten Kontaktaufbau führen, könnte aber von anderen als aufdringlich und unangenehm empfunden werden. Diesem Problem kann dadurch begegnet werden, dass die Versuchsleiter sich einfach so verhalten, wie sie sich natürlicherweise verhalten würden, um die Teilnehmer dann im Anschluss darum zu bitten, die Interaktion mit dem Versuchsleiter zu bewerten. Ich war an Ganzfeld-ASW-Studien und an PK-Studien beteiligt, in denen diese Methode angewandt wurde. Die Teilnehmer füllten nach der Interaktion, aber noch vor der Psi-Aufgabe, einen Fragebogen aus (mit der Zusicherung, dass ihre Antworten von einem unabhängigen Begutachter ausgewertet würden und der Versuchsleiter keinen Zugriff auf die Daten haben werde). In einer Studie (Sherwood, Roe, Holt & Wilson, 2005) fanden wir heraus, dass Erfolge in Ganzfeld-Experimenten mit einem höheren Grad an Entspannung, Optimismus und Zuversicht bei den Sendern und den Versuchsleitern sowie mit einem höheren Grad an Zuversicht bei den Empfängern zusammenhing. Bei einer PK-Aufgabe (Roe, Davey & Stevens, 2006) fanden wir einen Zusammenhang zwischen guten Ergebnissen und positiver Stimmung,

positiver Einstellung gegenüber der Versuchsaufgabe sowie Entspannung und besonders dem Vertrauen des Versuchsleiters in den Erfolg des Experimentes. Dies legt in starkem Maße nahe, dass wir in Bezug auf diejenigen Aspekte der Interaktion sensitiv reagieren können, die eine direkte Auswirkung auf die Psi-Fähigkeit der Teilnehmer haben.

In diesem Kontext lohnt es sich über die wertvollen Erkenntnisse von Caroline Watt bezüglich der Absichten und Interaktionsstile der psi-permissiven und psi-hemmenden Versuchsleiter Schlitz und Wiseman (Watt, Wiseman & Schlitz, 1998) nachzudenken. Watt führte nach Abschluss zweier der weiter oben beschriebenen Studien Interviews mit den beiden Versuchsleitern, um zu überprüfen, ob sich vielleicht Unterschiede im Untersuchungsverhalten feststellen ließen, die zu den unterschiedlichen Ergebnissen geführt haben könnten. Es kam heraus, dass Schlitz vor jedem Versuchsdurchgang ein Vorbereitungsritual durchführte, das ihr half, sich auf den Durchgang und den Teilnehmer zu fokussieren. Zudem setzte sie sich einen klaren Vorsatz für die Sitzung. Bei der Begrüßung der Teilnehmer bemühte sie sich darum, die Interaktion individuell anzupassen und den Kontakt aufzubauen sowie das Phänomen zu normalisieren und Erfolgserwartungen zu erwecken, die auf vorhergehenden Erfahrungen und Forschungsergebnissen basieren. Sie überlegt (S. 23): „Ich habe versucht, [den Teilnehmern] jeden möglichen Grund dafür aufzuzeigen, sich optimistisch zu fühlen, dass diese konkrete Sache, die wir gerade gemeinsam taten, tatsächlich etwas bewirken könnte.“

Während des Abschlussgesprächs rundete sie die Interaktion ab, indem sie die erhobenen Daten besprach und interpretierte. Im Gegensatz dazu beschrieb Wiseman, wie seine Vorbereitung auf praktische Fragestellungen beschränkt war, wie z. B. sicherzustellen, dass die notwendigen Gerätschaften und Materialien organisiert werden. Seine Vorbesprechung mit den Teilnehmern war eher „sachbezogen“ und trotz der Tatsache, dass er seinen Skeptizismus nicht angesprochen hatte, ist es wahrscheinlich, dass er mit seiner Haltung klar erkannt wurde: „Manchmal sagte ich: ‚Mir ist bewusst, dass sich das nach einem etwas merkwürdigen Experiment anhört, aber lass es uns einmal ausprobieren.‘ Und wenn der Teilnehmer skeptisch war, sagte ich: ‚Na gut, lass es uns versuchen.‘“ Tatsächlich vermutete Wiseman, dass sein Skeptizismus für seine Versuchsteilnehmer offensichtlich gewesen sei: „Es gibt wahrscheinlich große Unterschiede in dem Erstgespräch mit ihnen, allein schon darin, wie sehr sie losgingen mit dem Gedanken daran, dass dies tatsächlich ein funktionierendes Verfahren sei“ (S. 23). (Smith [2003a: 75] merkt an, dass die Teilnehmer von Schlitz einen stärker ausgeprägten Glauben an Psi aufwiesen und vermutet, dass der Wiseman-Schlitz-Effekt dadurch beeinflusst worden sein dürfte, dass die Versuchsleiter den Versuchsteilnehmern ihre Erwartungen mitteilten.) Wiseman erklärte zwar klar und deutlich, worum genau es in der Studie ging, versuchte aber nicht, Kontakt mit den Versuchsteilnehmern aufzubauen: „Nein, größtenteils gab es da nicht besonders viel sonstiges Geplauder. Ich würde es eher als geschäftsmäßig beschreiben, aber nicht als unfreundlich“ (S. 21). Auf die spontanen Erfahrungen, die berichtet wurden, wurde

nicht eingegangen, obwohl damit ein positiver Vorsatz (z. B. ‚ich werde es schaffen‘) hätte gesetzt werden können: „Wenn jemand von ihnen damit begann, ihre Erfahrungen mit mir durchzusprechen, empfand ich es als ziemlich schwer, darauf einzugehen, da ich ziemlich skeptisch bin, was solche Dinge angeht. Und man will ja auch nicht konfrontativ sein, also nickt man einfach und sagt: ‚Ah, das ist interessant. Wie auch immer, zurück zum Experiment“ (ebd.). Diese Beschreibungen zeigen offenkundig, wer von den beiden erwartungsgemäß als psi-permissiv und wer als psi-hemmend gelten kann.

### *Fazit*

In diesem Artikel habe ich versucht zu zeigen, dass Replikation in der Parapsychologie weniger zuverlässig als in den Naturwissenschaften sein könnte aufgrund der zahlreichen zwischenmenschlichen Faktoren, die einen Einfluss auf das Ergebnis haben können. Diese Faktoren wurden unter der Überschrift Versuchsleitereffekte zusammengefasst, um zu zeigen, wie subtil das Verhalten und die Erwartungen der Versuchsleiter das Ergebnis der Studie gestalten können. Diese Zusammenhänge zu verstehen, könnte das Auftreten der Phänomene, die wir erforschen wollen, besser vorhersagbar machen. Meiner Auffassung nach hängt dies davon ab, dass wir akzeptieren, dass die Sozialwissenschaften nicht wie die Naturwissenschaften funktionieren, da der Forscher nicht von dem System zu trennen ist, das er oder sie erforscht; dementsprechend ist die Forschungspraxis auch eine Kunst, bei der es darum geht, das eigene Bewusstsein dafür zu schärfen, wie man die Teilnehmer ermächtigt, ihr eigenes Potential umzusetzen. Es geht nicht nur um die Beherrschung spezieller Techniken oder Methoden.

Natürlich sind viele innerhalb der Sozialwissenschaften gewillt, dieses eher chaotische (und potentiell eigenartige) Verständnis der Versuchsleiter-Versuchsteilnehmer-Dyade zu verwerfen und bevorzugen stattdessen ein simpleres Modell, bei dem Teilnehmer nur das tun und denken, was von ihnen im Rahmen der experimentellen Beeinflussung erwartet wird. Trotz eines zu Beginn noch großen Interesses an der Arbeit von Rosenthal und dem Anwachsen einer überzeugenden Beweisgrundlage für das Auftreten von Versuchsleiter-Effekten wird das Thema in der heutigen Literatur der Mainstream-Psychologie kaum behandelt. Tart (2016: 54) beobachtet:

Das Thema ist quasi aus der psychologischen Literatur verschwunden. Man könnte annehmen, dass man sich bereits in der Vergangenheit mit dieser Frage befasst hat und sie nun als gelöst verwerfen kann. Nein, man hat lediglich aufgehört, über sie zu schreiben – meine Vermutung war, dass eine solche Gefährdung für die gebotene Objektivität psychologischer Studien einfach zu bedrohlich war und niemand darüber nachdenken wollte. (...) [V]or gerade einmal einhundert Jahren haben sie uns aus dem Institut für Philosophie freigelassen und wir fürchteten uns davor, dass sich unsere Behauptungen,

richtige (Natur-)Wissenschaftler zu sein, als unzutreffend herausstellen könnten und sie uns wieder zurück ins Institut für Philosophie schicken würden – verständliches, menschliches Verhalten, schlechte Wissenschaft.

Allerdings.

(aus dem Englischen von Max Fuhrmann)

### Literatur

- 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. In *Proceedings of Presented Papers: The Parapsychological Association 43rd Annual Convention*, (S. 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* (pp. 184–222). London: Routledge & Kegan Paul.
- Palmer, J., & Millar, B. (2015). Experimenter effects in parapsychological research. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Hrsg.), *Parapsychology: A handbook for the 21<sup>st</sup> century* (S. 293–300). Jefferson, NC: McFarland & Co.
- 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 (Hrsg.), *Research in parapsychology 1974*, (S. 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 Psychological 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. In *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention* (S. 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 (Hrsg.), *Psi wars: Getting to grips with the paranormal* (S. 69–84). Charlottesville, VA: Imprint Academic.
- 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 [Brief]. *Journal of the Society for Psychological 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, 8–25.
- West, D. J., & Fisk, G. W. (1953). A dual ESP experiment with clock cards. *Journal of the Society for Psychological 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. In *Proceedings of Presented Papers: The Parapsychological Association 42nd Annual Convention* (S. 471–479).
- Wiseman, R., & Smith, M.D. (1994). A further look at the detection of unseen gaze. In *Proceedings of Presented Papers: The Parapsychological Association 37th Annual Convention* (S. 465–478).