Experimenter Misconduct in Parapsychology: Analysis Manipulation and Fraud

James E. Kennedy
Version of February 9, 2014
Reproduction is authorized per Copyright Notice at the end of this article.

Key conclusions: Experimenter misconduct has occurred many times in parapsychology and is a constant threat. Undetected cases are likely. Routine practices for confirmatory experiments should include: (a) prospective public registration of experiments, (b) experimental procedures that make intentional or unintentional data alterations by one person difficult, and (c) sharing data for analyses by others. Research organizations and funding sources should require these practices. A registry for parapsychological experiments is available at https://koestlerunit.wordpress.com/study-registry/.

Experimental research in parapsychology has certain characteristics that could be considered warnings of possible fraud or other misconduct in science. Most parapsychological experiments have not obtained statistically significant outcomes (Kennedy, 2013) and most experimenters have found that many of their experiments were nonsignificant—sometimes seemingly capriciously nonsignificant (Kennedy, 2003). However, a few experimenters have reported significant results on almost every experiment. These experimenters provide a disproportionate amount of the overall evidence for psi.

These experimenter differences have long been recognized in parapsychological research (Kennedy and Taddonio, 1976; Office of Technology Assessment, 1989; Palmer, 1997; Rao, 2011, pp. 170-197; White, 1976). The potential for controversy is enhanced by the fact that the cause of the experimenter effects remains a matter of debate. The tendency for skeptics to obtain nonsignificant results is well known, but whether skepticism is more a cause or result of nonsignificant outcomes has not been resolved.

J.B. Rhine believed that only certain experimenters have the knack for successfully doing parapsychological experiments. He argued that parapsychological experiments
should be done by those with the knack for obtaining significant results (Rhine and Pratt, 1957, page 132). Those without the knack should find something else to do.

Given the prominent experimenter effects in parapsychology, speculations about misconduct have been common within parapsychology, as well as by outside critics. As described below, evidence of misconduct has been found in some cases.

The primary purpose of this paper is to describe established methods to reduce the threat of experimenter misconduct and to increase confidence in experimental results. These methods are consistent with standard practices in pharmaceutical research and are appropriate in other situations when research findings can be expected to be professionally challenged. This paper addresses methodology and is not intended to explicitly or implicitly take a position on whether paranormal phenomena exist.

This paper discusses two categories of experimenter misconduct: analysis manipulation and data manipulation. Analysis manipulation focuses on practices during the analyses and reporting of data that create a biased, misleading impression of the results. Data manipulation focuses on overt fraud by an experimenter. The related topic of deception by subjects in parapsychological experiments has been discussed by Hansen (1990).

Analysis Manipulation

The term analysis manipulation is used here for a range of practices that can significantly distort reported experimental findings. These practices include various forms of selective reporting and misleading descriptions of methodology, such as:

- Planning to analyze multiple different hypotheses for an experiment, but only reporting the results that support the experimenter's expectations, without mentioning or correcting for the other analyses.
- Planning a vague hypothesis and then determining the specific statistical test and precise hypothesis as the data are being explored during analyses.
- Reporting exploratory or post hoc findings in a way that can be mistaken for planned analyses.
- Failing to report experiments with results that do not support the experimenter's expectations.
- Adapting the description of the methodology or findings to conform to the comments of referees during publication.

The point about multiple hypotheses can easily occur in a typical psi experiment. An experimenter may plan to analyze for overall psi and also plan to analyze the relationship between psi scores and several other factors. If one of the analyses comes out significant but not the other analyses, the experimenter can be tempted to report the significant result as a planned analysis without mentioning or correcting for the fact that it was selected from several analyses. One symptom of this practice is an experimental
design that has an obvious main analysis, but the experimental report presents only some type of more complicated relationship between psi and other variables. Unfortunately, this practice can be difficult to discern from published reports.

When the planned hypothesis is vague, the main analysis for the experiment has a significant post hoc component. This substantially increases the likelihood of obtaining results that the experimenter wants and is difficult to detect in the published report. The practice of vague hypotheses can be combined with multiple hypotheses to create situations where some type of significant result is likely by chance.

Small studies are particularly susceptible to not being reported if the results are nonsignificant. Small studies typically have low statistical power and therefore are inconclusive. Also they tend to be more exploratory. However, small studies that obtain significant results typically are reported, which introduces significant biases in the literature.

The point about adapting to a referee’s comments is based on my experience as a referee for parapsychological articles. In some cases a revised manuscript presented methodology or findings sufficiently differently from the original manuscript that I wondered what had actually occurred for the experiment. However, in those cases there was little that could be done except to take the author’s word about which version was more accurate.

The need to eliminate these types of practices has been noted in many writings in parapsychology (e.g., Akers, 1984; Bösch, Steinkamp and Boller, 2006a, 2006b; Hyman & Honorton, 1986; Kennedy, 2004; Milton, 1999; Murray, 2011; Office of Technology Assessment, 1989; Rhine, 1974a; Schmeidler & Edge, 1999; Stokes, 1997; Watt, 2005).

These types of practices are very powerful at generating misleading experimental findings. They can occur virtually unconsciously as the experimenters enthusiastically think about what their data mean and how to efficiently present the findings. Many other areas of statistical research are susceptible to analysis manipulation. However, most of those areas have less controversial implications and are less subject to experimenter effects than parapsychology.

**Addressing Analysis Manipulation**

Analysis manipulation is inexcusable because this misconduct can be easily prevented. Numerous writings and discussions in parapsychology over the years have pointed out the need for some type of study registration prior to conducting psi experiments (e.g., Bösch, Steinkamp and Boller, 2006a, 2006b; Hyman & Honorton, 1986; Kennedy, 2004; Milton and Wiseman, 2001; Murray, 2011; Rhine, 1974a; Schmeidler & Edge, 1999; Watt, 2005).

Proper registration would virtually eliminate the problems from analysis manipulation. Exploratory or pilot studies could continue without registration as occurs at present, but would be clearly designated as exploratory. Prospective, public
registration of experimental research is rapidly becoming required practice in medical research (De Angelis, et al., 2004; U.S. National Institutes of Health, 2012). Note that the registration must be done before data collection begins and in a way that is accessible to the public. A registration process is not acceptable if the registrations have restricted access or if the experimenter can decide after the data are analyzed whether to make the registration openly available.

Compared to academic psychology, medical researchers much more frequently (a) prospectively register studies, (b) formally designate hypotheses as exploratory or confirmatory, (c) recognize the limited value of small, underpowered studies, and the much greater value of well-powered confirmatory experiments, and (d) use formal power analysis to establish the sample size in experiments. The better methodology for medical research is not surprising because medical research typically has much greater practical implications for human health and for life and death decisions than research in academic psychology. Parapsychology has generally followed the methodological practices of academic psychology.

Given the controversial nature of parapsychology, the methodological practices of medical research would be more appropriate than the typical practices of academic psychology. Registering studies is a much needed first step. Unfortunately, my efforts a few years ago to generate interest in a registry found that some of the most successful experimenters were the least interested in developing a registry. That experience was consistent with my suspicion that analysis manipulation has been common for certain highly successful experimenters in parapsychology.

As described in the conclusions later, a study registry has now been implemented for parapsychological experiments. The registry is most important for confirmatory experiments, but exploratory studies can also be registered. At a minimum, the registry will document the overall study purpose and design, the planned hypotheses, and the planned sample size. If these factors cannot be specified in advance, then the experiment is more exploratory than confirmatory, and the limited evidential value should be recognized. Registering more information is highly recommended. The optimal practice is to describe specific statistical tests and prepare a detailed written protocol.

Registration of experiments is important for skeptical experimenters as well as for proponents of psi. Skeptics sometimes use analysis manipulation to negate positive results (Kennedy, 1981). Analysis manipulation applies to any experimenter biases, not just biases for significant results.

**Data Manipulation**

Experimenter fraud is an established factor in scientific research (Stroebe, Postmes, and Spears, 2012). The extent of occurrence of fraud is unknown because undetected instances are likely and some organizations probably avoid publicizing cases of fraud. Fraud is generally assumed to be more likely when the risks of detection are low and
when there are financial incentives and professional pressures for definitive experimental results. The prominent cases of experimenter fraud are often by researchers who are described as brilliant rising stars with impressive credentials and affiliations (Stroebe, Postmes, and Spears, 2012).

In early 1974 J.B. Rhine (1974a) published a paper on experimenter fraud in parapsychology. In the paper he stated “I have selected a dozen cases to illustrate fairly typically the problem of experimenter unreliability prevalent in the 1940’s and 1950’s” (page 104). He also stated “Fortunately, the culprits have thus far been caught (at least in our ‘known’ cases) before serious damage has been done” (page 105). In addition, he described three more recent cases of fraud or clearly inappropriate experimenter behavior that made the results unsuitable for publication. One of his main points in this paper was that “we have been able to do quite a lot to insure that it is impossible for dishonesty to be implemented inside the well-organized psi laboratory today” (page 105).

Unfortunately, in the next issue of the same journal, Rhine (1974b) reported that the man he had appointed as Director of his laboratory had been exposed by coworkers as committing fraud. When the final report was issued (Rhine, 1975), the fraud by W. J. Levy was found to be extensive and all his research was dismissed. Levy was one of the people who seemed to obtain significant results on every experiment and was considered a brilliant rising star. I was involved in exposing Levy and the efforts to determine the extent of the fraud. More information on the Levy case can be found at Kennedy (2012).

Another prominent case of fraud in parapsychology was S.G. Soal in Britain. Soal’s work with two special subjects guessing cards in the 1940s had been considered some of the best evidence of psi (Beloff, 1993). The experiments were observed by many researchers. One of Soal’s coworkers is reported to have accused him of altering data, but the matter was not pursued due to the threat of legal action. However, later analysis of the data in the 1970’s reported unexpected findings that might be consistent with fraud. The matter is generally considered to have been resolved by Betty Markwick (1978) who analyzed Soal’s data and provided strong evidence for manipulation of the data (Beloff, 1993).

In addition to the 17 cases noted above, there are other cases with suspicions of wrongdoing but without adequate evidence. For example, in a private conversation after the Levy incident, J.B. Rhine told me that at one point a researcher at the lab told him that he believed Levy’s research methods were improper. At the same time, Levy told Rhine that he believed the other researcher was doing work that was improper. And, a third researcher at the parapsychology lab told Rhine that he had doubts about the integrity of both Levy and the other man. Rhine’s point was that he did not find the arguments convincing for any of these speculations. I have also heard stories about another case of accusations of fraud without convincing evidence, and have heard suspicions of fraud mentioned in passing in various conversations about other researchers. Gardner Murphy (1961, pp. 282-284) commented that he had doubts about
the integrity of some parapsychological research and discussed the ethical dilemmas in publically expressing such suspicions.

Carl Sargent's ganzfeld research is another situation with speculations of fraud, but without convincing evidence. While visiting Sargent's lab, Blackmore (1987) reported observing some irregularities in procedure that might be consistent with a poorly defined and implausible (in my opinion) hypothesis of fraud. Sargent (1987) and his coworkers (Harley and Mathews, 1987) denied any significant wrongdoing and argued that the irregularities were inconsequential random deviations. Sargent responded with indignation and refused to provide copies of the data for verification, even when requested by the Parapsychological Association. Soon Sargent left parapsychological research. As implied by Beloff (1993, pages 283-284), Sargent's behavior in this incident unfortunately may give the appearance that there was something to hide even though convincing evidence was not obtained.

Expressing suspicions of fraud without convincing evidence creates an untenable situation that tends to discredit everyone involved. These cases often generated extremely bad feelings and significantly affected the careers of the accusers as well as the accused. Decision makers were faced with either ignoring potential misconduct or adversely affecting a person's credibility and career without convincing justification. Unfortunately, if effective measures to prevent fraud are not part of the research culture, convincing resolution can be difficult to obtain and these untenable situations can be expected.

Based on my experience exposing fraud, it is not surprising that many suspected cases do not reach the point of convincing resolution. The need to maintain normal interactions with a close colleague while covertly planning and conducting steps for his exposure requires a degree of acting and compartmentalization that many scientists do not have. For me it was very difficult. Many pivotal decisions must be made in secrecy and under stress. In addition to the strategy and technical details for collecting unequivocal evidence, multiple people need to be involved to establish overwhelming credibility. Decisions must be made about who can handle the acting and extreme secrecy, how they should be approached, the risks of possible compromising communication, and the roles for the various people. These distasteful steps are necessary to resolve the matter rather than creating another intrinsically untenable situation with no clear resolution.

I have found that working in an environment with routine practices to prevent fraud is much preferable to my experiences in parapsychology.

The motivations associated with fraud sometimes appear to defy common sense. It is beyond my comprehension how Levy could work so hard and be so dedicated to findings that he was fraudulently producing. Stokes (1997, pages 95-96) reported a similar reaction. On one line of research I found effects that were later recognized as artifacts of his fraudulent activities. When I originally discovered these effects, Levy developed
hypotheses about interesting new properties of psi and encouraged research on the effects. His behavior appeared to embrace the artifacts of his fraud as real effects.

My overall impression is that Levy’s fraud was not rationally planned and was not a coherent effort to achieve a particular goal. For the maze computer game experiments, Levy developed the computer program to analyze the data. When we were investigating the extent of his fraud, I thought it was possible that he manipulated the analysis programming to produce fraudulent results. A colleague thought I was wasting time investigating this because Levy would not have done something that was so easily detected. My investigation revealed that we were both wrong. Levy had simply fabricated the published results and made no effort to make the permanently stored data or analysis program match the published figures.

The psychological factors for Soal were equally strange (Beloff, 1993, pages 146-148; Mauskopf and McVaugh, 1980). Soal reported consistently nonsignificant results for experiments over many years and was reported to have become cynical about the positive results claimed in the U.S. At the persistent insistence of a colleague, Soal reportedly examined his data for displacement effects and found that two subjects consistently called a target that was one ahead or one behind the designated target. This effect was then confirmed with many subsequent fraudulent experiments.

Of course, these cases should not be taken as evidence that all fraudulent experimenters have irrational or careless behavior. It is, unfortunately, very possible that more rational, careful fraudulent experimenters have successfully avoided detection.

**Addressing Data Manipulation**

Two effective strategies for reducing fraud are (a) experimental procedures that incorporate checks by coworkers and (b) making raw data available to others. A recent analysis of cases of scientific fraud reported that most frauds are detected by whistleblowers inside an organization and that “fraudsters are usually reluctant to make available the data they allegedly collected” (Strobe, Postmes, and Spears, 2012, p. 682). The authors concluded that “whistleblowers are likely to remain the single most effective instrument against scientific cheating” (p. 682).

Independent replication and peer review for publication are not deterrents to fraud. These practices have generally not been effective at detecting even extensive fraud (Strobe, Postmes, and Spears, 2012), and do not pose a significant risk of detection for those contemplating fraud. The experiences with fraud in parapsychology described above also support the importance of coworkers and independent data analyses, and the ineffectiveness of replication and peer review. In parapsychology, the prominent experimenter effects make replication virtually worthless as an indication of possible fraud.

The goal of designing studies such that fraud by one experimenter would be difficult and risky appears to be reasonable. Rhine (1975), Akers (1984) and Dalton, Delanoy,
Morris, Radin, Taylor, and Wiseman (1996) reached similar conclusions. Without such practices, a person can commit fraud with little risk of adverse consequences. Experimenter fraud should not be easy and tempting in parapsychological experiments. The more extreme goal of claiming that experimental designs can make fraud by one experimenter impossible appears to me to be problematic for typical experimental work. Great effort is required to imagine and anticipate all possibilities, particularly in this age of technology. Similarly, collusion by multiple experimenters is also possible, but for typical experiments the risks are not worth the extraordinary effort that would be required to address this possibility through experimental design. However, extremely controlled experiments are useful as special cases. Some of the most carefully controlled experiments in the history of parapsychology have been done by Schmidt with various observers (Schmidt, Morris, and Rudolph, 1986; Schmidt and Stapp, 1993).

For reasonable measures in typical psi experiments, the basic principle is to have multiple experimenters involved that (a) have duplicate copies of the randomization and outcome data, (b) check or observe each other, and/or (c) sometimes switch roles. These practices are not time-consuming once they become established habits for experimental design. Experiments involving extensive technology are more difficult to control for experimenter fraud and considerable technical expertise is required among different experimenters. However, methods such as obtaining random numbers from an independent remote site by internet or phone connection could be implemented. Data would be exchanged with the remote site in a way that makes copies of the critical data in two places, and therefore not easily subject to undetectable manipulation by one person. Such processes could be routinely implemented with little overhead after their initial development.

The experimental procedure can also be observed or audited by other experienced researchers and methodologists to enhance confidence in the results. As indicated from the discussion of Levy and Soal above, assuming competent, rational behavior is not always the optimal approach for identifying likely instances and methods of fraud. Sharing raw data is optimal scientific practice. Analyses by others can result in valuable scientific discoveries as well as detecting various types of methodological problems. Funding agencies and journals increasingly require data sharing (American Psychological Association, 2013). The development of online scientific data repositories is rapidly increasing (Marcial & Hemminger, 2010). An internet search for “open science” provides many articles and position papers describing the value of open data sharing. Digital data are conducive to data sharing and the effort is minimal if sharing is anticipated throughout data processing rather than an add-on step at the end.

My observation has been that most parapsychological researchers have provided data when requested by others. However, experimenters have sometimes refused to share data. Sargent’s ganzfeld data discussed above is one case. Skeptic David Marks (2000) describes some remote viewing controversies when data sharing was refused.
When biased data fishing is likely, an original investigator could reasonably require that a recipient register the planned analyses prior to receiving copies of data. The planned analyses would be publicly registered on a study registry and include adjustments for multiple analyses. The original investigator could also require a contract that prevented the recipient from providing the data to others. In general, the high potential for post hoc data fishing must be recognized with open data sharing. One good strategy would be to make data from an initial study available for public exploratory use, but retain data from a confirmatory study as proprietary that would be provided only for confirmatory analyses registered at a study registry.

Given the controversial nature of psi results, I believe that the benefits are worth the effort to design experiments that make experimenter fraud difficult. As noted below, controls for intentional data manipulations also protect against unintentional data changes and are expected in pharmaceutical research.

Conclusions

Experimenter misconduct has occurred many times in parapsychology and is a constant threat. It detracts from the scientific acceptance of the field and hinders progress by diverting resources to invalid hypotheses. The possibility of significant undetected fraud and analysis manipulation in parapsychological research has become increasingly plausible to me over the years. The lack of implementation of effective practices to detect and deter misconduct invites such behavior and makes undetected cases likely.

Experimenter misconduct must be recognized as an appropriate topic of discussion in parapsychology. Such discussions should not be taken personally. This topic is an essential aspect of proper scientific research.

My attitudes toward experimenter misconduct have been influenced by about 20 years of work in medical research. The need for study registration to avoid analysis manipulation is openly recognized. Study registration is legally required for pharmaceutical research (U.S. National Institutes of Health, 2012) and prospective study registration that is publicly accessible is rapidly becoming a required condition for publishing medical research (De Angelis, et al., 2004).

In pharmaceutical research, regulatory agencies audit key sites where data are collected and processed. I managed the software infrastructure for data management and analyses at a company and was the first person the FDA auditor wanted to interview. The auditor asked about every significant step in the development, validation, and use of the software systems and repeatedly asked what steps were taken to verify that intentional or unintentional data alterations did not occur. For example, a laboratory transferred certain data electronically and a programmer imported and reformatted the data. The auditor asked “How do you know the programmer did not change the data?” I explained that the laboratory sent another copy of the data directly
to another person, and a third person compared that copy to the electronic data output by the programmer. Of course, we had documentation for that comparison.

The auditor assumed that intentional or unintentional data alterations by one person should be difficult. Two independent copies of key data are a multiple-experimenter procedure that provides an important level of confidence when research findings are challenged.

Double-checking a colleague’s work is standard procedure in pharmaceutical research. A surprising number of mistakes are discovered. Regulatory auditors expect documentation of this double-checking. These verifications are an established part of the research culture and are not interpreted as questioning a person’s integrity or competence. Undetected intentional data manipulation is very unlikely in this environment.

In addition, the raw data for each study are provided to FDA. Data are collected, managed, and analyzed with the expectation that competent, critical professionals will examine the data in detail. Auditors also verify that the data match the original medical records.

Study registration and multiple experimenter designs have not yet become standard practice in parapsychology. The more informal methodology of academic psychology has been followed rather than the more systematic, convincing methodology of medical research. Given the controversial nature of parapsychological research, the more convincing methodology is appropriate. Fortunately, these methods can be implemented with relatively little additional effort.

The Koestler Parapsychology Unit at the University of Edinburgh now provides a simple public study registry, as well as information about the development of other registries. The registry website is:
https://koestlerunit.wordpress.com/study-registry/.

Parapsychological research organizations and funding sources should require prospective registration of studies, multiple experimenter designs, and data sharing. These requirements for confirmatory experiments would provide the greatest return on investment in research. Parapsychological journals should strongly promote these practices. In addition, a methodologically oriented colleague can be invited to observe or audit an experiment. These practices should also be standard study quality rating factors in meta-analyses.

Research data in parapsychology should be collected, managed, and analyzed with the expectation that the data will have detailed, critical scrutiny by others. The optimal scientific approach is to make all or part of the raw data openly available. However, when biased post hoc analyses are likely, an original investigator may reasonably require that the recipient register the planned analyses publicly, including corrections for multiple analyses, prior to receiving copies of the data.

An appropriate working assumption is that an experimenter who has competently conducted a study and is confident of the results will readily provide the data for
independent analyses. If an experimenter is unwilling to provide the data for independent analyses, an appropriate assumption is that the experiment has questionable methodology and the experimenter has something to hide.

References


---


The author authorizes and grants license that the content of this article ("Experimenter Misconduct in Parapsychology: Data Manipulation and Fraud") may be displayed, reproduced, distributed, and published by anyone without specific permission from the author or compensation to the author, provided that (a) the original author (James E. Kennedy) is identified, (b) the URL or a link to the original document ([http://jeksite.org/psi/misconduct.pdf](http://jeksite.org/psi/misconduct.pdf)) is provided, and (c) the original content of the article is not modified. This authorization applies for any media and applies worldwide for the duration of the copyrights. Uses of this article that do not comply with these stipulations require permission from the author. The author also intends and authorizes that this article becomes public domain upon the death of the author.

---

[Other Methodology Articles](#)