

EXPERIMENTER FRAUD: WHAT ARE APPROPRIATE METHODOLOGICAL STANDARDS?

By J. E. Kennedy

ABSTRACT: Discussions of experimenter fraud in parapsychology have missed a key lesson from the Levy case. The standard procedure for handling scientific fraud is an after-the-fact (post hoc) investigation. Post hoc investigations cannot be expected to be effective in parapsychology because signs of fraud in the data can be attributed to psi, as happened with Levy. In parapsychology, compelling evidence of fraud usually requires direct covert detection of fraud as it occurs during an ongoing experiment, as in the Levy case. However, such covert measures by colleagues are not a practical strategy for addressing fraud and are not expected in other areas of science. The standard that experimental procedures should make fraud by one experimenter very difficult or impossible has long been advocated in parapsychology but has not been implemented in recent decades. This standard was implemented in my experience working in regulated medical research and should eliminate the vast majority of cases of fraud—which start when one experimenter finds data manipulation or fabrication easy and tempting with very little possibility of detection. This standard provides a systematic and effective way to address experimenter fraud and should become part of the new standards for research in the behavioral sciences.

Keywords: experimenter fraud, experimenter misconduct, research methodology, research standards, confirmatory research

The recent extensive discussions of methodological issues for psychological research have included the topic of experimenter fraud (John, Loewenstein, & Prelec, 2012; Simonsohn, 2013; Strobe, Postmes, & Spears, 2012). In extending these discussions to parapsychology, Stokes (2015) and Kennedy (2014) raised the possibility that experimenter fraud may be more extensive than is generally realized and may significantly compromise the research findings, particularly when combined with other forms of methodological bias. Palmer (2016) and Roe (2016) challenged these points.

Virtually all discussions of fraud in parapsychology describe the Levy case as the most definitive evidence of experimenter fraud. As one who was involved in the Levy exposé, my observation is that the various discussions miss the key factors that were involved and the important implications of the Levy case. In addition, writers often make surmises about what happened without direct knowledge. For example, Rogo's (1985) description contains numerous errors and is not a reliable source of information (Kennedy, 2017).

In this paper I discuss key points and implications of the Levy case that have not been discussed before and describe my experiences working in research environments where measures to prevent fraud were standard procedure. I also make recommendations for dealing with fraud based on my experiences combined with the findings from published studies of fraud.

As a frame of reference, it is useful to describe three alternative positions or expectations with regard to experimenter fraud.

1. The expectation of *fraud-proof experiments* is based on the assumption that research can and should be conducted in a way that completely precludes the possibility of experimenter fraud. As Palmer (2016) pointed out, this is the standard for extreme skeptics such as Hansel (1966, 1980). Palmer also pointed out that this expectation is currently widely rejected in science. For me this expectation is eliminated by the fact that no measures could prevent collusion among experimenters to produce fraudulent results.

2. The expectation that *fraud by one experimenter should be very difficult or impossible* is based on the recognition that the vast majority of cases of fraud start when an experimenter is alone with the data and finds that data manipulation or fabrication is easy and tempting with very little possibility of detection. This standard has long been recommended in parapsychology (Akers, 1984; Dalton et al., 1996; Rhine, 1974, 1975) and was a methodological standard in my experience working in regulated medical research (described below).

3. The expectation that *procedures to prevent experimenter fraud are unnecessary* is usually based on the argument that independent replication will reveal and overcome fraud. Palmer (2016) raised a related argument that procedures to prevent experimenter fraud could create a paranoid work environment that should be avoided. He also argued that discussion or modeling of possible fraud for specific studies are implicit accusations of fraud and therefore are not ethical in the absence of strong evidence that fraud has occurred.

Background on Experimenter Fraud

Experimenter fraud is an established factor in scientific research (Broad & Wade, 1982; John et al., 2012; Stroebe et al., 2012; Retraction Watch, n.d.). The extent of occurrence of fraud is unknown because undetected instances are likely and institutions have often been reluctant to make cases of fraud known publicly.

Independent replication and peer review for publication have generally not been effective at detecting even extensive fraud (Broad & Wade, 1982; Strobe et al., 2012) and do not pose a significant risk of detection for those contemplating fraud. The primary symptom of fraud is inconsistent results among experimenters, but such differences are virtually never attributed to fraud.

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 et al., 2012, p. 682). The authors noted that “whistleblowers are likely to remain the single most effective instrument against scientific cheating” (p. 682). This recent analysis confirmed the same basic points made 30 years earlier by Broad and Wade (1982).

The normal process for handling experimenter fraud is an investigation by a committee after suspicions of fraud have been formally raised (Gross, 2016; Strobe et al., 2012). The committee examines publications and asks for raw data and other research records. Evidence or signs of fraud are typically found in the data and publications, including inconsistencies, data patterns that are artifacts of fraud, and/or data that are “too good to be true” (Strobe et al., 2012). Investigations of fraud are expected to take 10 months if all goes smoothly, but in practice, longer times have been common (Gross, 2016).

Gross (2016, p. 700) observed that “there appear to be no systematic empirical studies of the characteristics of perpetrators of scientific misconduct and no good evidence for any common characteristics.” He pointed out that the cases that get extensive publicity usually involve highly ambitious researchers who rise rapidly in elite institutions. However, these highly publicized cases cannot be assumed to be representative of all cases of experimenter fraud.

Three categories of fraud can be distinguished: detected, suspected, and undetected. In cases of detected fraud, initially suspected or observed fraud is investigated and unambiguously resolved as fraud. In cases of suspected fraud, the evidence of fraud is not fully resolved, even though apparent fraud may have been observed by a colleague. Suspected fraud includes cases that are not investigated and remain at the level of rumor as well as cases that are reported and investigated but have inadequate evidence to determine whether fraud did or did not occur. Undetected frauds are cases that do not reach the point of suspicion by colleagues. Reliable data obviously cannot be obtained about undetected fraud.

Surveys have been conducted asking scientists about admitted, observed, or suspected fraud (Fanelli, 2009; John et al., 2012). The accuracy of the findings is questionable for such surveys because the respondents may be biased about this topic. The generalizability of the samples is also questionable. In addition, the surveys cannot address undetected fraud.

However, the surveys may provide insights about the rate at which cases of suspected fraud are reported, investigated, and resolved. In commenting on one of the more methodologically sound surveys, Titus, Wells, and Rhoades (2008) stated “Extrapolating the survey results projects an alarming picture of under-reporting” (p. 981) They argued that all research centers should have the policy that any suspected researcher misconduct must be reported and must be thoroughly and fairly investigated .

My Experiences Exposing Experimenter Fraud

As noted above, I was involved in exposing the fraud of W. J. Levy. The experiment in which Levy was exposed was officially my experiment, and Levy was my co-experimenter. I also had the leading role in investigating the extent of his fraud for three lines of research (Kennedy, 1975a, 1975b, 1975c). Contrary to the incorrect comments by Palmer (2016), Doug Stokes did not have a role in the Levy exposé.

The Levy exposé was different from the typical case of exposed scientific fraud because direct evidence of fraud was obtained as the fraud occurred. Jim Davis, Jerry Levin, and I established a hidden recording of the output of the RNG used in the experiment before the point in the circuit where Levy pulled a plug to introduce bias. Recordings were made during an actual experiment without Levy’s knowledge. Davis also covertly observed the equipment during the period Levy was pulling the plug.

As noted above, suspicions of scientific fraud are normally handled by an investigation after the fact (post hoc) without direct experimental evidence as the fraud occurs. For the 40 cases of fraud summarized by Strobe et al. (2012), only one is described as a “sting operation” in which a colleague trapped a fraudulent researcher, as we did in the Levy case.

Evidence of fraud from post hoc investigations will usually be unconvincing in parapsychology because a fraudulent researcher can claim that the signs of fraud in the data are actually psi effects. The most conspicuous artifacts of Levy’s fraud were long strings of consecutive hits that had extremely low probability of occurring by chance (Kennedy, 1975b). Levy presented these strings as psi effects, and this interpretation had become accepted at the lab. Within the worldview of parapsychology, the claim that signs of fraud are actually psi effects is nearly irrefutable and cannot be resolved by post hoc analysis. For other areas of science, fraudulent researchers do not have psi as a virtually indisputable alternative explanation for signs of fraud in the data.

For the Levy case, we needed to obtain direct, experimental evidence of fraud as it was occurring to establish that the effects were not due to psi. We expected that accusations without such evidence would lead to a subsequent post hoc investigation that would produce a prolonged, intense debate with an inconclusive outcome. Evidence of fraud found in post hoc analysis would be considered to be possible psi effects. In the end, the negative impacts for the accusers would be as great as or greater than for the accused. Before we openly raised the issue of fraud, we needed to have indisputable evidence that some fraud had actually occurred.

Covertly obtaining direct definitive evidence as fraud is being conducted will usually be necessary for resolution of fraud in parapsychology but is generally not a practical goal. The effort to obtain such evidence is beyond what is reasonable in a professional setting. The need to maintain normal interactions with a close colleague while covertly planning and conducting steps for his exposure and resulting ruined career requires a degree of acting and compartmentalization that many scientists do not have. For me it was very difficult. Many pivotal decisions had to be made quickly in secrecy and under stress. In addition to deciding the strategy and technical details for collecting unequivocal evidence, multiple people needed to be involved to establish overwhelming credibility. If it was even remotely feasible, Levy could claim the accusations were false and based on fabricated data. Decisions had to be made about who could handle the acting and extreme secrecy, how they should be approached, the risks of possible compromising communication, and the roles for the various people. In addition, it was sometimes necessary to deceive colleagues in order to keep the preparations secret.

These distasteful steps were necessary to resolve the matter unambiguously rather than creating an irresolvable situation with suspicions but no compelling evidence, as has occurred for other cases of

suspected experimenter fraud in parapsychology. In fact, such a situation had occurred previously with Levy. When Jerry Levin first observed Levy behaving suspiciously near some wires that could be used to manipulate the results, Jerry responded by covering the wires with tape to prevent potential fraud. Jerry did not clearly observe fraud and had only suspicions. However, taping the wires let Levy know that he was suspicious and effectively eliminated the possibility of resolving Levy's fraud in the line of research Jerry was conducting.

I found Jerry's suspicions to be unconvincing and dismissed them—until I later observed Levy apparently manipulating data in another line of research. My observations would have been adequate to initiate an investigation but did not provide the type of indisputable evidence that would be needed to overcome Levy's counterclaims that the accusations were mistaken or fabricated and that the effects were actually due to psi. Carefully planned, indisputable experimental evidence was needed as Levy actually manipulated the data, and multiple people needed to be involved.

Based on my experience exposing fraud, I think it is very unlikely that instances of experimenter fraud in parapsychology will be convincingly resolved. Obtaining convincing evidence of fraud in parapsychology is much more difficult than in other areas of science because the normal process of conducting a post hoc investigation will usually not be effective. Signs of fraud can easily be explained away as psi effects in parapsychology, but not in other areas of science. Another case in which data analyses found patterns that would normally be construed as signs of data manipulation but are ambiguous if PK is considered plausible is described in Kennedy (1980a, 1980b). Compelling evidence during the actual manipulation of the data—a sting operation—is needed to establish that the effects were not due to psi. However, that typically requires covert effort that is not practical for scientific research.

It is usually much easier to avoid dealing with experimenter fraud than to make the effort to fully resolve the matter. Even when clear evidence of fraud is found, the effort to deal with the fraud is very time-consuming and distracts researchers from their main interests. The investigation of the extent of Levy's fraud took about a year, which, as noted above, is common for investigations of fraud. In addition, the adverse effects for the work environment are often significant. For the Levy case, the exposé would clearly create major disruption of the work environment at the beginning of the summer study program that Levy had organized. This disruption would be very detrimental for everyone at the lab, including those exposing Levy. In fact, the initial reaction of one of the three people involved in the exposé was to suggest that a long-term discrete investigation be conducted for several months or longer that would not disrupt the work environment, particularly over the summer. The other person and I vetoed that idea.

Broad and Wade (1982) argued that it is likely that only the most extreme, careless frauds have been detected. That conclusion is consistent with the experience in parapsychology. Levy's fraud appears to have become pervasive and irrational. For one experiment, Levy published fabricated results even though the original data and analysis programs were stored on backup tapes and provided completely different results (Kennedy, 1975a). Those of us involved in the exposé and subsequent investigations did not anticipate such irrational behavior.

For the four major lines of research Levy had conducted, fraud was exposed as it occurred in one and the data for a published study was clearly fabricated in another (Kennedy, 1975a). Strong circumstantial (post hoc) evidence of fraud was found in the other two lines of research (Kennedy, 1975b). Those skeptical of psi will interpret the circumstantial evidence as unequivocal evidence that fraud occurred in those studies. However, those of us involved in the Levy exposé believed that circumstantial evidence alone would not be accepted within parapsychology as compelling evidence that the effects were due to fraud rather than due to psi as claimed by Levy.

My Experiences Preventing Experimenter Fraud

My attitudes toward experimenter fraud have also been influenced by about 20 years of work in medical research. In my experience in regulated medical research, measures to prevent unintentional or intentional (fraudulent) data alterations were an accepted part of the research culture. 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 unintentional or intentional data alterations did not occur.

For example, after learning that 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?” The questions were carefully phrased to include both intentional and unintentional data changes. 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 did not ask about possible errors by the person checking the data or about collusion between the programmer and the person checking the data.

The auditor assumed that intentional or unintentional data alterations by one person should be difficult or impossible. Two independent copies of key data can meet this criterion and provide 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. When the costs of making a mistake are high, people want their work verified.

I found that working in an environment with routine practices to prevent fraud was much preferable to my experiences in parapsychology. In fact, the strategy for exposing Levy involved duplicate records and a colleague observing Levy’s actions during the experiment. These are the same basic procedures that are used to prevent fraud. In research environments with open efforts to prevent fraud (and also prevent unintentional errors), these procedures are expected and are considered good methodology. However, in environments without such measures, undetected fraud can be easy and tempting, and discussion of these practices can be considered inappropriate implicit accusations of fraud or incompetence.

If I would have told the auditor that I considered questions about the integrity of the programmer to be unethical and inappropriate, and that I assumed the programmer did not change the data and believed that efforts to verify that assumption created a bad, paranoid work environment, I would have failed the audit and been fired—appropriately so. Those arguments were not viable in the research culture.

Confronting Experimenter Fraud

The first and most fundamental question is whether the research culture allows the topic of experimenter fraud to be discussed without being taken as personal accusations. More generally, “for the ideologists of science, fraud is taboo, a scandal whose significance must be ritually denied on every occasion” (Broad & Wade, 1982, p. 142). These types of idealistic arguments are no longer viable. “As unpalatable as it is, to complete the culture change initiated in the second half of last century, we have to accept the fact that fraud can happen in our midst and that we have to look out for it” (Stroebe et al., 2012, p. 684).

In a healthy research environment experimenter misconduct (fraud and biased methodological practices) is considered an appropriate and necessary topic of discussion, including about specific studies by specific experimenters. Such discussions reflect a high priority on good methodology and must not be taken as personal accusations.

The next question is: Can we accept that independent replication and peer review generally are not effective at detecting or deterring fraud? Stroebe et al. (2012) and Broad and Wade (1982) reached that conclusion in their studies of fraud, and it has been true for the two prominent cases of fraud in parapsychology (Levy and Soal—see Beloff, 1993, for a description of the Soal case). In fact, the long-established finding in parapsychology of consistent differences among experimenters could be taken as a symptom of experimenter

misconduct (fraud and/or biased methodology). The exclusion of such considerations when discussing these experimenter differences brings into focus how ineffective independent replication is for deterring or detecting fraud. As is clear from Stroebe et al. (2012), replication is at best an extremely inefficient, slow, and costly strategy for dealing with fraud and does not deter fraud. The resources required to conduct well-powered confirmatory studies are usually substantial. The resources for conducting multiple confirmatory studies of a fraudulent finding will often be a significant diversion of the limited resources available for behavioral research. A researcher may initially rationalize fraud as necessary to obtain funding for a more effective research program—which was one of Levy's explanations for his fraud. Unsuccessful independent replications will virtually never be identified as indicating fraud, and thus an experimenter contemplating fraud need not be concerned about the threat of detection.

One of the most important questions is: Can we accept the fact that we simply do not know how much undetected experimenter fraud actually occurs? Cases of detected fraud are a proportion of the total cases of actual fraud and the magnitude of undetected fraud is unknown. The common arguments that detected fraud is rare in parapsychology and occurs at the same or lower rates as in other areas of science (Bierman, Spottiswoode, & Bijl, 2016; Broughton, 1991; Roe, 2016) provide no useful conclusions about the occurrence of undetected fraud in parapsychology or in other areas of science. Bierman et al. (2016) ignored undetected fraud and considered only detected, suspected, and admitted fraud in their simulations—which therefore underestimated the actual occurrence of fraud by an amount that is unknown.

Stroebe et al. (2012, p. 682) commented that the cases of detected fraud in their report “are likely to be the tip of an iceberg of fraudsters.” Titus et al. (2008) made a similar point. Broad and Wade (1982) acknowledged that the actual rate of experimenter fraud is unknown, but thought it is likely that for every case of major detected fraud, “a hundred or so go undetected” (p. 87). These speculations do not provide strong conclusions, but they do indicate the magnitude of the uncertainty.

Stokes (2015) and I think it is likely that a substantial amount of undetected fraud has occurred in parapsychology and in psychology given past research practices. Palmer (2016) and Roe (2016) argued that these are speculations without convincing evidence. However, they also did not provide convincing evidence that substantial undetected fraud has not occurred. Their papers focus on detected fraud and may give readers the impression that they believe the possibility of undetected fraud in past research can be ignored—which was my definite impression from their papers.

If research was conducted with measures to prevent and to detect experimenter fraud, the argument that undetected fraud is negligible would be plausible. However, in the absence of such measures, I do not see a scientific basis for this argument. In personal communication (September 12, 2016), Palmer emphasized that he did not intend to argue that undetected fraud can be ignored. He believes there is not convincing evidence to conclude that substantial undetected fraud has or has not occurred in parapsychology. He considers it possible that Stokes's estimates about fraud may be correct, but those estimates currently must be taken as speculations, not convincing conclusions.

Unfortunately, the uncertainty about the extent of undetected experimenter fraud implies corresponding uncertainty about the validity of the research findings. That was the main point Stokes and I were attempting to make.

My primary purpose in making that point was to bring into focus the need to implement measures to prevent fraud. Both Palmer (personal communication, September 13, 2016) and Roe (2016) indicated that vigilance about the possibility of experimenter fraud is needed and that some measures to address experimenter fraud are appropriate. There appears to be a consensus on this point, although exactly what should be done remains a topic of discussion.

A Methodological Standard for Addressing Experimenter Fraud

Considering all these factors, I believe that the methodological standard of making fraud by one experimenter impossible or very difficult is the optimal practice for research. Experimenter fraud should not be easy and tempting. Implementation of this standard would eliminate the vast majority of

cases of experimenter fraud. I believe that lack of implementation of effective practices to detect and deter experimenter misconduct (fraud and biased methodological practices) invites such behavior and makes undetected cases likely. The research culture in psychology now accepts that methods to prevent questionable research practices are needed. Measures to prevent experimenter fraud should be included in the methodological standards. I consider this standard to be appropriate throughout the behavioral sciences.

As noted above, this standard has long been recommended in parapsychology but has not been implemented in recent decades. Measures to prevent fraud are particularly warranted in parapsychology given the controversial nature of the phenomena, the traditional differences among experimenters in producing effects, and the difficulty in distinguishing between signs of fraud and psi effects. Special experimental designs with extraordinary measures to prevent fraud have also been described (Palmer, 2016; Schmidt, Morris, & Rudolph, 1986; Schmidt & Stapp, 1993); however, these measures are not practical for most research.

My experience has been that it is relatively easy to implement this standard once appropriate research habits have been developed. Measures to prevent fraud are needed for confirmatory research, but are optional for exploratory research by a researcher who plans to conduct one or more confirmatory studies before the findings are published.

Practical recommendations for implementing this standard and implementing other related methodological practices are discussed in Kennedy (2016). One key practice is to make duplicate copies of each component of the data early in the data collection process and handle the copies in a way that would be very difficult or impossible for one experimenter to alter all copies. Ideally, the secure copies will be made before any experimenter has unblinded information that could be used to alter the study results. When that degree of blinding is not possible, two experimenters should be present at any step during the data collection and processing that would allow an experimenter to alter or fabricate data without detection. The experimenters should explicitly and actively have the intention of verifying that intentional or unintentional data alterations do not occur. For automated experiments, documented validation of the software and hardware is needed and if properly done will detect both intentional (fraud) and unintentional problems (Kennedy, 2016).

A healthy, competent research culture will recognize the need to implement such measures as standard procedures. Parapsychological experiments have produced successful results with such measures (Rhine, 1974). Palmer's (2016) concerns about implied accusations of fraud and his speculations about creating a paranoid work environment are not applicable for this type of research culture. All forms of potential research bias should be openly recognized and addressed. The idea that measures to prevent fraud are implicit accusations of fraud is closely associated with the idea that preregistration and other measures to prevent bias are implicit accusations of intentional bias and that measures to prevent unintentional errors are implicit accusations of incompetence. These types of sensitivities do not have a place in a healthy research culture.

Routine measures to prevent fraud are preferable to the enhanced emphasis on after-the-fact accusations and investigations, which Titus et al. (2008) advocated. Their strategy is based more on happenstance than on a systematic approach. It also asks researchers to take actions that will usually create a major burden for the researchers, including investigations that take a year or longer to complete and that may have inconclusive outcomes. I also expect that reliance on accusations and investigations would produce significant discord in a research environment.

To have reasonable hope of competently evaluating suspicions of fraud in parapsychology, an investigating committee must implement covert detection measures during actual experiments, as was done in the Levy case. That is decidedly not an optimal general strategy for addressing fraud. On the other hand, my experience in research environments with routine measures to prevent fraud has been that the issue of fraud is systematically and effectively addressed with negligible discord. Systematic prevention is vastly preferable to after-the-fact accusations.

Making the raw data available to others for independent analyses is also a useful but secondary strategy for deterring and detecting fraud. Data could be fabricated or altered in a way that does not

leave convincing signs of fraud. As noted above, the cases of detected fraud may be the more extreme, careless frauds. Fraudulent researchers who are more careful may not leave conspicuous signs of fraud. Also, accusations of fraud based on post hoc analyses will too often be circumstantial and irresolvable, particularly in parapsychology.

The possibility of incorrect accusations must be recognized and addressed when accusations are based on statistical analyses. The usual concerns about Type I and Type II errors are applicable, and are enhanced for post hoc analyses. In addition, the statistical methods for screening tests are different than for typical experimental research and should be thoroughly understood if an analyst plans to check or screen a number of studies for statistical evidence of fraud. Anyone making accusations of fraud based solely on statistical analysis would be wise to consult an attorney about the legal liabilities and standards of evidence for libel and slander. The adverse consequences for both sides from the inevitable false accusations that will sometimes occur if statistical methods to detect fraud are widely applied reinforce the point that measures to prevent fraud are much preferable to after-the-fact accusations, including the statistical methods described by Simonsohn (2013). Making the data available to others does not eliminate the need for procedures that prevent fraud.

Efforts to address fraud should avoid any assumptions that the motivations for fraud will be simple and identifiable or that the behavior of those committing fraud will be rational and predictable. The two prominent cases of experimenter fraud in parapsychology cannot be understood in terms of straightforward motivations and rational behavior (Kennedy, 2014). As noted above, there is currently no good evidence for personal characteristics that can be used to predict experimenter fraud. Measures to address fraud should be uniformly applied to everyone.

The most defensible alternative that I can see is to argue that parapsychologists should ignore fraud and focus on developing experiments that can be replicated by any competent researcher. This argument is based on the idea that psi will not be widely accepted in science until virtually anyone can demonstrate the phenomenon. I question whether this idea is true. However, it is clear that parapsychology has not yet achieved the goal of highly replicable results. In order to pursue that goal, the field needs to have experimental results that are worth supporting. That requires studies with good methodology. Funding sources should recognize that measures to prevent fraud and other research biases are good investments and should be a requirement for funding. In addition, I think it is likely that psi is only associated with certain people and under certain conditions. If that is true, good methodology will be essential in making progress in parapsychology.

References

- Akers, C. (1984). Methodological criticisms of parapsychology. In S. Krippner & M. L. Carlson (Eds.), *Advances in Parapsychological Research 4* (pp. 112–164). Jefferson, NC: McFarland & Co.
- Bierman, D. J., Spottiswoode, J. P., & Bijl, A. (2016). Testing for questionable research practices in a meta-analysis: An example from experimental parapsychology. *PLoS ONE*, 11(5):e0153049, 1–18. Retrieved from <http://dx.doi.org/10.1371/journal.pone.0153049>
- Beloff, J. (1993). *Parapsychology: A concise history*. New York, NY: St. Martin's.
- Broad, W., & Wade, N. (1982). *Betrayers of the truth*. New York, NY: Simon and Schuster.
- Broughton, R. S. (1991). *Parapsychology: The controversial science*. New York, NY: Ballentine.
- Dalton, K., Delanoy, D., Morris, R. L., Radin, D. I., Taylor, R., & Wiseman, R. (1996). Security measures in an automated ganzfeld system. *Journal of Parapsychology*, 60, 129–147.
- Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE*, 4, 1–11. Retrieved from <http://dx.doi.org/10.1371/journal.pone.0005738>
- Gross, C. (2016). Scientific misconduct. *Annual Review of Psychology*, 67, 693–711
- Hansel, C. E. M. (1966). *ESP: A scientific evaluation*. London, UK: MacGibbon & Key.
- Hansel, C. E. M. (1980). *ESP and parapsychology: A critical re-evaluation*. Buffalo, NY: Prometheus.
- John, L.K., Loewenstein, G., & Prelec, D. (2012). Measuring the prevalence of questionable research practices with incentives for truth telling. *Psychological Science*, 23, 524–532. doi:10.1177/0956797611430953.
- Kennedy, J. E. (1975a). Memo to J.B.R. Re: Maze. Retrieved from <http://jeksite.org/psi/fmaze.pdf>

- Kennedy, J. E. (1975b). Summary of Levy egg work. Retrieved from <http://jeksite.org/psi/fegg.pdf>
- Kennedy, J. E. (1975c). Summary of rat implantation work. Retrieved from <http://jeksite.org/psi/frat.pdf>
- Kennedy, J. E. (1980a). Learning to use ESP: Do the calls match the targets or do the targets match the calls? *Journal of the American Society for Psychical Research*, 74, 191-222. Retrieved from <http://jeksite.org/psi/jaspr80a.pdf>
- Kennedy, J. E. (1980b). Ambiguous data result in ambiguous conclusions: A reply to Charles T. Tart. *Journal of the American Society for Psychical Research*, 74, 349-356. Retrieved from <http://jeksite.org/psi/jaspr80b.pdf>
- Kennedy, J. E. (2014). Experimenter misconduct in parapsychology: Analysis manipulation and fraud. Retrieved from <http://jeksite.org/psi/misconduct.pdf> .
- Kennedy, J. E. (2016). Is the methodological revolution in psychology over or just beginning? *Journal of Parapsychology*, 80, 156-168. Retrieved from http://jeksite.org/psi/methods_predictions.pdf
- Kennedy, J. E. (2017). Notes on a case of scientific fraud in parapsychology. Retrieved from <http://jeksite.org/psi/fraud.htm>
- Palmer, J. (2016). Hansel's ghost: Resurrection of the experimenter fraud hypothesis in parapsychology [Editorial]. *Journal of Parapsychology*, 80, 5-16.
- Retraction Watch (n.d). Retrieved from <http://retractionwatch.com/>
- Rhine, J. B. (1974). Security versus deception in parapsychology. *Journal of Parapsychology*, 38, 99-121.
- Rhine, J. B. (1975). Second report on a case of experimenter fraud. *Journal of Parapsychology*, 39, 306-325.
- Roe, C. A. (2016). The problem of fraud in parapsychology. *Mindfield*, 8(1), 8-17.
- Rogo, D. S. (1985). J. B. Rhine and the Levy scandal. In P. Kurtz (Ed.), *A Skeptic's Handbook of Parapsychology* (pp. 313-326). Buffalo, NY: Prometheus.
- Schmidt, H., Morris, R., & Rudolph, L. (1986). Channeling evidence for a psychokinetic effect to independent observers. *Journal of Parapsychology*, 50, 1-15.
- Schmidt, H., & Stapp, H. (1993). PK with prerecorded random events and the effects of preobservation. *Journal of Parapsychology*, 57, 331-349.
- Simonsohn, U. (2013). Just post it: The lesson from two cases of fabricated data detected by statistics alone. *Psychological Science*, 24 1875-1888. doi: 10.1177/0956797613480366
- Stokes, D. M. (2015). The case against psi. In E. Cardeña, J. Palmer, & D. Marcusson-Clavertz (Eds.), *Parapsychology: A handbook for the 21st Century* (pp. 42-48). Jefferson, NC: McFarland.
- Strobe, W., Postmes, T., & Spears, R. (2012). Scientific misconduct and the myth of self-correction in science. *Perspectives on Psychological Science*, 7, 670-688. Retrieved from <http://pps.sagepub.com/content/7/6/670>.
- Titus, S. L., Wells, J. A., & Rhoades, L. J. (2008). Repairing research integrity. *Nature*, 453, 980-982. Retrieved from http://ori.hhs.gov/sites/default/files/gallup_commentary.pdf

Broomfield, CO, USA
 jek@jeksite.org

Abstracts in Other Languages

French

LA FRAUDE DE L'EXPERIMENTATEUR : QUELS SONT LES STANDARDS METHOD-OLOGIQUES APPROPRIES ?

RESUME : Les discussions sur la fraude de l'expérimentateur en parapsychologie ont manqué de tirer une leçon-clef du cas Lévy. La procédure standard pour gérer la fraude scientifique est une investigation dans l'après-coup (post hoc). On ne peut attendre de ces investigations post hoc qu'elles soient efficaces dans le champ de la parapsychologie car les signes d'une fraude dans les données peuvent être attribués au psi, comme ce fut le cas avec Lévy. En parapsychologie, l'obtention de preuves évidentes d'une fraude nécessite habituellement une détection directe sous couverture de celle-ci au cours de l'expérience, comme dans le cas Lévy. Toutefois, de telles mesures sous couverture réalisées par des collègues ne constituent pas une stratégie très pratique pour traiter la fraude, et elles n'ont pas cours dans les autres domaines de la science. La procédure expérimentale standard pour rendre la fraude par un expérimentateur très

difficile voire impossible a été proposée depuis longtemps en parapsychologie, sans être intégrée au cours des dernières décennies. Elle a toutefois été introduite dans mon expérience de travail dans la recherche médicale réglementée et devrait éliminer la grande majorité des cas de fraude – ceux où l'expérimentateur trouve facile et tentant de manipuler ses données avec très peu de risques d'être détecté. Ce standard fournit un moyen systématique et efficace de s'occuper de la fraude expérimentale et devrait prendre place parmi les nouvelles règles communes de la recherche dans les sciences du comportement.

German

EXPERIMENTATORBETRUG: WAS SIND DIE GEEIGNETEN METHODOLOGISCHEN MASSSTÄBE?

ZUSAMMENFASSUNG: Diskussionen über den Experimentatorbetrug in der Parapsychologie lassen eine Hauptlektion aus dem Fall Levy vermissen. Die Standardvorgehensweise für den Umgang mit wissenschaftlichem Betrug ist eine Untersuchung nach der Tat (post hoc). Man kann nicht erwarten, dass post hoc-Untersuchungen in der Parapsychologie wirksam sind, weil Hinweise auf Betrug in den Daten dem Wirken von Psi zugeschrieben werden können, wie es bei Levy geschah. Ein zwingender Betrugsnachweis in der Parapsychologie erfordert gewöhnlich eine direkte heimliche Aufdeckung des Betrugs, wie sie während eines laufenden Experiments im Falle Levy geschah. Solche heimlichen Maßnahmen seitens von Forscherkollegen stellen jedoch keine praktische Strategie zum Betrugsnachweis dar und sind in anderen Wissenschaftsgebieten nicht üblich. Der Standard, dass durch experimentelle Maßnahmen ein Betrug durch einen Experimentator sehr schwierig oder unmöglich gemacht werden sollte, war lange in der Parapsychologie vertreten, wurde aber in den letzten Jahrzehnten nicht mehr realisiert. Dieser Standard wurde nach meiner Erfahrung in der regulierten medizinischen Forschung eingeführt und sollte die Mehrzahl der Betrugsfälle verhindern – die beginnen, sobald ein Experimentator herausfindet, wie leicht und verlockend die Manipulation oder Fälschung von Daten ist und wie gering die Wahrscheinlichkeit, dabei entdeckt zu werden. Dieser Standard stellt eine systematische und wirksame Art dar, dem Experimentatorbetrug zu begegnen, und sollte Teil der neuen Forschungsstandards in den Verhaltenswissenschaften werden.

Spanish

FRAUDE DEL EXPERIMENTADOR: ¿CUÁLES SON LAS NORMAS METODOLÓGICAS APROPIADAS?

RESUMEN: Las discusiones sobre fraudes en parapsicología perpetrados por el experimentador han hecho caso omiso a una lección clave del caso Levy. El procedimiento estándar para el manejo del fraude científico es una investigación posterior (post hoc). No se puede esperar que las investigaciones post hoc sean efectivas en parapsicología porque los signos de fraude en los datos pueden atribuirse a psi, como sucedió con Levy. Empero, en la parapsicología la evidencia convincente de fraude usualmente requiere la detección encubierta directa de fraude durante un experimento en curso, como en el caso Levy. Sin embargo, estas medidas encubiertas de los colegas no son una estrategia práctica para abordar el fraude y no se practican en otras áreas de la ciencia. La norma de que los procedimientos experimentales deben hacer que el fraude por un experimentador sea muy difícil o imposible se ha propuesto durante mucho tiempo en la parapsicología, pero no se ha aplicado en las últimas décadas. En mi experiencia, esta norma se implementa en la investigación médica regulada y debería eliminar la gran mayoría de los casos de fraude, que comienzan cuando un experimentador encuentra que la manipulación o fabricación de datos es fácil, tentadora, y con muy pocas posibilidades de detección. Tal estándar proporciona una manera sistemática y eficaz de abordar el fraude del experimentador y debe convertirse en parte de las nuevas normas para la investigación en las ciencias del comportamiento.

Copyright notice. This article was originally published in the *Journal of Parapsychology*, 2017, Volume 81, pages 63-72. The *Journal of Parapsychology* holds copyright for the final article. The author retained rights to post the final article on a limited number of websites. This article may be downloaded for personal use and links to the article are allowed, but the article may not be published or reposted on a different website without permission from the *Journal* or from the author.

[Other Methodology Articles](#)