

[Letter on Meta-Analysis in Parapsychology]

(Submitted for publication. Version of 8/5/2007)

To the Editor,

Caroline Watt's (2005) thoughtful and stimulating presidential address to the Parapsychology Association raised several points that could help parapsychology. For example, she pointed out that the proposals in a recent article by me on "A Proposal and Challenge for Proponents and Skeptics of Psi" (Kennedy, 2004a) would enhance the evidential value of meta-analysis. That point is true, but it may be useful to extend the discussion beyond enhancements to meta-analysis.

In particular, effective implementation of the proposals I suggested could virtually eliminate the need for meta-analysis. If appropriate power analyses were incorporated into the design of studies as recommended in my paper, 80% of clearly identified confirmatory or pivotal studies would be expected to be statistically significant, assuming that psi experiments conform to the assumptions for standard statistical research. A few such studies would provide strong evidence for psi without the need for meta-analysis and the associated controversies over alternative methods, criteria, and outcomes.

In the areas of medical research I currently work in, large, well-designed studies are given greater weight than meta-analyses. The situation was well summarized in a recent book on statistical methods in cancer research:

Our inclusion of [meta-analysis] in a chapter on exploratory analyses is an indication of our belief that the importance of meta-analysis lies mainly in exploration, not confirmation. In settling therapeutic issues, a meta-analysis is a poor substitute for one large well-conducted trial. In particular, the expectation that a meta-analysis will be done does not justify designing studies that are too small to detect realistic differences with adequate power. (Green, Benedetti, & Crowley, 2003, p. 231)

Among the medical researchers I have worked with in recent years, the conclusions of a meta-analysis are typically accepted only to the extent that they are supported by statistically significant results from large well-designed studies. In making the transition from parapsychology to other areas of research, the greatest adjustment for me was to start recognizing the fundamental importance of power analysis and the value of large well-designed studies.

To put the matter in concrete terms, I know of no well-designed ganzfeld studies. As noted in the previous article (Kennedy, 2004a), a reasonable power analysis for ganzfeld studies indicates a sample size of at least 192. I know of no ganzfeld studies with a pre-planned sample size of at least 192. There are some cases where a series of studies were combined to reach such a sample size, but these appear to have been done in a post hoc manner, and often combined exploratory studies that had variations in methodology or design. For example, the widely cited Bem and Honorton (1994) article is a meta-analysis of a series of studies from one laboratory. Based on a power analysis from previous research, each of those studies was severely

underpowered and therefore would be considered poorly designed by the standards I work with now. I do not believe that post hoc meta-analysis can compensate for poor designs. In addition, the variability among the studies combined with the negative correlation between effect size and sample size raise doubts about the meta-analysis. A negative correlation between effect size and sample size is normally diagnostic of bias in a meta-analysis (Egger, Smith, Schneider, & Minder, 1997).

Of course, these arguments and proposals assume that psi conforms to the properties of standard statistical research. As noted in the previous article (Kennedy, 2004), I have come to expect that the results of large confirmatory psi studies will not be more reliably significant than the results of small exploratory studies, which is contrary to the basic assumptions of statistical research including meta-analysis. In addition to the references in my previous article, a more recent meta-analysis of PK studies with electronic random number generators found that the z scores (significance level) did not increase with sample size and that effect size was negatively related to sample size (Bosch, Steinkamp, & Boller, 2006). The authors of the meta-analysis proposed that the pattern of results was due to publication bias rather than PK, but admitted that they could not provide convincing evidence for that hypothesis. However, it is noteworthy that a similar pattern of results occurred in the Bem and Honorton (1994) ganzfeld meta-analysis when publication bias presumably could not have been a factor because it included all studies of a certain type from one laboratory.

I suspect that the basically universal disregard in parapsychology for power analysis and for the value of large studies reflects the fact that most psi researchers implicitly (perhaps unconsciously) recognize that psi does not conform to the assumptions for standard statistical research. However, efforts to provide convincing evidence for psi will fail if the experimental results have unexplained properties that are inconsistent with the statistical foundations for the claimed evidence. Cautious scientists will continue to favor methodological problems as the most likely explanation, particularly if the results are unpredictable and appear to be associated with certain experimenters.

The finding that z score does not increase with sample size implies that the standard methods for data analysis including binomial tests, t tests, and analysis of variance do not have their usual meaning and applicability in psi research. The experiment as a whole may be the appropriate unit of analysis rather than the individual trials or subjects as assumed for those tests. The hypothesis of goal-oriented psi experimenter effects is logically consistent with the basic assumptions for psi research and has strong empirical support that the outcomes of psi experiments are typically (but not always) unrelated to sample size (Kennedy, 1995). Appropriate statistical methods for this type of phenomena remain to be developed. Using statistical assumptions that do not fit the phenomena will inevitably result in failure to make scientific progress.

A two-stage statistical strategy may be needed. The first stage would be based on normal statistical methods to provide evidence that something anomalous occurred. The second stage would utilize more novel statistical assumptions appropriate for the phenomena. The concepts of goal-oriented psi (Kennedy, 1995) and evasive psi (Kennedy, 2004b) may be useful starting points for developing relevant methods.

References

- Bem, D.J., & Honorton, C. (1994). Does psi exist? Replicable evidence for an anomalous process of information transfer. *Psychological Bulletin*, 115, 4-18.
- Borsch, H., Steinkamp, F. & Boller, E. (2006). In the eye of the beholder: Reply to Wilson and Shadish (2006) and Radin, Nelson, Dobyns, and Houtkooper (2006). *Psychological Bulletin*, 132, 533-537.
- Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple graphical test. *British Medical Journal*, 315, 629-634.
- Green, S., Benedetti, J., & Crowley, J. (2003). *Clinical Trials in Oncology* (2nd ed.). New York: Chapman & Hall/CRC.
- Kennedy, J.E. (1995). Methods for investigating goal-oriented psi. *Journal of Parapsychology*, 59, 47-62.
- Kennedy, J.E. (2004a). A proposal and challenge for proponents and skeptics of psi. *Journal of Parapsychology*, 68, 157-167.
- Kennedy, J.E. (2004b). What is the purpose of psi? *Journal of the American Society for Psychical Research*, 98, 1-27 (also available at <http://jeksite.org/psi.htm>).
- Watt, C. (2005). 2005 Presidential address: Parapsychology's contribution to psychology: A view from the front line. *Journal of Parapsychology*, 69, 215-232.

J. E. Kennedy
jek@jeksite.org

Return to: [Paranormal Phenomena Articles](#)