

(Original publication: *Journal of Parapsychology*, 2014, Volume 78(2), pages 273-274 by J.E. Kennedy)  
(copyright was not exclusively transferred to the *Journal*)

## [Letter on Methodology for Presentiment Studies]

To the Editor:

The recent article by Dalkvist, Mossbridge, and Westerlund (2014) on expectation bias in presentiment studies discussed an important methodological problem but included a controversial recommendation and two key comments that are not correct. Presentiment studies investigate whether physiological measures indicate that a person can unconsciously and precognitively anticipate a random stimulus. The most common strategy for analyzing the data has been to compare the average values of the observed physiological measures preceding the different types of random stimuli.

This analysis strategy reverses the traditional analysis for a typical ESP experiment, such as a participant pushing a button to predict which light will be randomly selected. The traditional analysis uses the button press or response to predict the random light or target. As described by Burdick and Kelly (1977, p. 93):

The response array is taken as fixed (in fact, it is immaterial where it came from, and this underlies the great generality of the method). The statistical problem is to evaluate the probability of obtaining a number of hits as large or larger than that observed, *given the response array*.

The random targets are the outcomes that are assumed to be variable and to follow a probability distribution.

This strategy was developed in the 1940s when it was realized that the responses cannot be assumed to be independent as is required for the outcome variable for standard statistical analysis. One of the most well-known examples is the stacking effect that occurs when a single target sequence is used with multiple participants. The responses are not considered independent because “the respondents may tend to possess shared guessing habits” (Burdick & Kelly, 1977, p. 92). The basic principle is that habits and associated nonindependence may occur within any sequence of responses generated by a human, and among any such sequences, including when feedback is not given. As noted above, any habits are immaterial when the traditional analysis strategy is used. The random targets are the outcome variable and are statistically independent (if generated with replacement).

The reversed strategy used in presentiment studies treats the random stimuli (targets) as fixed and uses the physiological measures (responses) as the outcome variable. This strategy results in concerns about dependence among responses. Given the immediate feedback on each trial, the analysis can have biases resulting from the physiological measures reflecting the properties of the particular random target sequence for a participant (Dalkvist, Mossbridge, & Westerlund, 2014; Kennedy, 2013).

Expectation bias as discussed by Dalkvist, Mossbridge, and Westerlund is a type of dependence that can occur in cases with immediate feedback. The traditional assumption for ESP research has been that human responses can also have other forms of dependence that the experimenters may not anticipate and that may or may not involve feedback.

One option for presentiment studies is to apply the traditional analysis strategy by using the physiological measures to predict the random stimuli. When properly done, this eliminates problems of dependence. A previous article recommended this option for confirmatory research and discussed the requirements for proper application (Kennedy, 2013).

Another option is to use the physiological measure as the outcome variable and to build statistical models that attempt to handle the potential dependence. However, the nature, amount, and effects of dependence are difficult to establish, and it is difficult to show that a statistical model adequately handles potentially complicated dependencies among responses (Kennedy, 2013). Dalkvist, Mossbridge, and Westerlund recommended that a simple model be used, but they provided little discussion about methods to evaluate whether a model adequately corrects the means and error variance for all pertinent dependencies.

From my perspective, this modeling approach cannot be expected to provide convincing evidence for a controversial effect (Kennedy, 2013). In addition to making this debatable recommendation, Dalkvist, Mossbridge, and Westerlund made two comments that are not correct.

Dalkvist, Mossbridge, and Westerlund inaccurately stated that bootstrap methods are “free from any statistical assumption” (p. 93) and can be applied in cases with dependencies among observations. Contrary to their comment, standard bootstrap methods are based on the assumption that the original observations are independent (Efron & Tibshirani, 1993, pp. 27, 31, 45, 396; Good, 2005, p. 23). Bradley Efron, the initial developer and promoter of bootstrap methods, commented: “There is no easy solution to problems of dependence . . . problems of dependence do not appear to be well understood and are an important area for further research” (Efron & Tibshirani, 1993, p. 396). As implied by this comment, simple models cannot be assumed to solve dependence problems.

Dalkvist, Mossbridge, and Westerlund also inaccurately said that my recommendation to use the physiological measure to predict the random stimuli is flawed by very low power because it does not adjust for the effects of the previous stimulus on the physiological measure. However, adjustments for the previous stimulus can be done with this analysis. The physiological values used in the final analysis for presentiment studies typically are derived by relatively complex processing after all data have been collected. The processing is usually described as baseline adjustment, normalization, data reduction, and/or artifact rejection.

The optimal strategy for confirmatory research is to have all decisions, derivations, adjustments, and criteria for the physiological data for a trial use only data collected prior to feedback for the trial (Kennedy, 2013). Any incorporation of data after feedback introduces potential for bias. When the physiological measures are used to predict the random stimuli, the trials can be stepped through in the sequence that they occurred, with the prediction criteria developed from previous studies or from previous trials in the current experiment. The prediction criteria can include adjustments for the stimulus on the previous trial. Models similar to those discussed by Dalkvist, Mossbridge, and Westerlund may be useful in developing the prediction criteria, but a confirmatory hypothesis test will be based on applying the criteria to other data that were not used in developing the criteria. This strategy avoids contamination by data after feedback for a trial—and also avoids other dependence problems.

The current situation with presentiment research reminds me of free-response research in the late 1970s. Many free-response studies had been done with methodology that was not optimal. The initial discussions of the methodological issues (e.g., Kennedy, 1979a, 1979b) resulted in arguments that some of the methodological concerns were unlikely to have significant effects and/or could be easily counteracted. However, after a few years of debates (e.g., Hyman, 1985), it became apparent to virtually everyone that the controversial methods should be avoided in future research (e.g., Hyman & Honorton, 1986).

## References

- Burdick, D. S., & Kelly, E. F. (1977). Statistical methods in parapsychological research. In B. B. Wolman (Ed.), *Handbook of parapsychology* (pp. 81-130). New York, NY: Van Nostrand Reinhold.
- Dalkvist, J., Mossbridge, J., & Westerlund, J. (2014). How to remove the influence of expectation bias in presentiment and similar experiments: A recommended strategy. *Journal of Parapsychology*, 78, 80-97.
- Good, P. (2005). *Permutation, parametric and bootstrap tests of hypotheses* (3rd ed.). New York, NY: Springer.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.
- Hyman, R. (1985). The ganzfeld psi experiment: A critical appraisal. *Journal of Parapsychology*, 49, 3-49.
- Hyman, R., & Honorton, C. (1986). A joint communiqué: The psi ganzfeld controversy. *Journal of Parapsychology*, 50, 351-364.
- Kennedy, J. E. (1979a). Methodological problems in free-response ESP experiments. *Journal of the American Society for Psychical Research*, 73, 1-15. Retrieved from <http://jeksite.org/psi/jaspr79a.pdf>
- Kennedy, J. E. (1979b). More on methodological issues in free-response psi experiments. *Journal of the American Society for Psychical Research*, 73, 395-401. Retrieved from <http://jeksite.org/psi/jaspr79b.pdf>
- Kennedy, J. E. (2013). Methodology for confirmatory experiments on physiological measures of precognitive anticipation. *Journal of Parapsychology*, 77, 237-248. Retrieved from <http://jeksite.org/psi/jp13b.pdf>

J. E. Kennedy

[jek@jeksite.org](mailto:jek@jeksite.org)