

Comment on “Failing the Future: Three Unsuccessful Attempts to Replicate Bem’s
‘Retroactive Facilitation of Recall’ Effect”

Daryl Bem
Cornell University

I believe that Ritchie, Wiseman, and French have made a competent, good-faith effort to replicate the results of one of my experiments on precognition reported in the article “Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect.” To actively encourage such replications, I make all materials necessary to conduct them to all who are interested in doing so, including Ritchie, Wiseman, and French.

Nevertheless I consider it premature to conclude anything about the replicability of my experiments on the basis of this article. First, in mainstream psychology it usually takes several years before enough attempted replications of a reported effect have accumulated to permit an overall analysis (often called a “meta-analysis”) of the evidence—20 years in the example described below. It usually takes busy researchers several months to find the time to design and run an experiment outside their primary research area, and my article was published only a year ago.

In their article, Ritchie et al. mention that their experiments were “pre-registered.” They are referring to an online registry set up by Wiseman himself, asking anyone planning a replication to pre-register it and then to provide him with the data when the study is completed. As he noted on the registration website: “We will carry out a meta-analysis of all registered studies...that have been completed by 1 December 2011.”

By the deadline, six studies attempting to replicate the Retroactive Recall effect had been completed, including the three failed replications reported by Ritchie et al. and two other replications, both of which successfully reproduced my original findings at statistically significant levels. (One of them was conducted in Italy using Italian words as stimuli.) Even though both successful studies were pre-registered on Wiseman’s registry and their results presumably known to Ritchie et al., they fail to mention them in this article. I consider this an important omission. (I also note that Ritchie et al., describe their

replication attempt as three independent studies, but the total number of sessions they ran was the same as the number I ran in my own original experiment and its successful replication.)

Second, it takes several years and many experiments to figure out exactly which variables in an experiment affect the results. Consider, for example, an attempt to assess the replicability of a well-known effect in mainstream psychology known as the “Mere Exposure Effect,” first brought to the attention of psychologists in 1969: Across a wide range of contexts, the more frequently humans or other animals are exposed to a particular stimulus, the more they come to like it. Twenty years later, a meta-analysis of over 200 mere exposure experiments was published, showing a significant overall effect; it is now widely accepted as a “real” and replicable phenomenon. But that same meta-analysis reveals that the effect fails to replicate on simple stimuli if other, more complex stimuli are presented in the same session. It fails to replicate if too many exposures are used, if the exposure duration is too long, if the interval between exposure and the assessment of liking is too short, or if participants are prone to boredom. As a result, the meta-analysis included many failures to replicate the effect; several of them actually produced results in the direction opposite to prediction. In short, it takes many more than three replication failures to conclude anything about an alleged effect.

I can imagine Ritchie et al’s. refuting this point with the argument that their failed replications were exact replications of my procedures, even using my own computer program to conduct the experiment. But there is one potential difference they fail to discuss: The possible effects of experimenters’ expectations and attitudes about the experimental hypothesis—as demonstrated by Robert Rosenthal in mainstream psychology several years ago. This has also been found to be a source of non-replication in psi (ESP) experiments. Even if the principal investigator is not the actual experimenter, he or she can easily transmit skepticism and negative expectations to the actual experimenters (again, as amply demonstrated in the mainstream studies in which research assistants were led to expect particular outcomes). Ritchie, Wiseman, and French are well known as psi skeptics, whereas I and the investigators of the two successful replications are at least neutral with respect to the existence of psi.

In fact, Wiseman himself has participated in a test of this experimenter effect. In three separate psi experiments, Wiseman and psi-advocate Marilyn Schlitz teamed up to conduct an experiment in which they used the same subject pool, identical procedures, and were randomly assigned to sessions. Schlitz got a significant psi effect in two of the three experiments whereas Wiseman never did.

I believe that such experimenter effects will turn out to be the most important variable in determining the success or failure of future replications of my experiments as well. This does not mean that psi results are unverifiable by independent investigators, but that we must begin to systematically include the experimenters' attributes, attitudes and expectations as variables in the experiments.