A group of 450 students rated the credibility of descriptions of commonly used psychotherapy and placebo conditions. The four commonly used placebo conditions were found to be significantly less believable than the two therapies, raising questions about the adequacy of extant placebos to serve their intended methodological purpose. [The Science Citation Index® (SCI®) and the Social Sciences Citation Index® (SSCI®) indicate that this paper has been cited in over 170 publications.]

Thomas D. Borkovec
Department of Psychology
Pennsylvania State University
University Park, PA 16802

June 3, 1986

When I began graduate studies in 1966, experimental research on psychotherapy was in its infancy. In that same year, one of my mentors, Gordon Paul, published a classic study comparing the efficacy of systematic desensitization and insight therapies.1 His study established a standard in the field for methodological rigor. Paul taught us many different things; commitment to excellence, scientific integrity, and constant critical analysis leading to further improvements in research methodology were among the ones for which I have been most grateful.

In this context, my classmates and I spent hours debating various methodology issues. As we were taught, the demonstration of the effectiveness of a therapy technique required a comparison to nontreated and placebo groups. The purpose of the placebo was partly to rule out client expectancy as a rival interpretation of outcome effects. One of us (I don’t remember whether it was me) noted that commonly used placebos did not seem very believable as treatments. If the placebo given was not believable to the client in comparison with the therapy condition, then that placebo was not accomplishing its intended purpose as a control.

A few years later at the University of Iowa, Sid Nau and I decided to investigate the problem because the issue seemed of fundamental methodological importance. The investigation was very simple. We asked students to read descriptions of "new therapies for speech anxiety" and to rate each on a 10-point scale, graded in terms of believability of the method as a therapy procedure. We found that the two real therapies used were significantly more credible than four commonly used placebos. So placebo effects were perhaps not being adequately ruled out in most existing outcome studies.

We recommended that care be taken in constructing credible placebos and that credibility ratings be obtained on every condition in any outcome study to verify equivalent believability. Our arguments in the abstract seemed persuasive from a methodological perspective, and for that reason, many of us routinely began to administer the credibility scales in our studies.

Our original scale was the first to be published, so our article remains the most convenient citation for authors using some form of the scales. Although the logic for their use seemed compelling, popularization of the argument, and thus the widespread use of the scales, awaited a review of its possible implications by Kazdin and Wilcoxon wherein they argued that the efficacy evidenced for systematic desensitization was perhaps flawed by a general failure in that literature to control for credibility differences among conditions.2

As Paul has recently suggested,3 the notion of credibility has been at times inaccurately overgeneralized and misused. My greatest concern is that none of us have yet demonstrated that variation in credibility does in fact relate to variation in outcome. Our original argument remains logically but not empirically justified. Perhaps a future author of a Citation Classic will someday resolve this issue.