

18

BE AS CAREFUL AFTER
YOUR STUDY IS RUN
AS YOU ARE BEFORE

Richard E. Petty
Ohio State University

A critical aspect of engaging in scientific work is being extremely careful. Whether you are mixing chemicals or psychological ingredients, you want to do it with the utmost diligence so you can provide the strongest test of your hypothesis. Hours and hours are spent *prior* to the implementation of most studies in an attempt to get the materials (independent and dependent variables) just right so that they represent the constructs of interest. Against this background, the research mistake I describe has to do with the importance of being just as careful *after* the data are collected. The mistake I describe is one that stands out in my mind because (1) it came very early in my research career (my second year as an assistant professor), and (2) it taught me a very valuable lesson.

As was my common practice then and now, the research was done in collaboration with a graduate student. Our goal was to test a new theoretical framework against the prevailing view regarding the role of personal relevance in persuasion. Prior research indicated that people tended to reject counter-attitudinal messages more when the message addressed a topic of high rather than low personal relevance. As you might imagine, as the topic becomes more personally important, people can become more defensive in guarding their initial views.

In contrast, our theory said that personal relevance makes people more interested in thinking about the information presented. Specifically, the extra thinking from the personal connection to the proposal can result in more message rejection when the message is easy to argue against. But, we argued, if the message presented very cogent arguments, the extra thinking could lead to greater acceptance. Thus, our theory predicted an interaction between personal relevance and argument quality. If the arguments were weak, high relevance would lead to greater rejection—the typical effect that was found previously. However, if the arguments were strong, more thinking under high relevance would lead to more message acceptance—the opposite of what was found previously.

To examine our hypothesis, we presented undergraduates with a counter-attitudinal proposal on a novel topic—that they should be required to take a comprehensive exam in their major in order to graduate. The students did not like this proposal at all. Then, we varied two things: (1) the quality of the arguments in the message and (2) the personal relevance of the proposal. To vary argument quality,

we pretested many arguments and selected those that elicited mostly favorable thoughts as the strong ones. The weak arguments were selected so that they elicited mostly negative thoughts (counterarguments). To vary relevance, we told students in the high relevance condition that the exams were being proposed for next year, in which case all of them would be affected. In the low-relevance condition, they were told that the exams were being proposed for 10 years in the future so none of them would be affected.

We then made four tape recordings for participants to hear that represented each of the four conditions in our 2 (Argument quality: weak vs. strong) X 2 (Personal relevance: 1 vs. 10 years) experimental design. We chose an audio message rather than a written one to ensure that all participants would have some exposure to the arguments. After listening to the appropriate message over headphones, students wrote their thoughts about the message. Then, attitudes toward the proposal were assessed (e.g., rating of how good/bad it was on a 7-point scale).

The study seemed to run smoothly enough, and within a few weeks, all of the data were collected. Now it was time to look at our results, and I scheduled a meeting with the graduate student in charge. At the meeting, we went over the data for each cell of the design. First, we had replicated the standard personal relevance effect when weak arguments were presented—the students were less favorable when personal relevance was high than low. So far, so good. Then, I was stunned when I saw the strong arguments results. Here, high personal relevance produced more *rejection* of the advocacy than in any other condition. This was quite different from what we hypothesized, so we next eagerly read the students' thoughts from that condition.

The participants' thoughts in the strong arguments–high relevance condition were not at all what I expected. Rather than providing cogent counterarguments to the strong arguments, the thoughts were statements like, "I couldn't wait for this damn study to end;" "This experiment is torture;" and "That tape was so annoying." A look at thoughts in the other conditions revealed that they were just relevant reactions to the arguments presented. What had gone wrong? I asked the graduate student to play each of the tapes for me over headphones just as the participants had heard them. Three of the tapes were fine. The tape for the high relevance–strong argument condition, however, had a loud, high-pitched screeching sound every 15 seconds or so. Indeed, it was like torture to listen to it! To this day, I have no idea how this screeching was introduced, but it clearly explained why participants' attitudes in this condition were so negative.

So, what was my mistake? First, I had not observed each of the instantiations of the experimental conditions from the participants' point of view prior to running the study—nor had my graduate student. But, second and more important, I only scrutinized the study materials extra carefully after the study was completed because the data had not come out as hypothesized. What would have happened, I wondered, if the screeching tape happened to be in the high relevance–weak argument condition?

That would have helped our study come out as hypothesized—but due to an artifact in the experimental materials. I realized it would be unfortunate if extra checking did not take place just as diligently if the study had confirmed our hypothesis.

Prior scholars of research methods have written about how researchers can be biased in favor of their preferred outcome. True enough. But there is nothing like a personal experience to bring the lesson home. That is, if a study comes out “correctly,” researchers are less likely to double-check experimental conditions, or redo data coding, or engage in additional data analysis, or look for outliers, or read participants’ written responses for indications of suspicion. They are less likely to do all of the things that they might try if the data did *not* come out as predicted. The lesson, of course, is that all of these double-checking protocols should be in place regardless of how the data come out—in support or not in support of one’s theorizing. This is something that we now aim to do with all of the studies we conduct. And, by the way, we reran the screeching study getting rid of the problem, and it was eventually published (Petty & Cacioppo, 1979).

REFERENCE

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology*, *37*, 1915–1926.

CRITICAL THINKING QUESTIONS

1. Why is it just as important to be careful after as before conducting an experiment?
2. What are some of the reasons that bias can enter into the scientific process after data are collected?
3. How can you avoid confirmatory bias when dealing with the results from your own studies?