

The Importance of Credibility in Research Design, Implementation, and Evaluation

Benjamin Burke, MS
November 2018

In 1971, Philip Zimbardo and his team conducted the Stanford Prison Experiment to see what happens when ordinary people are put in an extraordinary situation. Eighteen college-aged men were randomly assigned a role: *prison guard* or *inmate*. The study examined the influence of the presumably evil prison environment, testing if simply being in the environment would turn typical people into harsh, power-abusing people. Shortly, the “guards” became cruel and even abusive to the “prisoners” to the extent that the study was shut down after six days (11 days early) due to the chaotic and traumatic nature of the experiment. The study is regarded as one of the most infamous scientific experiments in modern history. Despite being partially responsible for the establishment of ethical considerations in human sciences study, it is also a sterling example for why credibility should be evaluated when interpreting research.

Credibility refers to how much a person can trust the findings of a research study, based on how carefully it was designed, implemented, and evaluated. A variety of areas should be considered when determining credibility, usually with emphasis on the study’s methodology. Methodology includes factors, such as the appropriateness of the study design, the representativeness of the sample, and the analytic approach. Credibility is important to evaluate because poor methodology can lead to poor results, which influence real life implications. Consider the Stanford Prison Experiment as an example.

First, consider the recruitment process. The experiment recruited participants using a newspaper advertisement to request subjects for a “psychological study of prison life.” However harmless as this may sound, this phrasing could have attracted certain people to the study. In 2007, a study was conducted on sample selection that made two similar newspaper advertisements: one included information about being a prison study and the other did not. The ad that included the prison information yielded a sample of people with higher levels of aggressive and socially domineering personality traits than the ad without prison information. This suggests that Zimbardo’s sample may have been more aggressive than the average

individual, potentially explaining the hostile behaviors observed in the experiment.

Second, Zimbardo was a biased study participant, instead of an objective investigator. He posed as the prison superintendent and created an environment where the prisoners felt powerless and humiliated. The study team coached the participants and described the prison environment as “evil,” thus, calling into question the results that emerged. Imagine you are trying to measure taste preferences for a soda. Your test subjects try the drink, and then are only asked to list the things they didn’t like about the drink. The only data you will gather will be about people’s negative reactions because of the biased nature of the study’s design. In a similar way, the prison experiment was designed to produce abuses of power, and the results demonstrate that finding. Hopefully, at this point the argument for evaluating credibility is becoming clear.

Because of the methodological issues of the study, the findings were skewed and lacked trustworthiness. Unfortunately, they were applied to influence real life implications. Shortly after the study, the results were used to influence Congressional prison reform policy and had an impact on the national narrative of prisons and human reality as a whole. The effects have been far reaching and all based on biased, highly questionable findings. Research is regularly utilized to inform local and national policy, as well as to inform practice; however, it can also be an illustration for the old saying: With great power, comes great responsibility (phrase commonly attributed to both FDR and Spider-Man). Researchers have the responsibility to produce sound science, and careful evaluation of research is necessary to ensure that findings are trustworthy. Without such rigor, the mistake could be costly for decades to come.

Military REACH regularly summarizes and evaluates newly published research in, what we call, TRIP reports. Credibility is a key dimension of evaluation. See our library to understand how we measure credibility, and stay tuned in the coming months to learn about our other dimensions of evaluation - *contributory* and *communicative*.