

Threats to Validity

Find this and many other dissertation guides and resources at www.dissertationrecipes.com

If you plan to do or review an experimental study, then you need to be aware of the following threats to validity and a means to reduce bias. People tend to have an unquestioning and accepting view of statistics and give it more credibility than it deserves. A good consumer and producer of statistics is cognoscente of all of the following!

The Hawthorne Effect: A tendency of participants of research to change their behavior simply because they are being studied. So called because the classic study in which this behavior was discovered was in the Hawthorne Western Electric Company Plant in Illinois. In this study, workers improved their output regardless of changes in their working condition.

John Henry Effect: A tendency of persons in a control group to take the experimental situation as a challenge and exert more effort than they otherwise would; they try to beat the experimental group. This negates the whole purpose of a control group. So called because this was discovered at the John Henry Company where a new power tool was being tested to see if it could improve productivity. The workers using the old tool took it as a challenge to work harder to show they were just as good and should get the new tool.

Rosenthal or Pygmalion Effect: Changes in participants' behaviors brought about by researcher expectations; this is known as a self-fulfilling prophecy. The term originally comes from Greek Mythology and was popularized by G.B. Shaw. Best know in controversial study by Rosenthal and Jackson in which teachers were told to expect some of their students' intelligence test scores to increase. They did increase based solely on the teachers' expectations and perceptions.

Double Blinded procedure: A means of reducing bias in an experiment by ensuring that both those who administer a treatment and those who receive it do not know (are blinded to) which study participants are in the control and experimental groups.

Response Shift Bias: Howard defined the response-shift bias as "a treatment produced change in a subject's awareness or understanding of the variable being measured" (1982, p. 320). For example, participants in diversity training workshops might view themselves as free of racial stereotypes prior to the workshop. During the workshop, however, they may become acutely aware of prejudicial thoughts that had been out-of-awareness. Thus, participants' preratings might be high (i.e., very nonprejudiced), whereas postratings might be low (i.e., very prejudiced). Such ratings would mask the effects of the workshop. Because the two ratings in the then-post design are "made in close proximity, it is more likely that both ratings will be made from the same perspective and thus be free of response-shift bias" (Howard, 1982, p. 320).

The novelty effect is the tendency for performance to initially improve when a new treatment or technology is instituted, not because of any actual improvement, but in response to increased interest in the new treatment or technology.

Sensitization (from the pretest or posttest), Pretest sensitization (pretest sets the stage); A treatment might only work if a pretest is given. Because they have taken a pretest, the participants may be more sensitive to the treatment. Had they not taken a pretest, the treatment would not have worked. Posttest sensitization (posttest helps treatment "fall into place"); the posttest can become a learning experience. "For example, the posttest might cause certain ideas presented during the treatment to 'fall into place' If the participants had not taken a posttest, the treatment would not have worked.

Inadequate power: The power of a statistical test is the probability that the test will reject a false null hypothesis (that it will not make a Type II error). As power increases, the chances of a Type II error decrease. The probability of a Type II error is referred to as the false negative rate (β). Therefore power is equal to $1 - \beta$. Power analysis can either be done before (a priori) or after (post hoc) data are collected. A priori power analysis is conducted prior to the research study, and is typically used to determine an appropriate sample size to achieve adequate power. Post-hoc power analysis is conducted after a study has been completed, and uses the obtained sample size and effect size to determine what the power was in the study, assuming the effect size in the sample is equal to the effect size in the population.

Treatment diffusion: occurs when a comparison group learns about the program from program participants. The comparison group might then set up their own program by imitating that of the program group. This threat to internal validity will equalize the outcome between the groups. In this case it will be harder to tell if your program under study actually works.

Resentful demoralization: This threat may occur in intervention studies in which comparison groups not obtaining a desirable treatment become discouraged or retaliatory and, as a result, perform worse on the outcome measures.

Compensatory rivalry: occurs when the comparison group knows what the program group is getting and therefore, develops a competitive attitude with them and makes it harder to see an effect if one is there.

Mortality threat occurs when participants dropout of the study. If more patients in Washington, D.C. drop out of a program than patients in the Chula Vista, CA, then the observed differences between groups becomes questionable.

Compensatory equalization of Treatment: occurs when people in a control group demand to be reassigned to the other group, the group that is receiving the program or treatment.

Response shift: a treatment produced change in a participant's awareness or understanding of the variable being measured.

The IV (Independent Variable) the predictor or variable that can be manipulated and the DV (Dependent Variable) the criterion variable, the effect of the IV *may be mis-reported*. Does smoking causes lower grades? Perhaps kids who get lower grades take solace by smoking.

