Threats to Validity

Maturation Effect: The observed effect is due to normal developmental processes or changes over time, not the program.

- Program A (implemented during the first semester of college) claims to have increased students' sense of independence. However, studies show students naturally gain more independence during their first semester of college even without an intervention.

History Effect: The observed effect is not due to the program, but to some other unaccounted for event.

- Program B claims to have reduced instances of sexual assault on campus. However, sexual assault prevention is a university-wide initiative and upon further investigation, the facilitators of Program B realize their participants also received programming related to sexual assault prevention in their residence halls. Could the reduction in instances of sexual assault be due to this residence life program instead?

Selection Bias: The observed difference between two groups at posttest is not due to the program, but to preexisting differences between the groups.

- Facilitators of Program C compare students who participated in their service learning program to students who did not and are pleased to find that their students are higher in civic engagement—clear evidence that the program works! Upon further investigation, however, they discover that students high in civic engagement were more likely to participate in their program in the first place. Thus, the difference between the groups was due to self-selection into the program, not the program's effectiveness.

Attrition: The observed effect may be biased due to a substantial amount of missing data (i.e., students failing to complete the program or take the posttest).

- Organization D finds that students' sense of belonging to their organization increased drastically from pretest to posttest—a major success! Upon further investigation, however, it becomes clear that students who felt lower sense of belonging dropped out of the organization and, thus, did not take the posttest. As such, the posttest results were artificially inflated.

Instrumentation Effect: The observed effect is due to changes in the instrument (or interpretation of scores), not the program.

- Program E recruits several raters to review ethical reasoning essays before and after a month-long ethical reasoning program. Shockingly, it seems students performed *worse* after the intervention. Upon further investigation, however, it becomes clear that the problem was with the raters. They became more critical over time, thus evaluating the posttest essays more harshly than the pretest essays.

Response Processes: Results cannot be trusted to reflect students' true ability because they are impacted by things like socially desirable responding and low motivation.

- After completing a 6-hour alcohol prevention workshop, students are fatigued and ready to leave. Unsurprisingly, when asked to complete a 100-item posttest (the only thing separating them from freedom) they speed through the test, responding randomly to the questions. Subsequent posttest results show students gained nothing from the workshop. Should these results be trusted?

For more information, consult:

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. New York, NY: Houghton Mifflin Company.

	Threats to Validity					
Data Collection Designs	Maturation Effect	History Effect	Selection Bias	Attrition	Instrumentation Effect	Response Processes
Posttest Only Design	X	X	N/A	N/A	N/A	X
Pretest-Posttest Design	X	X	N/A	X	X	X
Comparison Group Design (No Pretest)	?	?	X	X	N/A	?
Pretest-Posttest w/ Comparison Group Design	✓/?	✓/?	✓/?	X	✓	?
Pretest-Posttest w/ Comparison Group & Random Assignment Design	✓	✓	✓	X	✓	?

X = The design is highly susceptible to this threat.
? = This threat is a *possible* source of concern for the design. The design *may* provide partial protection against this threat.
✓ = The design provides strong protection against this threat.
N/A = The threat is not applicable for this design.

Table adapted from:

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. *Handbook of research on teaching. Chicago, IL: Rand McNally*.