

## TIP SHEET: BIAS IN PROGRAM EVALUATION

In statistics, bias is defined as “a systematic distortion of a statistical result due to a factor not allowed for in its derivation” (from *Oxford Dictionaries*, [http://www.oxforddictionaries.com/us/definition/american\\_english/bias](http://www.oxforddictionaries.com/us/definition/american_english/bias)). There are several types of bias – statistical and non-statistical – that are particularly relevant to the evaluation of social and educational programs. These are summarized in the table below:

TYPE OF BIAS	WHAT IT IS	WHY IT MATTERS
<b>Cultural Bias</b>	The bias that we as humans bring to our interpretation of the world as a result of the people, places, and things in our lives (objectivity is a myth!).	Failure to acknowledge one’s cultural biases as an evaluator can be harmful to stakeholders when working across differences of race, class, gender, sexuality, ability, etc.
<b>Confirmation Bias</b>	When we seek out and/or interpret information in a way that confirms and/or reinforces what we already believe.	We may make decisions based on inaccurate or incomplete information or interpretations of information.
<b>Selection Bias</b>	Occurs when participants in a program differ in a systematic way from non-participants (e.g. youth with leadership skills are more likely to choose a leadership program in the first place).	If we compare participants to non-participants, and fail to account for selection bias, we may overestimate the impact of our programs.
<b>Maturation Bias</b>	Occurs when phenomena that would have happened naturally are attributed to a program or intervention (e.g. children naturally develop self-management skills as they get older)	If we look at changes over time, and fail to account for possible maturation effects, we may overestimate (or underestimate!) the impact of our programs.
<b>Attrition Bias</b>	Occurs when participants drop out of a program before an evaluation is complete.	In a pre-/post- or post- only design, we may fail to account for the effects of a program on participants who have dropped out.
<b>Instrumentation Bias</b>	Occurs when the tools or techniques we use to measure are not accurate or consistent.	The conditions under which kids are assessed can influence the results of the assessment. Evaluation practices (assessor behaviors, survey instructions) should be as consistent as possible.
<b>Social Desirability Bias</b>	Occurs when survey respondents answer questions in a way they think is “right” or socially desirable rather than a way that is accurate or true.	We may not be measuring what we think we are measuring.
<b>Response Shift Bias</b>	Occurs when a program or intervention affects the standard or criteria that participants use to assess themselves (“you don’t know what you don’t know”).	Can produce untrue or counterintuitive results in pre- / post- research designs.

Bias can arise from sampling practices, research design, or from simple cognitive factors affecting all humans. In designing program evaluations, we should do our best to minimize bias. However, some degree of bias is *inevitable*, so it is important to acknowledge it when designing evaluation strategy, looking at evaluation results, and making claims about program impact.