Threat to internal validity	Relationship to evaluation design
History: An unanticipated event occurring during intervention implementation that is not measured or accounted for in the analysis.	Evaluation designs with more than one group (e.g., control group, comparison group) minimize the history threat to internal validity because the exposure to unanticipated events is relevant to both groups and differences between the groups can be attributed to the intervention.
Maturation: The outcomes observed may be attributed to normal developmental processes in the population or subpopulations of interest as a function of time.	Evaluation designs with more than one group (e.g., control group, comparison group) minimize the maturation threat to internal validity because the developmental processes are relevant to both groups and differences between the groups can be attributed to the intervention.
Statistical regression: A tendency of people who have extreme scores on measures (e.g., surveys, direct observation) to regress to the mean on subsequent administration of those measures.	Evaluation designs with more than one group (e.g., control group, comparison group) minimize the statistical regression threat to internal validity because the tendency to regress to the mean is relevant to both groups and differences between the groups can be attributed to the intervention.
Selection: The population or subpopulation members assigned to intervention and control or comparison groups should be equivalent before the intervention is implemented.	Evaluation designs with only one group (i.e., intervention group) or those with random assignment of individuals to group (i.e., intervention group and control or comparison group) eliminate or minimize the selection threat to internal validity because single group or the random assignment to group eliminates the second group or limits differences between the groups.
Experimental mortality: The differential loss of population or subpopulation members assigned to intervention versus control or comparison groups.	Evaluation designs with random assignment of individuals to group (i.e., intervention group and control or comparison group) minimize the experimental mortality threat to internal validity because random assignment increases the equivalence of groups, thus differences between the groups can be attributed to the intervention.
Testing: The baseline assessment (pre-intervention) may sensitize population or subpopulation members in unanticipated ways and affect their performance on the follow-up assessment(s).	Evaluation designs with more than one group (e.g., control group, comparison group) minimize the testing threat to internal validity because exposure to the baseline assessment is relevant to both groups and differences between the groups can be attributed to the intervention.
Instrumentation: The measures changed from the baseline assessment to the follow-up assessment.	Evaluation designs with more than one group (e.g., control group, comparison group) minimize the instrumentation threat to internal validity because the changes in measures are relevant to both groups and differences between the groups can be attributed to the intervention.