**Threats to Validity: A Primer**

|  |  |
| --- | --- |
| Threats to Construct Validity | |
| Type | **Explanation** |
| Inadequate explication of constructs | Failure to adequately explicate a construct may lead to incorrect inferences about the relationship between operation and construct. |
| Construct confounding | Operations usually involve more than one construct, and failure to describe all constructs may result in incomplete construct inferences. |
| Mono-operation bias | Any one operationalization of a construct both underrepresents the construct of interest and measure irrelevant constructs, complicating inferences. |
| Mono-method bias | When all operationalizations use the same method (e.g., self-report), that method is part of the construct actually studied. |
| Confounding construct with levels of constructs | Inferences about the constructs that best represent study operations may fail to describe the limited levels of the construct studied. |
| Treatment sensitive factorial structure | The structure of a measure may change as a result of treatment, change that may be hidden if the same scoring is always used. |
| Reactive self-report changes | Self-reports can be affected by participants motivation to be in a treatment condition, motivation that can change after assignment has been made. |
| Reactivity to experimental situation | Participant responses reflect not just treatments and measures but also participants’ perceptions of the experimental situation, and those perceptions are actually part of the treatment construct. |
| Experimenter expectancies | The experimenter can influence participant responses by conveying expectations about desirable responses, and those responses are part of the treatment construct. |
| Novelty and disruption effects | Participants may respond unusually well to a novel innovation or unusually poorly to one that disrupts their routine, a response that must then be included as part of the treatment construct definition. |
| Compensatory equalization | When treatment provides desirable goods or services, administrators, staff, or constituents may provide compensatory goods or services to those not receiving treatment, and this action must be included as part of the treatment construct description. |
| Compensatory rivalry | Participants not receiving treatment may be motivated to show they can do as well as those receiving treatment, and this must be included as part of the treatment construct. |
| Resentful demoralization | Participants not receiving a desirable treatment may be so resentful or demoralized that they respond more negatively than otherwise, and this must be included as part of the treatment construct. |
| Treatment diffusion | Participants may receive services from a condition to which they were not assigned, making construct definitions of both conditions difficult. |

|  |  |
| --- | --- |
| Threats to External Validity | |
| Type | **Explanation** |
| Interaction of the causal relationship with units | An effect found when certain kinds of units might not hold if other types of units had been studied. |
| Interaction of the causal relationship over treatment variations | An effect found with one treatment variation might not hold with other variations of the treatment, or when that treatment is combined with other treatments, or when only part of a treatment is used. |
| Interaction of the causal relationship with outcomes | An effect found on one kind of outcome observation may not hold if other outcome observations were used. |
| Interaction of the causal relationship with settings | An effect found in one kind of setting may not holds in other settings. |
| Context-dependent mediation | An explanatory mediator of a causal relationship in one context may not mediate in another. |

|  |  |
| --- | --- |
| Threats to Construct Validity | |
| Type | **Explanation** |
| Ambiguous temporal precedence | Lack of clarity about which variable occurred first may yield confusion about which variable is the cause and which is the effect. |
| Selection | Systematic differences over conditions in respondent characteristics that could also cause the observed effect. |
| History | Events occurring concurrently with treatment that could cause the observed effect. |
| Maturation | Naturally occurring changes over time that could be confused with a treatment effect. |
| Regression | When units are selected for their extreme scores, they will often have less extreme scores on other variables, an occurrence that can be confused with a treatment effect. |
| Attrition | Loss of respondents to treatment or measurement can produce artifactual effects if that loss is systematically correlated with conditions. |
| Testing | Exposure to a test can affect test scores on subsequent exposures to that test, an occurrence that can be confused with a treatment effect. |
| Instrumentation | The nature of a measure may change over time or conditions in a way that could be confused with a treatment effect. |
| Additive and interactive threats | The impact of a threat can be added to that of another threat or may depend on the level of another threat. |

|  |  |
| --- | --- |
| Threats to STATISTICAL Validity | |
| Type | **Explanation** |
| Low statistical power | An insufficiently powered experiment may incorrectly conclude that the relationship between cause and effect is not statistically significant. |
| Violated assumptions of statistical test | Violations of statistical test assumptions can lead to either overestimating or underestimating the size and significance of an effect. |
| Fishing and the error rate problem | Repeated tests for significant relationships, if uncorrected for the number of tests, can artificially inflate statistical significance. |
| Unreliability of measures | Measurement error weakens the relationship between two variables. |
| Restriction of range | Reduced range on a variable usually weakens the relationship between it and another variable. |
| Unreliability of treatment implementation | If a treatment is intended to implement in a standardized manner is implemented only partially for some respondents, effects may be underestimated. |
| Extraneous variance in experimental setting | Some features of an experimental setting may inflate error, making detection of an effect more difficult. |
| Heterogeneity of units | Increased variability on the outcome variable within conditions increases error variance, making detection of a relationship more difficult. |
| Inaccurate effect size estimation | Some statistics systematically overestimate or underestimate the size of an effect. |

[A close up of a sign

Description automatically generated](https://creativecommons.org/licenses/by-nc-sa/4.0/)

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](http://creativecommons.org/licenses/by-nc-sa/4.0/). The document has been created with colorblind safe colors.