

Six pitfalls, at a glance

For more information, see:

<http://www.eif.org.uk/publication/evaluating-early-intervention-programmes-six-common-pitfalls-and-how-to-avoid-them>

1 No robust comparison group



Problem: A robust comparison group is essential for concluding whether participation in a programme has caused improvements in outcomes. However, some studies do not use a comparison group at all; others use a comparison group which is not sufficiently robust, biasing the results.

Solution: Evaluators should endeavor to use a comparison group in impact evaluations. Ideally this should be generated by random assignment (as in a randomised control trial, or RCT), or through a sufficiently rigorous quasi-experimental method (as in a quasi-experimental design studies, or QED).

2 High drop-out rate



Problem: Attrition – the loss of participants during an evaluation – can introduce two problems: the study sample may become less representative of the target population, and the intervention group and control group may become less similar. These biases can result in misleading conclusions regarding programme effectiveness or the applicability of findings to the target population.

Solution: There are a range of measures to improve participants' cooperation with data collection, such as financial compensation. In addition, researchers can conduct analyses to verify the extent to which attrition has introduced bias and report any potential effects on the results.

3 Excluding participants from the analysis



Problem: Excluding participants from data collection and analysis due to low participation in the programme risks undermining the equivalence of the intervention and control groups, and so biasing the results. Bias can also arise from excluding control group participants who receive some or all of the programme that is being evaluated.

Solution: Evaluators should attempt to collect outcome data on all participants and include them in the final analysis of outcomes, regardless of how much of the programme was received. This maintains greater similarity between the intervention and control group, and so is less likely to produce bias.

4 Using inappropriate measures



Problem: Using measures which have not demonstrated validity and reliability limits our confidence in an evaluation's findings and conclusions. Validity is the extent to which a measure describes or quantifies what is intended. Reliability is the extent to which it consistently produces the same response in similar circumstances.

Solution: Researchers should use validated measures which are suitable for the intended outcomes of the programme, and appropriate for the target population.

5 Small sample size



Problem: If there are not enough participants in the study it is hard to have confidence in the results. Small sample sizes increase the probability that a genuinely positive effect will not be detected. They also make it more likely that any positive effects which are detected are erroneous. In addition, smaller sample sizes increase the probability that the intervention and control groups will not be equivalent in RCTs.

Solution: Researchers need to be realistic about the likely impact of their programme and potential attrition, and to use power calculations to identify the appropriate sample size. Use strategies to recruit the correct number of participants and retain them in the study, such as financial compensation. EIF will not consider evaluations with fewer than 20 participants in the intervention group.

6 Lack of long-term follow-up



Problem: Studies which do not assess long-term outcomes (at least one year post-intervention) – or do not assess them well – cannot tell us if short-term effects persist. Long-term outcomes are often the most important and meaningful outcomes, in terms of the ultimate goal of the programme.

Solution: Researchers should plan data collection to capture both potential short- and long-term outcomes. Guard against problems which are particularly likely to damage the quality of long-term outcome analyses: maintain comparison groups, attempt to minimise attrition, and conduct analysis to account for attrition.