



How to conduct a pilot impact study

A pilot impact study considers an intervention's potential for improving child outcomes by objectively assessing changes in children's behaviour, feelings, thoughts or attitudes. Essentially, piloting involves measuring participants' baseline level of an identified short-term outcome in your theory of change before running the intervention or service, and then measuring the same outcome in exactly the same way after the participant has completed their involvement in the intervention or service. This will give you a measure of distance travelled, and can indicate that the intervention or service has the potential to improve the outcomes stated in the theory of change, although a pilot impact study cannot prove that your service or intervention caused any measure of improvement.

Please note that the term pilot study can also be used to refer to pilot randomised controlled trials (RCTs), which are used to test whether it would be feasible to run a full-scale study for a particular intervention, and are designed to pilot the evaluation methodology rather than the potential to improve outcomes. Pilot RCTs are not covered in this section. For information on impact evaluation such as RCTs see step 6.

You may decide to combine your pilot impact study with your process evaluation (step 4). This can be a sensible approach to ensure you are not duplicating effort and overburdening participants by asking them to take part in multiple data collection exercises.

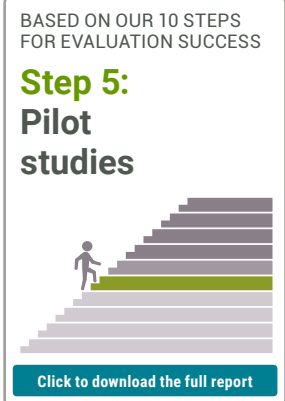
This section summarises the factors you need to take into consideration when piloting for outcomes. These include how to decide which assessment tools to use to measure your outcomes, what sample size you need to include in your pilot, and what to think about when interpreting the results.

Selecting appropriate assessment measures

Choosing appropriate assessment measures is crucial to the success of your pilot study. Positive pilot study findings depend on the use of assessment measures that are **appropriate for the child's age and consistent with the outcomes identified in the intervention's theory of change**. Therefore, the first step when determining which measures to use is to identify those measures that are designed to assess the short-term outcomes you have identified in your theory of change with the age group and abilities of your target population.

Similarly, pilot studies are unlikely to produce meaningful findings unless the measures used have been verified to be **valid and reliable**. Validity refers to whether the assessment tool consistently measures the behaviours, attitudes, feelings or thoughts that it was developed to measure. Reliability refers to the extent to which the measure will yield the same results if repeated multiple times. If a measure has not been shown to have high validity or reliability, you will not be able to draw any meaningful conclusions about your intervention or service.

We therefore **strongly encourage you to look for established measurement tools**, rather than trying to create your own. Established tools can give you the assurance and confidence in your results that other measures cannot. Furthermore, once you have identified the measures you want to use, we recommend that you do not try to modify or remove questions. Doing this can affect the validity, reliability and sensitivity of the measure, making the results meaningless. [See page 5](#) for guidance on where to look for established measurement tools.





Deciding on your sample size

When planning who to recruit to take part in your pilot study (your sample), you will need to consider the following:

- The representativeness of the sample, both in terms of size and characteristics.
- How to encourage and maintain participation throughout the pilot impact study.
- How to gain ethical approval.

Representativeness

A study's representativeness refers to the extent to which its findings are generalisable to circumstances and populations outside of the study. In other words, how representative your sample is will determine how confident you can be that the results from your pilot study would be similar to those seen in the entire target population.

One way to increase representativeness is to increase the sample size of your study (as discussed further below, we suggest a minimum of 20 participants). Another way to increase representativeness is to ensure that the sample in your study has the same characteristics as the intervention's target population. If you are measuring all participants as part of your delivery, rather than a subset purely for evaluation purposes, then you can be more confident in your representativeness. However, please note that due to its design, you will only be able to have limited confidence in the generalisability of a pilot impact study using a pre- and post-test method with just one group.

Recruitment and retention

For the purposes of your pilot study, recruitment and retention should be sufficiently high for an intervention to be viewed as feasible. It is up to you to determine what size your sample should be, based on what you have determined about the target population in your theory of change. However, we suggest a minimum sample size of 20 participants for a pilot study.

It should also retain enough participants to ensure that the findings are representative of the target population. We suggest that retention should also be no lower than 60%. This means that if you have a sample size of 20 then you will need to retain all 20 participants throughout the course of the study. You would need a minimum sample size of 34 if you think that attrition could be as high as 40%. You can enhance study retention by offering rewards or incentives for participants to remain in the study. For example, it is not uncommon for participants to be offered gift vouchers involving nominal amounts as a 'thank you' for participating.

For system-level pilot studies you will need to consider how families fit within the whole system, and how you can track them across multiple interventions/services. One way of doing this is to include a unique tracking number for individuals or families who enter the system. If all interventions or services use this number within the system, it is easier to track families throughout the system.

Key considerations when thinking about ethics and consent

There are ethical and consent-related considerations you will need to think about when planning your pilot impact study. The key considerations are:

- **How to gain ethical approval for your study.** Studies with parents and children should gain ethical approval from the local council as well as from any research organisation affiliated with the study (such as a university), or professional body (for example, all studies that recruit families through health services require NHS approval). See the links in the box at the end of this section on how to go about this.



- **How you will gain explicit informed consent from participants to take part in the evaluation.** A project information sheet will help you to provide the information needed to ensure those involved in your evaluation are providing informed consent. This should include details about the evaluation you are conducting and why, what questions you are likely to ask them, and what you will do with that information.
- **The GDPR requirements for the study if you are collecting personal information, including how you will use, store and process the data.** You will need to develop a privacy notice for your evaluation to gain informed consent for data processing. For more information, please refer to GDPR guidelines, such as those in the box below.

Information on ethics

Association for the Directors of Children's Services research governance: <https://adcs.org.uk/general-subject/article/reason-research-governance-guide-and-checklist>

Economic and Social Research Council. *Research with children and young people*. Retrieved from: <https://esrc.ukri.org/funding/guidance-for-applicants/research-ethics/frequently-raised-topics/research-with-children-and-young-people/>

NHS Health Research Authority. *Applying to a research ethics committee*: <https://www.hra.nhs.uk/approvals-amendments/what-approvals-do-i-need/research-ethics-committee-review/applying-research-ethics-committee/>

Information on UK GDPR in research

UK Data Service. *Applying GDPR in research*: <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/gdpr-in-research.aspx>

Interpreting your results

Once you have conducted your measurements before and after the completion of the intervention or service, we strongly advise you analyse the results using statistical techniques that are appropriate for measuring impact. This involves testing for statistical significance, which means that you can be confident that any change you see in your pre- and post-assessment measure is not the result of chance.

You might find it helpful to consult with a data analyst at this stage for guidance on how to complete analysis for your dataset. However, we advise that you keep this analysis as simple as possible for the pilot impact study, keeping in mind that the analyses must be appropriate for the nature of the data and the size of the sample. A paired sample t-test is an ideal analysis for pilot evaluations wishing to compare the pre- and post-scores of the same participants. Further guidance around how to complete this type of data analysis is provided in appendix D of **10 steps for evaluation success**.

A **lack of a statistically significant result** should not necessarily be interpreted to mean that the intervention is fundamentally ineffective. Rather, disappointing pilot study findings often mean that revisions to the intervention's content or evaluation design are necessary. Revisions to the intervention's content frequently involve revisiting the intervention design by addressing the three questions that were first considered when the theory of change was confirmed (see step 1). Disappointing findings can also occur because of weaknesses in the pilot study design, such as in choice of measurement or sample size. A second pilot study involving different measures and a larger sample size may be necessary.





A **statistically significant improvement** in at least one of the intervention's primary child outcomes is a substantial achievement and is viewed by many as a preliminary indicator that the intervention will prove to be effective after more rigorous testing has occurred. However, positive findings from a pilot study should never be interpreted to mean that the intervention is effective. The lack of an equivalent comparison group in most pilot studies makes it impossible to understand what might have happened if the intervention hadn't taken place. It is therefore very important that you are transparent when writing up your results and that you fully acknowledge the limitations of your study and any assumptions you have made. However, positive findings may indicate that the intervention or service is ready to have a more rigorous evaluation of impact. This is discussed in the next step.

This summary is based on two EIF guides:

- *10 steps for evaluation success*: <https://www.eif.org.uk/resource/10-steps-for-evaluation-success>
- *Evaluating early help: A guide to evaluation of complex local early help systems*: <https://www.eif.org.uk/resource/evaluating-early-help-a-guide-to-evaluation-of-complex-local-early-help-systems>

Where to look for established measures:						
Sources of information on validated measures	Focus of intervention/ service	Child outcomes				Other outcomes
		Physical	Social emotional	Cognitive	Behavioural	
<i>Foundations for Life: What works to support parent-child interaction in the early years?</i> Early Intervention Foundation: https://www.eif.org.uk/report/foundations-for-life-what-works-to-support-parent-child-interaction-in-the-early-years	Early years		X	X	X	Parenting
Child Outcomes Research Consortium: https://www.corc.uk.net/	Nonspecific		X		X	
The Education Endowment Foundation's assessment of the validity and reliability of social and emotional skills measures: https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/measuring-essential-skills	Nonspecific		X			
Reducing parental conflict measures selector. Early Intervention Foundation: https://www.eif.org.uk/resource/rpc-measures-selector	Parental conflict		X	X	X	Parent relationship outcomes
List of reviewed measures. California Evidence-Based Clearinghouse for Child Welfare: http://www.cebc4cw.org/assessment-tools/measurement-tools-highlighted-on-the-cebc/	Nonspecific		X			Parenting, domestic abuse, substance misuse
Early years measures database. Education Endowment Foundation (EEF): https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluating-projects/early-years-measure-database/early-years-measures-database/	Early years		X	X		
ETS Test Collection: https://www.ets.org/test_link/about	Nonspecific	X	X	X	X	Multiple
Deighton et al., 2014. Review of measures: https://capmh.biomedcentral.com/articles/10.1186/1753-2000-8-14	Nonspecific		X			
Denham and Hamre review, 2010: https://www.semanticscholar.org/paper/Compendium-of-Preschool-through-Elementary-School-Denham-Ji/e55c3929969c5b1ffb35966ead41a08b1e040aaf	Nonspecific		X			