

Types of Impact Evaluation: Widening Participation - Access and Participation

Adapted from: Office for Students: 'Access and participation standards of evidence' (2019)

The standards framework: The standards aim to facilitate robust and rigorous impact evaluation by providing a shared reference framework across higher education providers. In line with evaluation tools and standards developed elsewhere, the model is based on three types of evaluation which generate different types of evidence: the **narrative** of evaluation (knowing what you are doing and why); **empirical enquiry** (evaluation to measure changes generated by different activities and practices); and **consideration of causal claims** (what impact can you identify as a direct result of your activities?)

	Description	Evidence	Claims you can make
Type 1: Narrative	The impact evaluation provides a narrative or a coherent theory of change to motivate its selection of activities in the context of a coherent strategy	Evidence of impact elsewhere and/or in the research literature on access and participation activity effectiveness or from your existing evaluation results	We have a coherent explanation of what we do and why Our claims are research-based
Type 2: Empirical Enquiry	The impact evaluation collects data on impact and reports evidence that those receiving an intervention have better outcomes, though does not establish any direct causal effect	Quantitative and/or qualitative evidence of a pre/post intervention change or a difference compared to what might otherwise have happened	We can demonstrate that our interventions are associated with beneficial results.
Type 3: Causality	The impact evaluation methodology provides evidence of a causal effect of an intervention	Quantitative and/or qualitative evidence of a pre/post treatment change on participants relative to an appropriate control or comparison group who did not take part in the intervention	We believe our intervention causes improvement and can demonstrate the difference using a control or comparison group

The standards of evidence build on the work of Crawford et al. in 2017 that established a framework for types of evaluation of the impact of outreach (commissioned by the Office for Fair Access (OFFA) and the Sutton Trust).

Overview of Types of Impact Evaluation:

Type 1: Narrative		Type 2: Empirical Enquiry <i>(encompasses Type 1 and the following)</i>		Type 3: Causal claims <i>(encompasses Type 2 and the following)</i>	
✓ Yes Please	✗ No thanks	✓ Yes Please	✗ No thanks	✓ Yes Please	✗ No thanks
Coherent strategy	Disjointed activities	Clear aim of what activities seek to achieve	Aims developed after activity	Have a target as well as a control or comparison group	
Approach and activities underpinned by evidence from literature or other evaluations	No rationale for developing approach and activities	Select indicators of your impact	No concept of measuring success	Could use an experimental or quasi-experimental design	Using groups that are not comparable
Shared understanding of processes involved	The model of change is not shared	Quantitative or qualitative data – or both, ‘triangulation’ is good!	Information not systematically collected	Think about selection bias and try to avoid it	Selection bias in comparator groups
Reason for activity	Ad hoc activities	Pre/post data (minimum two points in time)	Only collect information once		
Clear conception of why the changes you seek to make are important	No understanding of needs of target groups	Analysis competently undertaken	Data not related to the intervention		
Programme reviews	No review or evaluation	Sharing of results and review of activity	Results not used to inform decisions		

The matrix below sets out an indication as to which types of evaluation might be appropriate for different types of activity, though this should not be regarded as restrictive and, in particular, will vary according to the nature of the project and objectives.

	Narrative Evaluation A coherent explanation of why the activity is expected to work and why	Empirical Enquiry Collects data to show those receiving an intervention have better outcomes	Establishing Causality Provides evidence of a causal effect
Multi-activity intervention programmes (e.g. transition support programmes)	◆ Important for all activities to inform programme choice and delivery	◆ Important for all activities to justify use of resources	★ Commended for resource intensive programmes for which an evidence base needs to be established and where there is access to reliable outcomes data and appropriate comparison groups
Intensive interventions (e.g. residential programmes)	◆ Important for all activities to inform programme choice and delivery	◆ Important for all activities to justify use of resources	★ Commended for resource intensive programmes for which an evidence base needs to be established and where there is access to reliable outcomes data and appropriate comparison groups
Long term interventions (e.g. mentoring programmes)	◆ Important for all activities to inform programme choice and delivery	◆ Important for all activities to justify use of resources	★ Commended for resource intensive programmes for which an evidence base needs to be established and where there is access to reliable outcomes data and appropriate comparison groups
One-off interventions (e.g. campus visits, subject taster sessions)	◆ Important for all activities to inform programme choice and delivery	★ Important for innovative projects for which an evidence base needs to be established	⊘ Not usually feasible unless part of a multi-intervention package
'Light-touch' intervention (e.g. information dissemination projects)	◆ Important for all activities to inform programme choice and delivery	★ Important for innovative projects for which an evidence base needs to be established	★ Commended in situations where it is possible to capture appropriate outcomes data and the effect of the intervention can be isolated

◆ Expected for all types of activities; ★ Commended for resource intensive and pilot interventions; ⚡ Highly commended if conditions allow and conducted appropriately; ⊘ May not be feasible unless special conditions apply.

The types of evaluation are not hierarchical –i.e. it is not a matter of trying to aim for a ‘higher’ type. Indeed, it is better to aim for a strong Type 2 evaluation as opposed to an unrealistic or badly executed Type 3 evaluation. A well-formulated Type 2 evaluation is often more realistic, especially in the context of outreach where there are good working relationships in place with stakeholders who can provide access to good quality data and insights.

What standard of evidence should I aim for?

The selection of evaluation approach is important since higher quality research designs can help to meet the challenge of attributing outcomes to the activity in question (as opposed to other influences), whereas lower quality designs reduce confidence in whether it was the activity that generated the outcomes. However, there is no simple answer to the question of what will provide the best evidence for any particular type of access and participation activity. It depends on what is being measured and in what context.

As a rule of thumb, the more **resource-intensive** an activity the higher one would wish the standards of evidence to be to show impact because it would be risky to continue to devote the level of resource unless the activity can be shown to have the beneficial impact it is aiming for.

Review and Reflect: Two worked examples of different types of evaluations

<p>EXAMPLE 1: Information through a football club</p> <p>Overview: The approach brings current students from a partnership of universities into a local football club to help with the coaching of the players and, at the same time, to deliver information, advice and guidance (IAG) about HE. These students are either studying sports science courses or are members of university football teams. Other activities such as university visits have been organised to reinforce the student coaches' work.</p> <p>About the evaluation: The research focused on the benefit of participation to their school performance and to their HE aspirations and awareness, their learning from the project and intention to consider HE study.</p> <p>Methodology: The opinions of participants about the project were surveyed by questionnaires and focus groups at the end of the project. The evaluation used a questionnaire that had been tested in the previous year to provide comparative data over time. Research with the student coaches and parents/carers complemented the project.</p> <p>Over to you: What type of impact evaluation do you think this intervention is? Why? What is good about this evaluation? How could the approach be improved?</p> <p>Possible Answers:</p> <p>What type of impact evaluation is this?</p> <p>Type 1</p> <p>Why?</p> <p>The partnership gathered data that shows some change amongst those receiving the intervention. It is developing a narrative account to motivate its selection of outreach activities in the context of a coherent outreach strategy.</p> <p>What is good about the evaluation?</p> <p>There is a holistic approach to understanding change for the participants in their context of sport and home.</p> <p>How could the approach be improved?</p> <p>The evaluation is based on research carried out shortly after the completion of the project, so only provides an assessment of views of the short-term impact. The introduction of longitudinal tracking is desirable to show the medium- and long-term impacts such as variations in attainment and HE progression rates. Comparative research, pre/post intervention, or gathering data on outcomes from a matched sample of participants and non-participants would help to confirm that the work is making a difference.</p>	<p>EXAMPLE 2: Individualised tuition for disadvantaged students</p> <p>Overview: The project supports young people from disadvantaged backgrounds to progress to selective universities through offering one-to-one academic tuition in schools with volunteer tutors and personalised university support and guidance.</p> <p>About the evaluation: The evaluation aimed to assess the effectiveness of the tutoring programme at raising pupils' GCSE grades.</p> <p>Methodology: For GCSE grades – a matched comparison group design using propensity score matching and pupil data. For university places – pre/post data showing the change in number of pupils attending top universities from each school, from before the project started working with them to after.</p> <p>Over to you: What type of impact evaluation do you think this intervention is? Why? What is good about this evaluation? How could the approach be improved?</p> <p>Possible Answers:</p> <p>What type of impact evaluation is this?</p> <p>GCSE outcomes are Type 3 and university places are Type 2.</p> <p>Why?</p> <p>GCSE outcomes: They can demonstrate causality using a control or comparison group. Propensity-score-matching is widely considered a robust approach to creating a comparison group, provided that the factors on which participants are matched are sufficiently comprehensive and meaningful. (N.B. The project was unable to include 'level of motivation' as a matching factor, but was able to provide evidence to successfully make the case that this does not significantly weaken the findings).</p> <p>University places: The evaluation design compares the outcomes of pupils in the project with pupils from the same school who did not take part in the previous period – this gives an interesting benchmark, but there is likely to be some systematic difference between those pupils who did and did not take part in different years.</p> <p>What is good about the evaluation?</p> <p>There are robust 'before' and 'after' measures for the intervention. GCSE grades are externally verified measures.</p> <p>How could the approach be improved?</p> <p>By introducing a matching technique that mitigates bias, they can confirm with more certainty and accuracy that the programme is having a positive impact. For future impact evaluations this was tackled by using UCAS Strobe data to compare the participant outcomes against a matched 'control' group. Using focus groups or interviews could highlight which aspect of the tutoring is particularly helpful.</p>
--	--