Description:

Tracking is a means of verifying the outcome achieved by participants, especially the outcome relevant directly to HE access, participation and success. Tracking participants helps evaluators to understand whether the final result the activity is aiming for is realised in practice (although does not necessarily prove causality). For example, if you are aiming to influence young people's progression in education, and/or their attainment outcomes, then it becomes crucial to have some element of longitudinal tracking to establish the educational destinations and achievements of your participants.

Tracking could be undertaken in different ways, which may be more or less available to providers depending on their contexts. Options include:

Keeping in contact with or following up your participants over time, for example, through telephone or postal surveys. If you are only working with a small group of close-knit participants then you may be able to keep in touch over time. This approach can help you to find out about outcomes in a direct and timely way, and you also have the possibility of canvassing perceptions on the influence of your provision. However, it is likely to be resource intensive, responses may be partial and there may be a problem of response bias if you have data for only a subset of all participants.

Using your partners' datasets (e.g. UCAS applications data from schools/colleges or attainment data from local authority sources). This data can be direct and timely but it is going to depend on the nature of the relationship with your partners and data sharing arrangement processes will need to be in place. If data comes from different sources (e.g. multiple schools) it can often be difficult to aggregate.

Linking your participant data to administrative datasets. For example, pupil-level information is available via the National Pupil Database (NPD), plus HESA provides information on HE applicants and entrants (either directly or via the NPD linked dataset). UCAS also has a data service including 'STROBE', which gives access to applicant tracking data and comparison group data at aggregate level based on applications and acceptances. Administrative data can prove to be a very helpful data source on outcomes, plus, pupil-background indicators can be included which can significantly improve the robustness of the evaluation design. However, this type of data often takes time and specialist expertise to access (and there are costs involved for HESA and UCAS data). For linked data there can be a time delay in receiving data, and you may need to put new processes in place for receiving and securely working with the data. Going through a tracking service can be a convenient option. Tracking systems – HEAT, EMWPREP, Aimhigher West Midlands – offer shared facilitated access to regular longitudinal datasets (with costs involved).

Using your own datasets (for example matching outreach participants to applicants on undergraduate courses and/or using your student records systems to track student and graduate outcomes). Since this type of data is available 'in-house' it can be quick and direct to obtain. The options are going to depend on the accessibility of individual level data within your institution and who you want to track. An issue for evaluation of outreach is that institutional data will not include outcomes for participants who progress to other HE providers.

Applications:

Tracking could be used to capture the intermediate and long term outcomes for outreach participants progress through education or could look at the progression, attainment and destinations of target groups in higher education students.

Strengths:

Although data on long term outcomes does not necessarily say anything about the influence of your activity on the outcomes, it can be used to make inferences. The strength of the conclusions will depend on the type of study and the methods of analysis used. When used as part of an experimental or quasi-experimental research design, tracking data can help to identify firm conclusions about the impact your work.

Using a tracking service such as HEAT or EMWREP allows you to record individual participants' interactions with other higher education providers (subject to consent). These systems also allow monitoring of key information such as an individual's background characteristics (gender, ethnicity, first in their family to HE, disability, care status etc). This data can then be used to analyse differences in attainment and/or progression to HE whilst controlling for other factors.

Weaknesses:

For schools outreach work it is often a challenge to put in place tracking that sustains across years,

particularly where the outreach is with younger students at a distance from HE, and for whom it is going to be hard to establish any direct connection between outreach and education progression in later years.

There are technical and logistical challenges to accessing and using different types of tracking data (as discussed above).

Mixed Methods: Tracking can provide important data to inform consideration of long terms outcomes for participants, but it is also useful to include a mix of methods to ensure that the evaluation gives consideration of perspectives on why and how the activity affects outcomes and the processes involved. Demonstrating a result was achieved is crucial to assess impact, but unless you know how and why the result was achieved it can be hard to replicate.

Expertise: Depends on the methods, but High for methods that involve linking to administrative datasets and use of experimental or quasi-experimental designs.

Ethical considerations: You will need to make sure that appropriate permissions are in place from those involved to record, store, and link their data. One of the trickiest parts is making sure participants fully understand the reasons for collecting their data, how it will be used and protected, and the length of time it will be kept. As well as a Data Protection statement you may want to consider briefings on consent as an introduction before collecting data from students.

If you are planning to link to administrative datasets, it is important to make sure that you have the correct consent in place enable you to access the type of individual personal data you need to identify them (which might entail collecting permission to hold data and permission to link/track individuals separately).

Once data has been linked, it can be stored in a 'de-identified' format, for example using unique reference numbers, which reduces the sensitivity of the data (but can still allow for nuanced analysis depending on the types of fields included in the dataset).

If you are identifying a control group, then permission is needed from this group as well as those who took part in your activities.

If you are linking to data from other organisations, i.e. from administrative sources, your partners or between departments, you will need appropriate data-sharing agreements to be in place.

Reporting protocols need to be applied, to avoid the risk of individuals being identified from data presented in evaluation reports (and other evaluation outputs).

Work planning: Planning for longitudinal tracking needs to be thought about early on before you start to collect participant data, not only to make sure that you have the correct permission to track in place, but also to ensure you collect enough data to be able to track people across datasets on. Usually you will need a range of identifiers if you are linking individuals across datasets – as a minimum full name, gender, date of birth, school, postcode.

It is also a good idea to collect data on factors which are likely to affect the outcomes for your target group so that these can be taken into account as part of your analysis (for example, gender, ethnicity, prior attainment and motivation have been shown to affect educational outcomes).

You might need to put in place data sharing agreement with others who work with your target groups if you intend to track on individual level data (e.g. with schools and colleges to access the Unique Pupil Number (UPN) in the learner record for your target group).

If you are planning on using quantitative methods at the analysis stage, then you want to aim for individual-pupil-level data rather than aggregated measures. If you only record outcomes across a whole group, then you will not be able to say which characteristics of individuals benefited.

There can be significant challenges in terms of negotiating permissions to access secondary sources of data such as the NPD. For example, making sure your data retrieval, data storage and data destruction procedures are in line with the requirements of the data providers. There are also time lags involved (it takes about 6 months to process NPD requests). Therefore, you will need to plan ahead and also

ensure sufficient resource and expertise is in place. Members of tracking services may be able to receive the data from this source (which could be individual level or aggregate level data depending on the permissions involved). You will still need skills to analyse and interpret the data.

Analysis:

Individual-level data supports tracking individuals across the student lifecycle, to assess rates of positive outcomes and is amenable to fine grained analysis, for example, to show if there were any demographic sub-groups with worse/better outcomes. If the programme includes different activities, remember to record the take-up and frequency so that you can consider how differences in the interactions affect the outcomes.

Depending on the population you have tracking data for, there is potential for quasi experimental evaluation. When using large-scale administrative datasets, for example a data extract from the NPD (linked to student outcomes), there is potential for making comparisons against similar learners who did not receive the intervention (for example, using Propensity Score Matching (PSM) or Case Control Matching (CCM) techniques). This work would then allow for assessment of the likely impact of the outreach on the participants involved compared to the matched comparison group.

Special consideration should be given to the issue of missing outcome data and how to handle this (e.g. participants who did not respond to follow-up or for which there was not enough data to track). Commonly participants without an observed outcome are excluded, but you will lose statistical power by reducing the sample size and bias may be introduced if the missing data is non-random (e.g. related to a participant's response to intervention). Depending on the reasons for the missing data, you could choose to assign outcomes, although care must be taken if you decide to do this so as to avoid bias.

The analysis should control for as many factors affecting the outcome as possible – you should be aiming to test the effect of the activity 'all other things being equal'. This may depend however on which fields you have available in your dataset, which may need to be specified at the data collection stage. It could be that you need to collect additional data in order to control for variables which may affect the outcome (e.g. prior attainment, engagement in education).

Reporting:

How you report the results of tracking data and the claims you can make will depend on the study design. A mixed methods study would allow you to report outcomes data quantitatively alongside qualitative data to describe the processes involved. An experimental or quasi-experimental design would allow you to test the strength of the relationship between the intervention and the outcome.

Useful Link(s):

DfE manages the application process for NPD and associated linked data: www.gov.uk/government/collections/national-pupil-database

HEAT, EMWPREP, and Aimhigher West Midlands offer a shared service facilitating access to regular longitudinal datasets: www.heat.ac.uk; www.emwprep.ac.uk; www.aimhigherwm.ac.uk

HEAT have published a video giving advice on sourcing data comparator group data for KS4 tracking data, see: https://vimeo.com/412251062/82df9543cb

The UCAS STROBE service offers access to applicant tracking data and comparison group data at aggregate level: www.ucas.com/data-and-analysis/data-products-and-services/strobe

Tracking participants: Case Study/Practice Example

Key Stage 4 attainment is associated with future HE progression and probably accounts for nearly all of the social stratification seen in later stages of education.[1] Attainment is sometimes seen as an early outcome to predict later HE progression and is often available within the life span of an outreach project. A study by a Uni Connect partnership was designed to test whether there was a correlation between participation in a pre-16 activity designed to support motivation and exam preparation and improved attainment at Key Stage 4, taking account of prior attainment. These activities were offered in partnership with an external provider who ran a series of workshops in schools focused on exam and revision skills.

The HEAT's Key Stage 4 report was used for this study, looking at participants who had received above a threshold number of contact hours for the activity. Firstly, results for participants were compared with the school average levels of attainment. The results showed that Attainment 8 Scores were on average +6.6 grades higher than the school average. Participants with medium prior attainment saw the biggest difference (+5.5 grades higher than the average for the same attainment band).

This initial analysis used the school average as the comparator, but it is likely that the activity participants were not representative of their classmates (targeting often determines this). The partnership wanted to identify a similar comparator group to the participant group, using a quasi-experimental research design, in order to generate Type 3 (causal) evidence. Analysis was undertaken looking at respondents to the Uni Connect baseline survey, who were also being tracked through HEAT (with consent). A group of students were identified as a 'non-treatment' group (not receiving any activities). Students from the activity group were matched with those in the non-treatment group, based on variables including Gender, Ethnicity, IMD and IDACI quintile and KS4 performance of school. These variables influence attainment.[2] Case Control Matching was undertaken in SPSS: a match was found for 85% of cases (unmatched records were discarded). The sample size was therefore reduced but as this was a large scale activity, the analysis was still viable. The results were in line with the earlier findings: participants achieved an average of +4 grades higher than the matched non-participant group (statistically significant at 5% level). The difference is below that compared to the analysis based on the school average (which probably overestimated the effect). Participants at all levels of prior attainment performed better, with participants from the highest prior attainment band showing the biggest improvement compared to peers with similar prior attainment.

Using matching techniques to build a comparator group does not mean that the analysis is able to account for all differences between the activity and non-treatment group. In particular, it is possible that there are differences in terms of their motivation to participate in outreach. Controlling for motivation should be considered when setting up experimental methods, and it is important that selection practices do not result in differences in motivation levels, or other factors, between the two groups.

These results give an indication that the activities may be positively supporting the learners, especially those at risk of not fulfilling their potential. However, setting out clear Theories of Change for raising attainment activities is also important because the link between raising motivation and attainment is debated.[3] [4]

[1] Crawford, C. (2014) The link between secondary school characteristics and university participation and outcomes, London: Department for Education.

[2] Sylva, K., Melhuish E., Sammons, P., Siraj, I., Taggart, B., Smees, R, Toth, K., Welcomme, W. and Hollingworth H. (2014) Students' educational and developmental outcomes at age 16. Effective Pre-school, Primary and Secondary Education (EPPSE 3-16) Project. Research Report. Department for Education.

[3] Cummings, C., K. Laing, J. Law, J. McLaughlin, I. Papps, L. Todd and P. Woolner, P. (2012) Can Changing Aspirations and Attitudes Impact on Educational Attainment? A Review of Interventions, York: Joseph Rowntree Foundation.

[4] Gorard, S. and B.H. See (2013) Overcoming Disadvantage in Education, Abingdon: Routledge