



# DATA COLLECTION / USE OF BASELINES

## SUMMARY

Measuring impact will always involve collecting data to show the pupils' initial state, their record of attendance, and their final state. Collecting the data can be problematical and the data collected needs to be accurate and reliable if conclusions are to be valid. This guide helps teachers to be aware of, and to overcome some of the pitfalls in data collection.

## INTRODUCTION

Measuring impact requires data. This data must be accurate and reliable, otherwise any conclusions drawn from it may be flawed. Impact can only be measured if there is data available to reflect the status of the participants both before and after the activity. Finally, there must be sufficient data to ensure that any conclusions have general application (very small samples are more akin to case studies). Data collection requires time and planning in order to be effective. Schools are data rich but access is not always straightforward and some of the data may be of questionable reliability.

This leaflet provides guidance as to what data is worth collecting and how to collect it. It also looks at commonly available baseline measurements and their advantages and disadvantages.

## WHAT DATA DO YOU NEED?

The purpose of measuring impact is to show that the activity undertaken benefits those that take part in defined ways. This can only be demonstrated if it can be shown that there was an increase in some skill, ability or desirable outcome over the course of the activity, and that this increase was actually related to the activity (i.e. it would not have occurred 'naturally'). The following pieces of data will therefore be necessary:

### **An attendance record**

Only those that have attended the activity regularly can benefit from it. It is therefore necessary to show that the pupils concerned have attended sessions.

### **Measurement of outcomes**

This may occur at the end of the activity or programme and/or some time after completion. If long term benefits are an objective, data will need to be collected over an extended period.

### **Pre-test or predictive data**

Progress can only be established by demonstrating that pupils have a higher level of skill / attainment after the activity than before, or that their achievement is greater than was predicted before the activity or programme took place. Such measurements are known as *baseline data*.

Sometimes the data collected may not be quantitative, but is converted into a quantitative measure afterwards, e.g. Pupils may be asked if they feel more confident in performing in public after a drama course. It might then be reported that (e.g.) "76% of pupils said they were now more confident in performing in public following the course". Note that such an outcome could be established in two ways. The pupils could give themselves a confidence 'score' before and after the course, or could simply report that they were *more* confident now (in which case the baseline measurement is their own perception before the course).

## Other information

Depending on what the purpose of the evaluative study is, other data may be needed. This may include any of the following (all of which will be available in a school):

- Whether in receipt of free school meals
- Gender
- Age
- Ethnic group
- School attendance
- Behaviour record
- Rewards record

## DIFFICULTIES WITH DATA COLLECTION

It is not always easy to collect the data you want, and sometimes the way the data is collected leads to it losing reliability. Problems and hints for getting around them, are listed below.

### Data is needed across a number of activities

This is usually attendance data. It relies upon all course tutors filling in and returning registers. Taking a register is always a requirement for health and safety reasons, yet on occasions this is not done. Tutors may forget to return registers to a central coordinator who is responsible for the data analysis. Schools with a delegated study support coordinator who has time to chase paperwork have less difficulty. A further method that has been used successfully is to have central registering for all study support activities, so that pupils register in one place before going to their activity.

### Necessary central data is incomplete

Predictive and national test data is not always in place for every student, particularly if they have moved to a different local authority in their school career. If the measurement of impact study requires this data, then those pupils who do not have it must be omitted from the sample.

### Self-reported data issues

Care must be taken with any self-reported data. Pupils are often conditioned to try to give the 'right' answer. If asked if a particular skill or capability has increased, they know that the expected or desired answer is 'yes', and will often give it. Scoring such aspects rather than just a yes or no response is more reliable. A further problem is that friends will often give similar or identical responses - a sort of 'group answer', which does not accurately reflect every individual's opinion. Low self esteem can sometimes mean that a pupil will give consistently low (or high) scores for any form of self assessment. Objective measurements should be used wherever possible, although self-reporting can provide some useful qualitative evidence from some individuals, even if it is not consistent.

### Getting useful data from questionnaires

It is sometimes difficult to get a good return of questionnaires, and the quality of pupil responses may be poor. There are various methods that can help here:

- Try to get the questionnaires done in school (assuming pupils are the respondents). Return rate will be poor if taken home. Form tutor sessions are often useful for this, or the questionnaire can be done in the activity itself if it relates to just one activity.
- Poor quality responses may be linked to poor survey design. Think carefully about exactly what you want to find out, and target questions accordingly.
- Keep the questions and the questionnaire as short as possible
- Use simple and direct language.
- Put the most important questions first, in case the whole questionnaire is not completed.
- Make sure any scoring system is clearly explained (e.g. If scoring between 1-5, some pupils will instinctively think 1 is the 'best', but others will think 5 is).

- Reduce the amount of writing - pupils can circle scores or multiple choice answers

### Time taken to collate survey data

It takes time to collate any data, but surveys can be particularly time consuming. A designated study support coordinator or data manager in the school will help, and it is also possible to use web-based survey analysis, which is sometimes free if requirements are limited. A useful survey of web based tools can be found at

<http://www.ucl.ac.uk/learningtechnology/opinio/survey-software.html>

## USE OF BASELINE DATA

Baseline data is necessary if the impact of an activity or a programme is to be measured. Baseline data can be acquired in various ways and all have their advantages and disadvantages. The different types of baseline data are outlined below.

### Pre and post tests

The same (or similar) test is administered before and after the activity, with the pre-test providing the baseline measurement. Care must be taken that the test focuses on the desired outcomes stated in activity objectives. Pre and post testing should only be used for a single activity. It is much less reliable for a programme because the time between the two tests allows other variables (apart from attendance at study support) to come into play.

**ADVANTAGES:** Easy to obtain, and generally accurate and reliable provide certain factors are borne in mind (see 'disadvantages').

### DISADVANTAGES

If the pre and post tests are identical, there is a possibility that recall may improve the post test (it is particularly important that the answers to the pre test are not revealed). If the pre and post test are not identical, you cannot be certain that they are of exactly equal difficulty, which could skew the results.

### Predictive data

Many schools and authorities use the Fisher Family Trust or other predictive modelling. The predictions relate to attainment in national tests and therefore are only used to measure the impact of programmes, not individual activities.

**ADVANTAGES:** As predictions go, these models are fairly accurate and reliable, taking into account a number of factors.

**DISADVANTAGES:** The predictions cannot be absolutely reliable for individuals, as many factors can influence attainment over a period of time. The models assume the status quo more or less being maintained over the period of the prediction. However, individual circumstances can change considerably e.g. family breakup, bereavement, prolonged illness, or significant improvement in personal or family circumstances etc. Such individual variations are much less important if very large samples are used (e.g. a whole year group).

### External tests

SATs and GCSE grades may be used as the basis of predictions (or as outcome measurement).

**ADVANTAGES:** External tests are standardised, linked to specific objectives and their marking is moderated to ensure equal standards across the whole entry. They are therefore generally accurate and reliable.

**DISADVANTAGES:** The tests are accurate and reliable but are a snapshot at a particular moment. Candidates may under-perform due to nerves, personal state on the day, and levels of motivation

prior to the test. They may score better than expected (or worse) due to luck with questions, or may get by by revising very hard (scoring well on recall questions) but with little understanding. It is not unusual for individuals to do significantly better or worse than expected from previous external tests. This can skew results in small samples but will have much less effect if the sample is large.

### **Internal testing**

Internal testing is sometimes used to make predictions or judge baselines. Pre and post testing (above) is internal testing, but here we are more concerned with tests and assessments that are not directly linked to an activity.

**ADVANTAGES:** Internal tests can be constructed to a particular agenda or set of objectives, so there may be more precision than in national tests.

**DISADVANTAGES:** Internal tests do not undergo the scrutiny that external tests do. Language level, clarity or layout may affect performance. Many internal tests are not moderated, so that if they are marked by more than one individual there may be inconsistency of marking. Awarding of grades or levels on internal assessments is usually much less reliable and rigorous than in external tests.

The reliability and accuracy of the data also depends upon what it is. For instance, if you want to establish a baseline for attendance at school, the data will be much more accurate and reliable than if you wanted a baseline for 'understanding scientific concepts' for instance.

In summary, when measuring human factors, all baseline and outcome data suffers from some level of uncertainty. It is simply necessary to choose the most reliable methods available, and to be aware of possible drawbacks.