

To compare or not to compare?

Kathy Sylva

The author is Professor of Educational Psychology in the Department of Educational Studies, University of Oxford, United Kingdom. In this article she reviews the importance of comparison groups, considerations that arise in sample size, and the place of qualitative and quantitative data in longitudinal studies.

A study that is following up participants after some years is a form of longitudinal study. In this kind of research, there are two questions that can be asked:

- what are children/families who participated in the programme like a number of years later?
- What would children/families be like if they had NOT participated in the programme?

The first of these questions is essentially looking for description, it is one way of

looking at outcomes, but you can never be sure what caused them. If this is your research question, you do not require a comparison group.

The second question cannot be answered without a comparison group. This will help to show the effects of a programme by comparing similar populations who did and who did not participate in the programme. Although the use of a comparison group does not mean that you can 'prove' that it was the programme that made the difference, it strengthens the case that it was.

Types of comparisons

There are basically three different types of groups that can be used for comparison:

1. the **control** group where the same kinds of individuals/families from the same kinds of neighbourhoods have been randomly assigned to be in the programme or not (the programme group is sometimes called an 'intervention' or 'treatment' group);
2. the **matched** group in which you first define important characteristics of the intervention group and then match them with individuals in another group that is as much like the first group as possible; and
3. the **comparison** group which is composed of people who are like those in the programme in many ways but did not participate for whatever reason. For example, they could be from a different village.

Of these, the control group gives you the best possibilities for making your comparisons some years later because the assignment into programme/control group was random right at the beginning. There are, however, some drawbacks, particularly the ethical issue of denying a programme to those who want it when the resources are available. If there are not sufficient resources available from the start, it might be possible to offer services at a later stage, in which case families might be willing to be randomly assigned in the first instance, knowing that *everyone* will participate in the programme at some time.

With the second two groups on the list, you cannot know for sure that it was the programme that made the difference, since you do not know whether the people in the programme are different from those who are not. The mere fact of being willing to participate already makes them different in some way.

An important difference between a matched and a control group is that the control group was randomly allocated. The matched group is a little stronger than the comparison group because you know how the two groups are alike and have selected them for their similarity.

The more we have control and matched groups in our studies, the better able we are to limit the influence of other factors. If we are trying to establish effects we need to 'triangulate', which means getting our data from two or more sources and matching or comparing them.

How big should the samples be?

The main constraint when you are trying to find out the effects of a programme are the influences on children/families that are beyond the sphere of the programme. Here is where numbers become important because if the numbers are large enough then unusual circumstances for one child do not influence the overall outcomes. If the numbers are small, then an unusual event for one individual impacts the sample in a disproportionate way. An example is a research study involving the use of a specific methodology to teach a small number of children to read. In the control group there is a child whose aunt is a teacher and suddenly moves in to her home and teaches her to read. In this case, the reading scores of a small group might be 'artificially increased' by the aunt who joined the family.

If both groups contain larger numbers, for example 60+ children, the impact of

one child is not great. If, however, the study includes a small number, for example less than 16 children, dramatic and unusual circumstances for one child may greatly influence group outcomes.

Determining the number of children to include in the sample depends on many factors, including the resources available (time, money, expertise and so on). Calculating the optimum number needs to be done in relation to the number in the total population. This is a complicated process and hard to do. Another approach, not quite so good, is to look at previous studies of the same kind and aim for more or less the same sample sizes. If the groups have been randomly assigned, then the sample sizes can be smaller; if the sample is to be non-random, then a larger sample is needed to take account of greater variation between the groups.

In any case, the sample needs to be representative of the kinds of children you are serving. For example, if you select for tracing only children who are **in** school several years later, this does not represent the whole group as there may well have been some that dropped out.

Qualitative or quantitative?

Whether to do a qualitative research study or a quantitative research study depends on the research questions, the objectives of the study, the audience being aimed at as well as the resources and skills available. And there is also the matter of sample size.

If the study is qualitative then it is impossible to do with 2,000 children. If the study is quantitative then samples of 60 and 60 can be sufficient. In the example the visiting aunt who was a teacher, her impact will be less with a sample of 60 in each group than one which has a smaller sample. You may know about her and it is possible to talk about her in the study, but she does not 'swamp' the outcome.

Both quantitative and qualitative methodologies struggle to interpret the evidence and account for differences between respondents. And in many ways, it is a false dichotomy. It is possible, for example, to have a fairly large scale quantitative study and also to draw a smaller sample from that study for an in-depth qualitative study. Combining the different methodologies should lead to more meaningful findings and interpretations. ○