Introduction to sampling

Introduction

Asking all your service users and partners to be involved in your evaluation may be both costly, time consuming and probably not even feasible - so the next best approach is to speak to some of them – also known as a sample.

Say, for example, you have worked with 100 people in a year; it will be very challenging to speak to all of them about their experiences, while a representative sample of 20 could tell you what you need to know.

Alternatively, you might decide that there is some information you need from everyone you have worked with (such as data on outcomes) and some information you only need from a sample (such as feedback on particular activities). This is a good approach; data on everyone gives you confidence in the scale of achievement while sample data tells you more about how these were achieved.

Sampling may also be applicable if you are conducting research with other groups such as staff, volunteers or stakeholders.

Sampling issues also relate to both quantitative and qualitative research[1], although statistical calculations of margin of error only apply when you are conducting quantitative research (we talk more about this later).

If you do decide to sample, you (and your stakeholders) will still want to be confident that the findings are representative and valid. It's a risk because a poor or unrepresentative sample will not deliver useful conclusions - regardless of whether you asked the right questions or conducted robust analysis.

Regardless of who you are researching it helps your credibility to sample correctly and to do everything you can to minimise bias. This guide summarises the different approaches you can take and aims to help you chose which method to use.

Terminology

When thinking about sampling it's useful to understand four terms:

Population

This is everyone who is eligible for the research. In service evaluations this is everyone who has used the service (or possibly everyone who is referred to the service). You can also think about this as the population you want to generalise about. For example, "50% of our service users take up volunteering opportunities".

[1] For more information on these terms please see our hints and tips document http://www.clinks.org/sites/default/files/EvaluationHintsTips.pdf

Sample frame

This is everyone who you could feasibly conduct research with. Often it will be the same as the population, but there are circumstances when it will differ. For example if some service users move away and are not contactable then they are in your population but not your sample frame.

Selected sample

This is everyone you will try to conduct research with. So if your sample frame has many hundreds of people in it your sample will be a selection from it.

Achieved sample

This is everyone you actually conduct the research with. It's different to the selected sample because there will be people who you can't get in touch with or refuse to take part.

Sampling is therefore a multi-step process and the challenge is that in moving from each step to the next, there is the possibility of introducing error or bias. A biased sample is one that does not reflect the views or experiences of the whole population and provides misleading or wrong findings. For instance, even if you are able to identify perfectly the population, you may not have access to all of them. And even if you do, you may not have a complete or accurate sampling frame from which to select. And, even if you do, you may not draw the sample correctly or accurately. And, even if you do, they may not all come and they may not all stay.

The aim of a sampling process is to minimise or eliminate these risks so that the achieved sample is as representative of the population as possible.

Minimising bias

Firstly, to minimise differences which can introduce bias between your population and your sample frame it is important to keep good records of everyone you work with. This includes their contact details, mobile numbers and if appropriate details of family members[2].

Secondly, minimising bias between the sample frame and the selected sample is the main business of sampling. Below are the main approaches to be aware of.

Simple random sampling

This is the best approach because it gives everyone an equal chance of selection, thereby minimising the chances of bias. To do it you need to list everyone in your sample frame and then make a random selection - like a raffle. These days the best way to draw a random sample is to use the 'RAND' function in excel[3]. This allocates a random number to everyone in your sample frame, you then simply need to rank the data by the RAND variable, thus reordering the list at random, and then take the top 10, 20 or 50 etc. people (see discussion of sample size below).

[2] Please see our guide on engaging users for more information http://www.clinks.org/sites/default/files/AcheivingUserParticpationResearch.pdf

[3] http://office.microsoft.com/en-gb/excel-help/rand-function-HP010342816.aspx

Stratified random sampling

A drawback to a simple random approach is that if you are interested in subgroups - particularly minority groups – you may not select enough people from these subgroups. Stratified random sampling addresses this. It is similar to random sampling but is preceded by dividing your sample frame into different sub groups (or strata) and then using the simple random approach *within* each group. It's a helpful approach if you want to explore differences in views/experiences by subgroups. For example, if your service user population was 90% men and 10% women a pure random sample would reproduce this profile, but if you stratify the sample and select randomly from amongst men and women, you will achieve a more equal balance that helps you compare the two groups.

The importance of reporting a response rate

Another advantage of both random and stratified approaches is that they enable you to calculate a response rate. This is a simple calculation of the proportion of people you selected for the research who actually took part - and is one of the main ways to assess the quality of research because a low response rate greatly increases the risk of bias. If you can quote a response rate it will add greatly to the credibility of your findings. A good response rate is generally over 50%, while 70-80% is exceptional because the majority of people have been interviewed and the risk of bias reduced.

But the overall response rate is not the only measure of success, you also need to consider differential response; whether some groups responded more than others. For example a 50% response rate in the general population looks good, but if only men took part and no women then the sample is clearly biased. The section below on checking for bias looks at this issue further.

Convenience sampling

Having described the best approach to selecting from your sample frame, we now describe the worst, which is convenience sampling. As the term implies, it simply involves using anyone who is available and willing to take part. It could mean just catching any users as they come into your service or approaching stakeholders or partners you know well. The advantage over random sampling is that it is easier; you don't have to work hard to get in touch with people and increase the response rate. The disadvantage is that those who are easier to engage in research are also probably more engaged with the service – so likely to give a more favourable view, or more likely to be biased in other ways.

Quota sampling

Quota sampling is a way to make convenience sampling more robust. It involves taking what you know about your population and designing a convenience sample to match. Say for example you have a good understanding of the profile of all your service users (be this age, gender, ethnicity, level of engagement with the service or something else), you would then use this knowledge to select service users to match that profile. You might then know that in wanting to speak to 30 service users, you want at least 5 to be female, another 5 to be from

minority ethnic groups and 5 to have been out of education for the last year. If someone refuses to take part in the research you would need to replace them with someone who has the same characteristics so that the overall profile remains the same.

The challenge is that getting your proportions right can be difficult as it isn't always easy to find up to date information on the population. Biases may also be introduced because even though some individuals fit the sampling criteria they might differ to the population by some other characteristic – in particular their willingness to engage with the research or the service – which is the key problem highlighted above.

Purposive sampling

The final main type of sampling is purposive sampling which – as the name suggests – is used when you want to focus your research on a particular group. For example, you may be interested in reaching people who failed to complete a programme to find out why. "Snowballing" is a type of purposive sampling used to engage 'hard to reach' groups. It involves individuals who have already taken part in the research using their social networks to find other people who could participate.

We have now described the main types of sampling from your sample frame. There are many others for different purposes. The following link is a good source for more information:

http://betterevaluation.org/plan/describe/sample

Preventing bias from attrition

The third point at which biases can occur is when moving from the selected sample to the achieved sample. This is because research can be affected by low response and attrition, where users are hard to reach, unresponsive or withdraw part way through. The issue of persuading people to take part in research is addressed more fully in other guidance[4].

Checking and correcting for bias

No matter how good your sampling approach you should always do what you can to check for bias. Remember, the aim is to ensure your sample is as representative of the population as possible, so you should take whatever you know about the population and compare it to the sample. For example, if you know that half your service users are men and half are women, then you will want the same proportion in your sample. Similarly, if you know that half your service users worked with you for three months and half for six months this should also be reflected.

The more you go through the process of checking your sample against the population (and report this), the more robust your research will appear.

If you do identify a bias you may decide to correct this by doing more interviews with groups who are under-represented. An alternative approach – applicable to quantitative studies only –

[4] http://www.clinks.org/sites/default/files/AcheivingUserParticpationResearch.pdf

is to weight the data. This is a statistical adjustment of the data that corrects biases. At its simplest it's quite easy to do; for instance if the split between men and women in your population is 50/50 but in your sample it is 75/25, then you could correct this by giving men in the sample a weight of 0.67 and women a weight of 2 to bring it back to 50/50. The drawbacks are that; technically, it reduces statistical confidence in the findings; you will need more advanced software like SPSS to do it; and weighting by more than two variables starts to get very complicated. If you are interested in the topic then you could read more information here:

http://help.pop.psu.edu/help-by-statistical-method/weighting/Introduction%20to%20survey%20weights%20pri%20version.ppt

Sample size and confidence levels

An appropriate sample size differs depending on whether you are conducting qualitative or quantitative research.

Quantitative research

In quantitative research, the larger the sample the more confident you can be that your findings are not down to chance and reflect the overall population. This is expressed statistically by margins of error (also known as confidence levels), which help you understand how sure you can be that the result of research with a sample of your population accurately reflects what the result would be if you had interviewed everyone.

To calculate a confidence level you need to use quite complex maths but luckily there are online calculators that can do this for you (see below). Based on the size of your sample and the proportion of the population interviewed they calculate the parameters within which you can be confident that you results are accurate.

So by way of illustration, if you interview 10 people out of a population of 100 service users the confidence level calculated is 30 percentage points. This means that if you found that 50% of the sample said they were satisfied with your service then the real figure for the whole population could be 30 percentage points lower or higher than the result from your sample, so in this case between 20% and 80%. These parameters are very wide because the margin of error in interviewing a small sample is much higher. However if you interview a sample of 80 out of 100 service users the confidence interval would be 5 percentage points meaning the true result is likely to be anywhere in between 5 percentage points lower or higher than the result you found in your sample (and therefore much more precise). The key point here is the more people you interview, and the higher the proportion of the population the more accurate your findings will be.

Finally it should also be noted that very small samples are always treated with caution. Generally any quantitative sample below about 30 should only be seen as indicative and not suitable for statistical analysis.

A good introduction to this topic can be found here:

https://www.whatisasurvey.info/chapters/chapter10.htm

There are a number of online calculators that will help you to calculated margins of error easier. For example:

http://ncalculators.com/statistics/confidence-interval-calculator.htm

http://www.surveysystem.com/sscalc.htm

Qualitative research

Statistical calculations are not applicable for qualitative research, which typically entails smaller sample sizes. Rather qualitative researchers talk about getting to a 'saturation point' – that is when repeated interviews result in the same themes and findings. This can mean you need only conduct as little as 10/15 interviews before you have enough findings to meet your objectives.

Whilst qualitative research is not attempting to be representative, it should still ensure that a good mix of the research population is included in the study. Any tendency towards being selective may mean the sample becomes biased towards those with certain experiences, such as those more likely to engage with the service. In general, the key messages in this paper about trying to select randomly and eliminate bias are equally applicable to qualitative research as they are to quantitative.

Conclusion

Perhaps the key message of this guidance is that research should not rely on speaking to those who are easiest to engage as this produces biased results. Rather, if any sampling is involved, then it should be based on a systematic approach to selecting people and careful comparison of the sample against the population you want to generalise about.

Sample size is important too, but it is perhaps not as important as ensuring the representativeness of the sample. This applies as much to qualitative research as it does to quantitative.

Finally, you should explicitly address the issue of bias when reporting your findings. Being upfront about concerns or limitations, and how confident you are in your findings is an important part of appearing credible.

Links to further sources

http://betterevaluation.org/plan/describe/sample

http://www.socialresearchmethods.net/kb/sampterm.php

http://qualtrics.com/wp-content/uploads/2013/05/Sampling.pdf