The Snowball Sampling Strategy in the Field of Social Sciences. Contexts and Considerations

http://dx.doi.org/10.12775/PBE.2023.029

Abstract
This article discusses the snowball sampling strategy that has been successfully used for decades in research in the field of social sciences. The focus of this article is only on the applica-
tion of the strategy in qualitative research, although it can be employed in quantitative research too. Despite its limitations and debatable applicability in some cases, it is methodologically justified and considered very effective in reaching hard-to-reach populations. Referring to the methodological literature, research in which it was adopted and our own experience, we reflect on the contexts and controversies that arise around its implementation.

**Keywords:** snowball sampling strategy, social sciences, qualitative research, methodology.

**Introduction**

> I would suggest that sampling strategies involving people [...] are more akin to opening a Pandora's box.
> (Curtis et al., 2000, p. 1008)

One of the main challenges when conducting research in the interpretive paradigm is the choice of a specific procedure, selecting respondents and their number. Especially if our explorations require “hidden,” “vulnerable” or “elite” groups to participate in the study, which is a challenge that educators, sociologists, political scientists and physicians usually face in their research projects. The complexity and difficulty in reaching these groups has resulted in the emergence of a method known as *snowball sampling*, which is considered a highly effective respondent selection technique, especially for the hard-to-reach population (Atkinson & Flint, 2001). For years, research on sensitive topics has used the social networks of individuals to gain access to both hard-to-reach and vulnerable populations (Becker 1963; Lindesmith, 1968; Vervaeke et al., 2007; Waters, 2015).

The snowball method has been used for many decades and is believed to have been invented by Coleman (1958) and Goodman, who studied the structure of social networks (Heckathorn, 2011). One of the earliest examples of its application has been used as exemplified by the classic *Outsiders* by Howard Becker (1963), who, using this method, reached the hard-to-reach community of “deviants,” which were people addicted to marijuana. A characteristic feature of this method when implemented in qualitative research is that its purpose is not to estimate the characteristics of the general popu-
lation, but to estimate the characteristics of the hidden population networks (Dragan & Isaic-Maniu, 2013).

The snowball method is most frequently used in ethnography, action research or individual case studies. It is classified as non-probability sampling, in which the research group is specific, and the research is conducted on a small scale. This method is economical from both a time and a financial perspective (Barbour & Schostak, 1999, p. 61). The snowball technique, also called a chain sampling, is based on the principle of graduated selection with a theoretical sampling, “in which the researcher examines individual cases of the phenomenon of interest in order to be able to define and elaborate on its various manifestations. The researcher samples individuals, institutions, documents, or wherever the theory leads the study” (Teddlie & Yu, 2007, p. 82). This technique is described by Marshall as “a more intellectual strategy than simple demographic stratification” (Marshall, 1996, p. 523). He believes that in this fashion the principle of maximum variability should be invoked, including deviant individuals, those with specific experiences (a sample of critical cases) and also key informants. Each of these groups can recommend further individuals for study (Marshall, 1996). There is an abundance of research which focuses on quantitative snowball approaches, which include representative quota sampling and representative purposive sampling, network sampling and respondent-driven sampling, each underpinned by statistical inference. Moreover, there is a plentiful pool of quantitative research which uses non-probability survey snowball sampling. This article, however, is focused on the use of snowball sampling for qualitative research.

Qualitative sampling strategies: putting the snowball in context

Before arriving at the decision to use snowball sampling, qualitative researchers must be aware of the broader context within which sampling strategies and practices are located. At a basic level, as Uwe Flick notes in the classic ‘Introduction to Qualitative Research’: “The general issue of sampling is how to select cases or examples from a wider population so that the research in the end can make statements that apply not just to the individual participant(s) of a study” (Flick, 2018, p. 173). The methods of obtaining a research sample – depending on the scientific field – are well described and categorized
in the methodological literature. Some are based on random selection and probability, some on purposive selection. It is not always possible to use traditional probabilistic methods (Griffiths et al., 1993; Johnston, 2014) to obtain a research sample, i.e., to use or build an available sampling frame, because it may be impractical or even impossible.

Sampling strategies are usually designed to take us beyond the unique individual case, but wider relevance manifests itself in different ways. Most obviously, in qualitative research, statements beyond the individual are more likely to be illustrative than representative (the latter being a gold standard in quantitative research).

There are three basic sampling steps in qualitative interviewing: 1) determining the target population; 2) determining the sample size from this target population; and 3) devising and operationalising a sampling strategy. In this section, we will now discuss these three steps in turn.

As far as the target population is concerned, qualitative research is more often than not focused on population sub-sets rather than the general population. In other words, participants are selected because they meet specific criteria (a purposive/stratified/quota sample), based on a theoretical rationale or research aim/objective, but there is no claim that those selected in this way represent a wider population or sub-population.

The second qualitative sampling step is determining the sample size. Taking a grounded theory approach (Glaser & Strauss, 1967), as is popular in qualitative research, the sample size will most likely be flexible as researchers make real-time judgements relating to the emergence and evolution of theory. In such a context, sample size will grow until theoretical saturation i.e. when new themes no longer emerge from new respondents a ‘saturation’ point will be reached and research will end. This is not the end of the story, however, in that some qualitative researchers find that corroboration can be useful and so the repeated emergence of similar themes can be an objective: and this can take sample sizes well beyond a theoretical saturation point.

Key decisions revolve around the level of homogeneity or heterogeneity across a target population and, related to this, determining the characteristics justifying inclusion or exclusion in the target population. This can be quite difficult. Not least, there may be a need for more respondents if the snowball becomes heavily skewed. For example, in cases where there is considerable
homogeneity, sample sizes will generally be smaller at the sample saturation point. However, when there is heterogeneity, sample sizes will most likely need to be larger in order to achieve sampling saturation. Drawing on insights from a homogenous sample, Guest et al. found that: “For most research enterprises […] in which the aim is to understand common perceptions and experiences among a group of relatively homogeneous individuals, twelve interviews should suffice” (2006, p. 79).

In some instances, snowball sampling can help to develop sample homogeneity, i.e. research investigating young low-wage migrants in London can introduce more young low-wage migrants in London into the research. This would be a classic snowball sampling strategy which underpins sample homogeneity, a major strength in snowball sampling. However, this can also be a weakness, because homogeneity may become skewed, especially if initial entry points are limited. Thus, whilst sample homogeneity may have been achieved, not all areas of a target population may have been reached and represented after snowball sampling. For instance, a young migrant working in a London hotel may give referrals only to his colleagues rather than a more wider group of young low-wage migrants across the London economy. Therefore, the sample will end up becoming significantly skewed and, even if representativeness was not an aim, there could still be issues associated with the sample population being too homogenous. As Chambers, Bliss & Rambur (2020, p. 847) state, this would become a ‘narrow network of acquaintances.’

A sample size of 12 is more modest than most qualitative research projects and there is certainly a sense in which there is a qualitative orthodoxy in terms of number of interviews. In our own research areas, there is relatively little informed sample size guidance apart from averaging. Saunders & Townsend (2016), to this end, found an average number of 33 interviewees across worker/workplace studies in general. This chimes with the ‘rule of thumb’ academic orthodoxy: of around 6–10 interviews for undergraduate dissertations; 10–20 interviews for Masters dissertations; and 30–50 interviews for PhD theses. The latter figure is also commonly used for academic research projects as well. Oliver Robinson (2014, p. 28) very usefully here identifies five types of sample homogeneity (Tab. 1).
Table 1. Five types of sample homogeneity

<table>
<thead>
<tr>
<th>Types of sample homogeneity</th>
<th>Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic</td>
<td>age, gender, ethnicity, socio-economic group</td>
</tr>
<tr>
<td>Geographic</td>
<td>sample rooted in particular locations</td>
</tr>
<tr>
<td>Physical</td>
<td>sharing a common physical characteristic such as an illness</td>
</tr>
<tr>
<td>Psychological</td>
<td>sharing a common psychological trait such as high IQ</td>
</tr>
<tr>
<td>Life History</td>
<td>sharing a past life experience such as migration when young</td>
</tr>
</tbody>
</table>

Source: Own elaboration.

Atkinson & Flint (2001, p. 1) distinguish between the formal and informal role of the snowball sampling strategy as a research method (Tab. 2).

Table 2. Formal and informal role of snowball sampling method

<table>
<thead>
<tr>
<th>Primary purposes</th>
<th>If the aim of a study is:</th>
<th>Used</th>
<th>Excellent tool is:</th>
</tr>
</thead>
<tbody>
<tr>
<td>FORMAL</td>
<td>Formalized and statistical sense when random sampling is not possible.</td>
<td>Quantitative Verification strategies</td>
<td>In accessing and interviewing control groups to further bolster the validity of the research findings</td>
</tr>
<tr>
<td>INFORMAL</td>
<td>Reaching a target population, by creating contacts with a respondent’s circle of acquaintances</td>
<td>Primarily explorative Qualitative Descriptive</td>
<td>For Interviews</td>
</tr>
</tbody>
</table>

As stated previously the third step in qualitative sampling is devising and operationalising a sampling strategy. As stated previously the ‘gold standard’ amongst quantitative researchers is random/ probabilistic sampling and it is possible for qualitative interview samples to be random, systematic-random (e.g. every nth person or street) or stratified-random (i.e. a random sample of a sub-population).

Representativeness through random sampling is not, however, the main driver behind most qualitative interview research. Whilst interview data may be generalisable to a population, or population sub-set, it may not be and, instead, emphasis is more on data that illustrates a particular theoretically-informed theme.

Within non-probability purposive/ stratified sampling there is thus a considerable degree of convenience or opportunism. Once a target population has been identified it becomes those who are most available and willing who are sampled rather than a random selection of people where every member of a target population has an equal chance of selection. This is why claims over representativeness are rare in qualitative interview research. It is in the sphere of purposive/ stratified/ quota sampling, based around convenience and opportunism, that snowball sampling comes into its own.

Snowball sampling: characteristics and types, strengths and weaknesses

Having outlined the three key sampling steps for qualitative interviewing, we will now turn particular attention to snowball sampling. The method of acquiring respondents using the snowball sampling strategy is well documented in methodological literature (Biernacki & Waldorf, 1981; Berg, 1988; Gile & Handcock, 2001; Browne, 2005; Noy, 2008; Isaic-Maniu, 2013; Naderifar et al., 2017; Geddes et al., 2018; Parker et al., 2019; Cantone & Tomaselli, 2022) and successfully applied in social sciences. This method is used both in qualitative research to gain access to potential interviewees h to search for survey participants, although in quantitative research it is less common due to the demand for large populations (Cohen & Arieli, 2011, p. 427).

Snowball sampling is one of the most popular methods of sampling in qualitative research, central to which are the characteristics of networking.
and chain referral. Despite being a frequently used and commended sampling strategy, due to “snowball” sampling [...] [being] a profitable means of recruiting research participants, [...] at the same time it tends to be profiled in a rather limited and superficial manner” within journal articles and research methods text books (Geddes et al., 2018, p. 347). A more detailed explanation of such a desirable sampling technique is thus required.

At its simplest, a snowball sample is where the researchers start with a small number of initial contacts (seeds) who fit the research criteria and who are invited to become participants within the research. Those who agree to become research participants are then asked to recommend other contacts who fit the criteria. Of those who agree, those participants then in turn recommend other potential participants, and so on. Researchers, therefore, use social networking to establish initial links and gather recruitment momentum, capturing an increasing pool of willing research participants in a chain-like fashion. When one starts with a few entry points and develops recruitment momentum from these entry points, we have genuine snowball sampling.

Many quantitative researchers describe the snowball method as marginal since the results obtained cannot be generalized (Atkinson & Flint, 2001). However, it is worth emphasizing – especially in the qualitative research paradigm – that there are important topics that need to be investigated, and this opportunity only appears if we use the snowball method (King & Keohane, 1994, p. 6). Furthermore, this method allows us to obtain unique data (Noy, 2009). Mark Handcock and Krista J. Gile (2011) point out that snowball sampling has evolved over the years and been used inconsistently in various scientific fields. In the literature on the subject, there are several most frequently quoted and consistent definitions:

Pooja Bhardwaj (2019, p. 162) distinguishes three types (Tab. 3). She uses medical examples to learn about the quality of life of people with a given disease in the school environment, but these examples are appropriate as part of health pedagogy.
Table 3. Types of snowball sampling method

<table>
<thead>
<tr>
<th>Types</th>
<th>In these types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear snowball</td>
<td>The collection of samples starts from collecting data from one and then that individual tells about the other and so in this way, a chain is formed and it continues till we get enough number of individuals to analyse.</td>
</tr>
<tr>
<td></td>
<td>A researcher conducting research on Crohn’s disease needs to source the people suffering from Crohn’s disease which can be problematic, so he/she asks one patient and obtains information about the other patient suffering from Crohn’s disease.</td>
</tr>
<tr>
<td>Exponential</td>
<td>In this type of snowball sampling, one patient gives multiple referrals, but the recruitment will be done only for one patient on the basis of the nature and type of the research study.</td>
</tr>
<tr>
<td>Exponential</td>
<td>In this, one individual will be giving information about more than one individual and those individuals in turn will be giving information about the others and in this way, with more and more referrals, the chain is formed, and we collect data.</td>
</tr>
<tr>
<td></td>
<td>To obtain data concerning Diabetic mellitus a researcher finds an individual who is suffering from this disease. There is a high probability that a researcher will get some information about other people he/she may know suffering from Diabetic mellitus.</td>
</tr>
</tbody>
</table>


One problem with using a snowball sample is when the snowball fails to roll. In other words, when recruitment does not gain momentum and new participants are not recruited. This could be due to a lack of recommendations or a lack of willing participants. With some research topics the research community can be characteristically tight-knit and closed. With some research topics, the focus of the research is considered too sensitive and risky, and/or potential participants may fear exposure. Waters (2015) identified such limitations within her research article on older adult drug users and reminds that a network has to exist in the first place for a snowball to gain momentum. Waters also highlights how insider status may assist with some research topics, in the gaining of momentum.

If the snowball fails to roll – that is, networking does not gain momentum – then there may be adaptations the researcher can consider. As with the research of Woodley and Lockard (2016), rather than relying on strong
or tight-knit social ties for securing interviewee recruitment, the researcher could network using weaker connections to less familiar acquaintances and other opportunistic interactions. Researchers could also look for a multitude of entry points. In these instances, it is more appropriate to claim horizontal network sampling rather than the vertical chain-referral snowball sampling (Geddes et al., 2018).

Sampling is usually considered complete once either a target sample size or saturation point has been reached. The issue, of course, is that if momentum has proved arduous, where multiple entry points have been used, and weaker ties drawn upon, then the sample may have lost focus and/or become skewed. In other words, the sample may not align with the original intentions of the research and the original target population may not have been captured. (should we explicitly mention a lack of validity here?). Nonetheless, researchers must remain diligent and ensure that recommended participants actually fit the research criteria.

As will be discussed later within this paper, snowball samples are renowned within the academic literature for capturing hard-to-reach populations and awarding a voice for underground, marginalised and repressed voices. On the other hand, the key dimensions of snowballing i.e. social networking, power relations and drawing upon social capital should be celebrated (Noy, 2008). Snowballing is thus a sampling technique which draws upon the power of social networking, for counter-narratives to be told (Woodley & Lockard, 2016). Nonetheless, it is importance to note the rarely documented value of snowball sampling for also accessing the everyday, mundane, and mainstream (Geddes et al., 2018; Parker et al., 2019). Parker points this Johnson (1997, pp. 283–284):

One potential threat to validity that researchers must be careful to watch out for is called researcher bias. This problem is summed up in a statement a colleague of mine once made to me. She said “The problem with qualitative research is that the researchers find what they want to find, and then they write up their results.” It is true that the problem of researcher bias is frequently an issue because qualitative research is open ended and less structured than quantitative research. This is because qualitative research tends to be exploratory. (One would be remiss, however, to think that researcher bias is never a problem in quantitative research!)
Researcher bias tends to result from selective observation and selective recording of information, and also from allowing one's personal views and perspectives to affect how data are interpreted and how the research is conducted.

It is also important to note that, in studies based on deliberate sampling, the interpretation of results is limited to the cases studied and generalizations are not valid. This is the “internal validity” of research results (Tongco, 2007). The authors are fully aware that qualitative studies, unlike quantitative studies, are considered “soft, descriptive, ‘feminine,’ ‘microscopic,’ less exact/precise/objective/rigorous, systematic […] and non-scientific” by many (Vissak, 2010, p. 378).

Despite the fact that snowball sampling may be the only method possible, it can pose many methodological challenges as Jaime Waters (2015, p. 372) notes. Referring to a number of researchers, she firstly indicates that this method depends on the researcher, i.e., their resources and contacts (Griffiths et al., 1993) and also their experience. Also, it may involve distortions caused by the researcher’s influence, as well as their negative attitude towards a respondent, despite the fact that the latter meets the requirements for the inclusion in the study. Snowball sampling thus faces some criticisms. As a network-based convenience form of sampling, it may be viewed negatively for not producing samples that meet the criteria of random sampling in the statistical sense (i.e., snowball samples depart from probability-based sampling approaches). In other words, the dominant characteristic of the snowball sample (i.e., the referral process) is dependent on a selection bias. Moreover, the basis for establishing the representativeness of samples may also be questioned. Overall, snowball sampling is criticized for its selection bias as well as a lack of external validity, generalisability, and representativeness and at a philosophical level, researchers who have a nomothetic aim (i.e. to develop approaches that enable generalisation) would either reject snowball sampling altogether or turn to the statistical representative quota sampling and representative purposive sampling.

Criticisms of the snowball technique draw attention to a so-called primary sampling, in which there is a danger of over-representing a single group (Harrell & Bradley, 2009, p. 32). In addition, as Amanda Wilmot points out, a research sample composed of people who know each other may not only
have similar views and experiences but also influence each other with regard to answering interview questions. Hence, she advocates a rigorous screening process for respondents (Wilmot, 2005, p. 224). It also entails sourcing people who are willing to talk, extroverts, and leaving out introverts who are reluctant to share experiences (Simkus, 2022). Hildenbrand, on the other hand, notes that when sourcing introverts, there is the problem of the researcher as an ‘acquaintance’ who, unlike a stranger, may find it difficult to gain interesting and new information: “the stranger the field, the more easily may researchers appear as strangers, whom the people in the study have something to tell which is new for the researcher” (Hildenbrand, 1995, p. 258). In this context, David L. Morgan notes that many of these risks can be avoided by starting a study with a set of several respondent-informants with a high degree of diversity, which ‘increases the likelihood that subsequent links in the snowball process will reach different segments of the total set of eligible participants’ (Morgan, 2008, p. 816).

Irrespective of the nomothetic/ideographic starting positions, the homogeneity/heterogeneity of the target population, sample size averages, and the degree to which one stops at the saturation point or goes further to corroborate, there is an additional set of pragmatics to consider when determining sample size. Specifically: the money available; the time available; the energy/motivation the researcher(s) has left; and the accessibility to and willingness of the target population are all important considerations. These may all check the final number of interviewees and may even mean a saturation point is never reached, or the sample falls below acknowledged averages. Depending upon one’s philosophical position, adherence to qualitative orthodoxy, and/or the funding promises made, one may or may not be troubled by such pragmatic limitations.

**Hard-to-reach populations**

The methods of obtaining a research sample – depending on the scientific field – are well described and categorized in the methodological literature. Some are based on random selection and probability, some on purposive selection. It is not always possible to use traditional probabilistic methods (Griffiths et al., 1993; Johnston, 2014) to obtain a research sample, i.e., to use or
build an available sampling frame, because it may be impractical or even impossible.

One of such limitations is the so-called hard-to-reach populations, which are defined as communities whose members may be reluctant to self-identify (Johnston, 2014). Such groups (depending on the cultural context) include: non-heteronormative people (Browne, 2015), addicted people (Shewan & Dalgarno, 2005; Waters, 2015), HIV-positive people (Dowsett, 2001), migrants (Johnston, 2014) and people with risk-taking behaviour (Thompson & Collins, 2002), criminals (Fitzgerald, 1996), seriously ill and people with rare diseases (Sudman & Freeman, 1988), homeless people, members of elite clubs, as well as people who provide sex services, rape victims, sect members, hackers (Felix-Medina & Thompson, 2004) or victims of natural disasters (Malilay et al., 1996). Therefore, it can be noted that recruiting these individuals for research is often very difficult due to both the lack of official demographic data and also because exposure can cause health and life risks and negative consequences for friends and families (Farquhar, 1999). (e.g., non-heteronormative individuals in countries where transgressing the dominant codes of heterosexuality is criminalized – i.a., Guinea, Tunisia, Barbados, United Arab Emirates, Kenya, Malesia) (Farquhar, 1999).

Sampling via the snowball method proves ideal when it is problematic to reach so-called “vulnerable groups.” In a study conducted in 2022 among female war refugees from Ukraine, Alekandra Boroń and Agnieszka Gromkowska-Melosik (2022) used a purposive sampling method using the snowball technique. The first stage was to contact the principals of the schools to which the war refugee children were referred, asking them to provide contact details with their mothers. In the second stage, centres hosting refugee women were contacted. In addition, female translators were included in the study, acting as insiders to recruit people for the research. Each refugee woman who went through the interview experience recommended another with whom she usually had telephone contact or lived or worked together. So the first contact was the school’s headmaster, then the Ukrainian teacher hired to look after the newly arrived children, this one in turn giving contact to one of the mothers. This to the next and so on. In one case, the first contact was a priest who had custody of the monastery buildings, part of which were made available to the refugee women. The refugee women’s children also participated in
the project by preparing artwork for a competition. During the meeting, they received toys and sweets. All individuals received written information about the purpose and conduct of the study, hence they could decide to participate. An additional element that strengthened refugee women’s sense of security and confidence was the inclusion of their children in the ‘Great Power in Ordinary Words’ project aimed at community integration and intercultural intervention. The cross-cultural nature of the project involved in-depth, spatio-temporal interaction between Ukrainian children and their mothers along with teachers and Polish schools as well as Polish university researchers.

This interaction took place within the space of literature and art during a face-to-face meeting, with Ukrainian children as the main protagonists. A ‘sudden community’ which was created, although temporary undoubtedly remained in the participants’ memory and was essential both for the children and the Polish and Ukrainian adults who participated in the meeting. The meeting space became an intercultural performative ‘situational site,’ temporary but significant from the perspective of the Ukrainian children’s experience. It seems to have constituted a time of momentary ‘exclusion from consciousness’ of tragic events, the pleasure of the moment, the comfort of the meeting, the feeling of ‘being there’ and ‘being important’ – all this gave the meeting a context of ‘intercultural interventionism’ – the main theme of the whole project.

This allowed the researchers to obtain a sample from all districts of Ukraine, with different status, education and length of stay in Poland. Simultaneously, this method of recruiting people for the interview allowed migrants to gain trust and a sense of security. This was despite the sense of destruction they felt at their entrance into a new culture, one which occurred due to the destruction of their homeland. Sometimes they feel disoriented in the new social situation, experience a loss of status, a sense of being in the world, as well as a variety of attitudes of people from the host country – from acceptance to discrimination (Summerfield, 2000, p. 420).

Nonetheless, researchers who have an ideographic aim (to give substantial voice to individual participants and allow an intensive analysis of each case), where generalisation, representativeness, and external validity are not sought after, snowball sampling is thus frequently advocated and employed by qualitative social researchers (especially interviewers and ethnographers)
as a form of non-random sampling. Thus, the case of say low-wage migrant workers in only one London hotel would not be seen as a problem. Indeed, and this sometimes occurs, less can be considered more in terms of sample size and a single interviewee may be deemed sufficient for the purposes of some research investigations.

Conclusions

The authors are aware that the snowball method described in this article is not comprehensive. Our main focus was directed towards a quality strategy. However, we believe that this description will serve as a starting point for students and researchers using snowball procedures. Those who employ this method must realize that it has its supporters and opponents, its strengths and weaknesses. However, according to Geoff Walsham’s statement, “there are no ‘correct’ and ‘incorrect’ theories [methods].” – it is the researcher who should be able to make a judgment as to whether a given theory [method] deserves interest and thus whether its continued use in an ongoing research exploration is justified (Walsham, 1995b, p. 295).

References


