



social research UPDATE

• In its simplest formulation snowball sampling consists of identifying respondents who are then used to refer researchers on to other respondents.

• Snowball sampling contradicts many of the assumptions underpinning conventional notions of sampling but has a number of advantages for sampling

populations such as the deprived, the socially stigmatised and elites.

• Snowball sampling has advanced as a technique and the literature contains evidence of a trend toward more sophisticated methods of sampling frame and error estimation.

• Although they violate the principles of sampling, the use of snowball strategies provides a means of accessing vulnerable and more impenetrable social groupings. However, the nature of similarity within social networks may mean that 'isolates' are ignored.

Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies

Rowland Atkinson and John Flint

Treading an uneasy line between the dictates of replicable and representative research design and the more flowing and theoretically led sampling techniques of qualitative research, snowball sampling lies somewhat at the margins of research practice. However, the technique offers real benefits for studies which seek to access difficult to reach or hidden populations. These are often obscured from the view of social researchers and policy makers who are keen to obtain evidence of the experiences of some of the more marginal excluded groups.

Policy makers and academics have long been aware that certain 'hidden' populations, such as the young, male and unemployed, are often hard to locate. Other groups such as criminals, prostitutes, drug users and people with unusual or stigmatised conditions (e.g. AIDS sufferers) pose a range of methodological challenges if we are to understand more about their lives. This issue describes the processes and the advantages and difficulties of using snowball sampling techniques.

Snowball Sampling

Snowball sampling may simply be defined as:

A technique for finding research subjects. One subject gives the researcher the name of another subject, who in turn provides the name of a third, and so on (Vogt, 1999).

This strategy can be viewed as a response to overcoming the problems associated with sampling concealed populations such as the

criminal and the isolated (Faugier and Sargeant, 1997). Snowball sampling can be placed within a wider set of link-tracing methodologies (Spren, 1992) which seek to take advantage of the social networks of identified respondents to provide a researcher with an ever-expanding set of potential contacts (Thomson, 1997). This process is based on the assumption that a 'bond' or 'link' exists between the initial sample and others in the same target population, allowing a series of referrals to be made within a circle of acquaintance (Berg, 1988).

Snowball sampling can be applied for two primary purposes. Firstly, and most easily, as an 'informal' method to reach a target population. If the aim of a study is primarily explorative, qualitative and descriptive, then snowball sampling offers practical advantages (Hendricks, Blanken and Adriaans, 1992). Snowball sampling is used most frequently to conduct qualitative research, primarily through interviews. Secondly, snowball sampling may be applied as a more formal methodology for making inferences about a population of individuals who have been difficult to enumerate through the use of descending methods such as household surveys (Snijders, 1992; Faugier and Sergeant, 1997).

Early examples of the technique can be seen in Whyte's *Street Corner Society* (1955) and Patrick's study of a Glasgow gang (1973), which used initial contacts to generate contexts and encounters that they could use to study the gang dynamic. A general move

Rowland Atkinson and John Flint are researchers at the Department of Urban Studies, University of Glasgow. Both have an interest in the spatial distribution and experience of social exclusion and have been commissioned to devise a methodology for tracing residents who leave regeneration areas in Scotland.

social research UPDATE

away from participant observation of this kind towards the use of snowball sampling techniques primarily for interview-based research has been seen more recently. Snowball sampling has been used in studies of drug users (Avico et al, 1998, Griffiths et al, 1993; Kaplan et al, 1987); prostitution (McNamara, 1994); pickpockets (Inciardi, 1977); aids sufferers (Pollak and Schlitz, 1988) and the seriously ill (Sudman and Freeman, 1988).

While some may seek to characterise the topics for which snowball strategies have been used as being trivial or obscure, the main value of snowball sampling is as a method for obtaining respondents where they are few in number or where some degree of trust is required to initiate contact. Under these circumstances, techniques of 'chain referral' may imbue the researcher with characteristics associated with being an insider or group member and this can aid entry to settings where conventional approaches find difficult to succeed.

A range of advantages have been claimed for snowball sampling. Firstly, it has enabled access to previously hidden populations. Often members of such populations may be involved in activities that are considered deviant, such as drug taking, or they may be vulnerable, such as the stigmatised in society, making them reluctant to take part in more formalised studies using traditional research methods. Trust may be developed as referrals are made by acquaintances or peers rather than other more formal methods of identification. Snowball sampling has been found to be economical, efficient and effective in various studies. For example, it has been shown to be capable of producing internationally comparable data in Avico et al's 1988 study of cocaine users in three European cities. It may be used to examine changes over time. Snowball sampling can also produce in-depth results and can produce these relatively quickly.

Perhaps one of the strongest recommendations for the snowball strategy stems from a distinction between descending and ascending methodologies (Van Meter, 1990). Traditional techniques such as household surveys, as 'descending' strategies, are associ-

ated with a largely quantitative tradition of the measurement of social problems that often suffer from a lack of responses from particular groups, often the young and unemployed males. Ascending methodologies, such as the use of snowball techniques, can be used to work upwards and locate those on the ground who are needed to fill in the gaps in our knowledge on a variety of social contexts. In this sense snowball sampling can be considered as an alternative or as a complementary strategy for attaining more comprehensive data on a particular research question.

While many have considered snowball strategies primarily as an aid to accessing the vulnerable or the deviant, some studies have used them to engage with the 'hard to reach' among urban elites. Saunders' (1973) study of urban politics is an example where a 'reputational' method was used. Respondents were asked who held power in the local arena. This led to a series of contacts and the establishment of a subjective indication of the relative local power bases. This suggests that snowball sampling has a wider applicability in sociological research than has hitherto been realised.

Difficulties

Snowball samples have a number of deficiencies.

1. Problems of representativeness and sampling principles

The quality of the data and in particular a selection bias which limits the validity of the sample are the primary concerns of recent snowball sampling research (Van Meter, 1990; Kaplan et al, 1987). Because elements are not randomly drawn, but are dependent on the subjective choices of the respondents first accessed, most snowball samples are biased and do not therefore allow researchers to make claims to generality from a particular sample (Griffiths et al, 1993). Secondly, snowball samples will be biased towards the inclusion of individuals with inter-relationships, and therefore will over-emphasise cohesiveness in social networks (Griffiths et al, 1993) and will miss 'isolates' who are not connected to any network that the researcher has tapped into (Van Meter, 1990).

The problem of selection bias may be partially addressed, firstly through the generation of large sample and secondly by the replication of results to strengthen any generalisations. At present, a statistical formalisation of snowball sample biases is not available (Van Meter, 1990). However, larger sample sizes may reduce bias. For example Pollak and Schlitz's study of AIDS sufferers produced a sample with representative proportions for age, class and size of town of residence (Pollak and Schlitz, 1988).

In addition to selection bias, there is also the issue of gatekeeper bias (see Groger, Mayberry and Straker, 1999). In their work they identified a difficulty when using nursing home staff as 'go-betweens' in obtaining the informed consent of caregivers. These 'gatekeepers' were sometimes reticent or protective toward those they cared for and sometimes hindered access for the researchers. Based on their experiences they attempt to draw a comparison between snowball sampling and scrounging sampling. They describe the latter in terms of:

desperate and continuing efforts, against mounting odds, to round out the collection of individuals with relevant types of experiences we know to exist but have not been able to capture (Groger et al, 1999:830)

While social scientists may vary in the degree to which they accept such a viewpoint we can recognise similar traits in quantitative approaches such as household surveys. Here we often find struggles to obtain adequate numbers of respondents by making repeated visits to minimise biases derived from low response rates.

The ideal number of links in a referral chain will vary depending on the purpose of the study. More links in each chain will generate substantial data about a particular sample, and may also allow access to those most difficult to identify (e.g. those respondents who require the greatest level of trust to be built up before participating). However, it is also more likely that members of such a large single chain sample will share similar and unique characteristics not shared by the wider population. Thus, there may be a case for initiating several discrete chains with fewer links, particularly where any inference

social research UPDATE

about a wider hidden population is considered important.

Attempts have been made to consider the statistical accuracy of samples obtained by snowball techniques using 'mark-recapture' techniques (Fisher et al, 1994; Shaw et al, 1996) to estimate homeless populations. Mark-recapture takes its name from the techniques used to estimate numbers of natural wild populations. A sample of animals is captured from the population and released, a second sample is taken and counted and the number of recaptures noted. Shaw et al (1996) give the example of eighty animals being marked and then, in a second sample of fifty animals, ten being found with marks. The overlap of ten animals creates a multiplier of 5 (50/10) leading to an estimate of 400 (5 times 80) for the size of the population.

In comparing registers, such as service agency data and police arrest data, the ratio of total agency population to the number of overlaps can similarly be used to estimate the size of an unknown population. However, while such techniques help us to test the relative robustness of a sample, the majority of snowball sampling applications have been used to look at vulnerable and hidden experiences which have been ignored by descending research methodologies.

2. Finding respondents and initiating 'chain referral'

By their very nature, members of a hidden population are difficult to locate. Often studies require some previous 'knowledge of insiders' in order to identify initial respondents. Such prior knowledge may not be readily available to researchers and it may be very time consuming and labour intensive to acquire. Under these circumstances it is possible that people in positions of relative authority or proximity may provide a route into the required population (Groger et al, 1999). For example, housing officers might be able to introduce a researcher to a tenant on incapacity benefit if the tenant agrees to be identified. It should be stressed that there are clear ethical implications for such work and that informed consent should be considered a prerequisite.

Criteria for membership of a sample will depend on the nature of the research question being posed. In the case of deviant activities such as drug taking it may be that some referrals will not be accurate. Secondly, referrals will largely depend on the subjective perceptions of initial respondents about the involvement of others in the same activity. Thus particular individuals (the most popular, long-term residents or those with wider social networks) are more likely to be identified than others. Much of snowball sampling rests on the assumption that social networks consist of groups with relatively homogenous social traits. However, there are limits to this and it largely depends on what characteristics are considered to be the most important. In the case of a particular disability, for example, it is assumed that someone in this group would know others to whom a researcher could be directed. However, some groups may themselves consist of highly atomised and isolated individuals whose social networks are relatively impaired. Young unemployed men have been viewed as a prime case in this respect and may therefore be very difficult to identify or to initiate contact with. It is therefore apparent that snowball samples are both time consuming and labour intensive (Griffiths et al, 1993; Faugier and Sergeant, 1997).

3. Engaging respondents as informal research assistants

Researchers may encounter initial hostility and suspicion from targeted individuals. There is also evidence of research fatigue, particularly amongst marginalised groups who have been subject to previous research (Moore, 1996). Establishing the trust of respondents is therefore essential. Often trust can only be built up slowly as the purposes of the study and the consequences for respondents of taking part become clearer as the study develops (Berg, 1988; Faugier and Sergeant, 1997). In both an ethical and a practical sense, respondents need to be reassured of the protection of the information they provide. This assurance of confidentiality can only be demonstrated over time. These points suggest that the initial respondents may act as invaluable assistants in obtaining

social research UPDATE is distributed without charge on request to social researchers in the United Kingdom by the Department of Sociology at the University of Surrey as part of its commitment to supporting social research training and development.

Contributions to **social research UPDATE** that review current issues in social research and methodology in about 2,500 words are welcome. All UPDATE articles are peer-reviewed.

social research UPDATE

the confidence or time of further respondents.

Conclusions

Snowball sampling techniques offer an established method for identifying and contacting hidden populations and, potentially, for their enumeration although often this may be a secondary concern. From this we may want to make a distinction between snowball strategies as a method of contact in a practical sense and as a method of sampling in more formalised and statistical sense. We have seen that this latter connotation has been viewed as problematic, but that statistical techniques are being used to improve the method. Snowball-based methodologies are a valuable tool in studying the lifestyles of groups often located outside mainstream social research. They may also be used to complement other research methodologies in the study of less stigmatised and even elite groups. Advances in the quantitative application of snowball techniques and the increasing need for ascending methodologies to fill in gaps in our knowledge of more obscure social situations suggest both a complementary and substitute role for snowball sampling. The real promise of snowball sampling lies in its ability to uncover aspects of social experience often hidden from both the researcher's and lay person's view of social life.

References

- Avico, U., Kaplan, C., Korczak, D. and Van Meter, K. (1988) *Cocaine epidemiology in three European Community cities: a pilot study using a snowball sampling methodology*, Brussels: European Communities Health Directorate.
- Berg, S. (1988) Snowball sampling, in Kotz, S. and Johnson, N. L. (Eds.) *Encyclopaedia of Statistical Sciences Vol. 8*.
- Faugier, J. and Sargeant, M. (1997) Sampling hard to reach populations, *Journal of Advanced Nursing*, vol. 26, 790-797.
- Fisher et al (1994)
- Griffiths, P., Gossop, M., Powis, B. and Strang, J. (1993) Reaching hidden populations of drug users by privileged access interviewers: methodological and practical issues, *Addiction*, vol. 88, 1617-1626.
- Groger et al (1999)
- Hendricks, V. M., Blanken, P. and Adriaans, N. (1992) *Snowball Sampling: A Pilot Study on Cocaine Use*, Rotterdam: IVO
- Inciardi, J.A. (1977) In search of the class canon: a field study of professional pickpockets, in Weppner, R. S. (Ed.) *Street Ethnography: Selected Studies of Crime and Drug Use in Natural Settings*, Beverley Hills: Sage.
- Kaplan, C. D., Korf, D. and Sterk, C. (1987) Temporal and social contexts of heroin-using populations: an illustration of the snowball sampling technique, *Journal of Mental and Nervous Disorders*, vol. 175(9), 566-574.
- McNamara, R. P. (1994) *The Times Square Hustler: Male Prostitution in New York City*, Westport: Praeger.
- Moore, R. (1996) 'Crown Street Revisited', *Sociological Research Online*, vol. 1(3) <http://www.socresonline.org.uk/1/3/2.html>
- Patrick, J. (1973) *A Glasgow gang observed*, London: Eyre Methuen.
- Pollok, M. and Schlitz, M. A. (1988) Does voluntary testing matter? How it influences homosexual safer sex, Paper presented at the Fourth International Conference on AIDs, Stockholm, Sweden, June 13, 1998.
- Saunders, P. (1979) *Urban politics: a sociological interpretation*, London: Hutchinson.
- Shaw, I., Bloor, M., Cormack, R. and Williamson, H. (1996) 'Estimating the Prevalence of Hard-to-Reach Populations: The Illustration of Mark-Recapture Methods in the Study of Homelessness', *Social Policy and Administration*, vol. 30(1), 69-85.
- Spren, M. (1992) 'Rare populations, hidden populations and link-tracing designs: what and why?', *Bulletin Methodologie Sociologique*, vol. 36, 34-58.
- Snijders, T. (1992) 'Estimation on the basis of snowball samples: how to weight', *Bulletin Methodologie Sociologique*, vol. 36, 59-70.
- Sudman, S. and Freeman, H. (1988) 'The Use of Network Sampling for Locating the Seriously Ill', *Medical Care*, vol. 26(10), 992-999.
- Thomson, S. (1997) Adaptive sampling in behavioural surveys, *NIDA Research Monograph*, 296-319.
- Van Meter, K. (1990) Methodological and Design Issues: Techniques for Assessing the Representatives of Snowball Samples, *NIDA Research Monograph*, 31-43.
- Vogt, W. P. (1999) *Dictionary of Statistics and Methodology: A Nontechnical Guide for the Social Sciences*, London: Sage.
- Whyte, W. F. (1955) (2nd ed.) *Street corner society: the social structure of an Italian slum*, Chicago: University of Chicago Press.

social research UPDATE

(ISSN: 1360-7898)

is published by

Department of Sociology
University of Surrey
Guildford GU2 5XH
United Kingdom.

Tel: 01483 300800

Fax: 01483 259551

Edited by Nigel Gilbert

(e-mail: gng@soc.surrey.ac.uk)

Summer 2001 © University of Surrey