Review Article

Types of Sampling in Social Research

Florentina Scârneci-Domnișoru[1]

1. Transilvania University of Brașov, Romania

The value and credibility of research results depend greatly on how the subjects or participants are selected. Nevertheless, the sampling process is often superficially described in many research reports and articles, and the chosen sampling procedure is rarely justified by researchers. One possible cause of this situation may be the inconsistencies present in the methodological literature regarding this stage of research. In an attempt to provide clarification, I propose a classification of sampling methods that can be used in studies involving multiple units of investigation, using the criterion of representativeness. Probability sampling and theoretical sampling are representative; they allow for internal generalization of results through statistical induction and analytical induction respectively. All other types of sampling described in the specialized literature are unrepresentative. For these latter types to be acceptable in scientific research, they must be systematic, controlled, and show a concern for increasing representativeness by reducing subjectivity in the process of selecting subjects / participants in the study. Applying these acceptability criteria forces greater transparency in the sampling procedures used and allows for more accurate evaluation of the research results.

Correspondence: papers@team.qeios.com — Qeios will forward to the authors

1. Introduction

The selection of subjects or participants for research is one of the most important decisions a researcher must make. It is a decision that has a major impact on research results, especially because, depending on the chosen sampling procedure, the results will either be valid only for those who were studied (interviewed, surveyed) or will be generalizable to the entire research universe.

Despite this, researchers often treat it rather superficially. Few still consider using sampling procedures that allow for generalization of the results, and the aim of trying to eliminate subjectivity from the selection process of subjects or participants has lost popularity. Any sampling procedure has become

acceptable; social researchers are no longer embarrassed to reveal that they chose only convenient subjects, including in their studies only those who were at hand. And if convenience sampling does not sound good when describing their sampling method, they use another label that better conceals the lack of objective rules in the selection process of subjects or participants: purposive sampling.

In fact, most researchers have reached the point where they no longer disclose the sampling procedure they used, let alone justify it: "Unfortunately, the vast majority of qualitative and quantitative researchers do not make clear their sampling decisions. Indeed, the exact nature of the sampling scheme is rarely specified. As such, sampling in qualitative and quantitative research appears to be undertaken as a private enterprise that is unavailable for public inspection"[1].

In my opinion, we should talk more about the sampling procedure and make it transparent, considering the effects that the way we choose subjects or participants has not only on the results of our research, but also on science and the production of social knowledge in general. The first step we should take is to clarify the descriptions of the sampling stage in methodological literature: what the types of sampling are, by what criteria they are classified, whether there are sampling methods specific to different types of research, and so on. This because, at least in qualitative research – and implicitly, in mixed method research – things are far from clear: "The methods literature regarding sampling in qualitative research is characterized by important inconsistencies and ambiguities, which can be problematic for students and researchers seeking a clear and coherent understanding"[2].

Through this article, I will attempt to provide a general classification of sampling types (thus applicable to any social research, regardless of its type), using the only criterion that truly matters in the selection of subjects or participants: representativeness.

2. Existing classifications of sampling and issues they raise

The literature addresses the subject of sampling as specific to different types of research. Thus, in quantitative research, the literature typically divides sampling into two categories: probability and non-probability. Probability sampling involves random selection of subjects. Four or five types of probability sampling are usually listed: simple random sampling, stratified random sampling, cluster random sampling, systematic random sampling, and multi-stage sampling (see [3] or [4]). Non-probability sampling, on the other hand, is defined in opposition to probability sampling, specifying that it uses criteria other than random selection. Non-probability sampling usually includes quota sampling,

convenience sampling, purposive sampling, and snowball sampling. But what are the criteria used to select participants in non-probability sampling? In my view, a negative definition is not sufficient to clarify things – it is not enough to say that the selection is non-random to fully describe a sampling procedure. And if the selection is not rule-based and is subjective (for example, involving subjects at hand), can such a selection be scientifically justified? There are authors who categorize additional types of sampling – typically used in qualitative research – as non-probability as well. For example, Neuman[4] also lists theoretical sampling or deviant case sampling among non-probability sampling.

In qualitative research, the label purposive or purposeful sampling is often used to describe the participant selection process, and methodology authors have identified up to 40 types of purposive sampling schemes (see [5]), including convenience sampling and snowball sampling. But are non-probability sampling and purposive sampling synonymous? Do they refer to the same procedure? Are convenience sampling and snowball sampling forms of purposive sampling (as seen in qualitative methods literature), or are they self-standing techniques (as seen in quantitative methods literature)? Is it possible to use the same sampling methods (e.g., convenience sampling or snowball sampling) in both quantitative and qualitative research? What are the criteria used to classify types of purposive sampling? Given that the number of these types seems to increase (at least) with each edition of Patton's book, is it possible that any selection procedure becomes acceptable once it is labeled?

In mixed-method research, the authors present the united sets of sampling techniques specific to qualitative and quantitative research as possibilities of subject or participant selection (i.e., the 4 or 5 probability sampling plus 4 to 40 purposive sampling schemes). For example, in 2007, Onwuegbuzie and Collins stated: "mixed-method researchers presently have 24 sampling schemes from which to choose. These 24 designs comprise 5 probability sampling schemes and 19 purposive sampling schemes [non-probability]" (p. 287). The authors listed among the purposive sampling variants: maximum variation, intensity, typical case, as well as politically important case and opportunistic or random purposeful. Can sampling techniques be combined within a single study (e.g., two random types, two or more purposive types, or one random and one purposive)? Is sampling a relevant issue in case study situations? If so, is sampling in case studies conducted using the same procedures as in studies with multiple units of investigation?

It is worth noting that online research has introduced some specific particularities related to subject / participant selection. However, the principles of sampling remain the same, and the types of online sampling listed in the literature are generally similar to those used in traditional research. Online

sampling often faces difficulties in the adapted applying of certain traditional sampling techniques; in some cases, older sampling techniques have no equivalent online (see [6]). Additionally, online studies more frequently use non-probability techniques such as opt-in or voluntary sampling, where the responsibility for selection shifts largely from the researcher to the subjects / participants themselves (they self-select to take part in the research). Should online sampling follow different criteria than traditional sampling? Which classic sampling techniques can be applied in online research, and which cannot? Are there sampling techniques specific to online quantitative research? What about online qualitative and online mixed-method research?

3. Proposal for a new classification of sampling in social research

I will not refer in this article to the situations where we study entire populations. In these situations (see, for example, the census), it is no longer a matter of sampling; every person in the research universe is studied. This type of research is perfect when and where it can be used. There are many (especially applicative) research situations where this is possible and even advisable; when this can be done it is good to do so. For instance, when studying all employees in a department, the management staff in a company, all children in a kindergarten, or the medical staff in a small hospital. Likewise, I will not address case studies here, whether concerning a person, a profession, or an institution. Choosing the cases by the typical case, extreme case principle, critical or deviant case, and the discussion of when and for what purpose such cases are studied will be the subject of another article. I will refer here only to research involving multiple units of investigation in which only a part of the research universe is studied (i. e., where sampling is applied).

The need to select subjects or participants arises from constraints related to time, budget, and efficiency. We do not have the time or sufficient funds to study the entire population of interest. But it is also not worthwhile (i. e., not cost-effective) to study the entire population if the results obtained from a sample are similar to those we would obtain by involving the entire research universe. So, the purpose of sampling is frequently forgotten when performing sampling in scientific social research. We do not sample to make things easier or more convenient, nor because of laziness or cunning. We sample to reduce costs. But we do not reduce costs at any price – certainly not at the price of transforming our research from a scientific study into something else: an essay, a report, or mere chatter. Therefore, when we choose the subjects to study, we must ensure that the results obtained from the sample can be extrapolated to the entire research universe: “Every scientific enterprise tries to find out something that

will apply to everything of a certain kind by studying a few examples, the results of the study being, as we say, 'generalizable'"(Howard Becker, apud [4]). In short, the only truly relevant criterion in classifying sampling procedures is whether the resulting sample is representative – that is, whether what was studied in a few can be extended to many.

Thus, the types of sampling discussed in the literature can be divided into two major categories: representative sampling, which allows for generalization of results, and unrepresentative sampling, which provides results valid only for the individuals who were surveyed / interviewed / observed. The first category includes probability sampling and theoretical sampling, and the second category includes all other types mentioned in the literature, the so-called non-probability sampling and purposive sampling.

3.1. Representativeness

Representative means typical, standing or acting for another. A sample is representative if it reflects the population from which it is drawn. Therefore, "in a representative sample, our goal is to create sample data that mirror or represent many other cases that we cannot directly examine"[4]. Babbie[3] shows that representativeness is "that quality of a sample of having the same distribution of characteristics as the population from which it was selected" (p. 133).

In its basic definition, sample representativeness is a continuum, with varying degrees of representativeness: "Representativeness is a relative notion, in the sense that a sample is more or less representative, not simply representative or unrepresentative"[7]. If we were to approach it mathematically (express it quantitatively), we would range from 0% representativeness (in a sample with zero individuals) to 100% representativeness (when the sample equals the population).

But what is the use of representativeness in research? Why do we talk about it? It matters because, depending on how representative a sample is, we determine whether the results obtained from the sample can be generalized to the research universe. That is, whether "the descriptions and explanations derived from an analysis of the sample may be assumed to represent similar ones in the population"[3]. It is clear that with 0% representativeness, generalization is not possible, which is the reason why researchers have sought an acceptable level of representativeness from which results can be generalized.

Rotariu and Iluț[7] argue that "the use of the term 'representative sample' is justified if it means that, with respect to all studied characteristics, the error d is below the acceptable limit, and the value of p is below

5% (or P above 95%)" (p. 127). So, we have agreed to use the expressions "representative sample" / "unrepresentative sample" even though representativeness is a gradual quality. By convention, in research activities, we use the term "representative sample" (when values of d and p are below accepted limits) to refer to samples that allow result generalization, and the term "unrepresentative sample" for cases where values of d and p exceed acceptable limits, and results apply only to the investigated individuals. Because of variations in d and p (above and below acceptable values), we can maintain the idea of gradation in representativeness and say that a sample may be more or less representative, or more or less unrepresentative.

Sample representativeness can only be calculated for probability samples, and "we cannot speak of the representativeness of a sample in general, but only in relation to a given characteristic"[7]. Therefore, a sample may have a certain level of representativeness for gender, another for age, and another for voting intention. In research, representativeness should be calculated for each characteristic individually, but it is considered that if the acceptable level of representativeness is achieved for the characteristics with the greatest dispersion (i.e., the main socio-demographic ones), then the sample is also considered representative for the other characteristics (such as opinions, expectations, intentions, or attitudes).

It should also be noted that the mathematical approach, the quantitative expression of representativeness, is not the only one possible. Thus, if we cannot calculate representativeness (based on d and p), this does not mean that our sample cannot be representative, or that at some point during data collection, it may not become representative enough for result generalization. Representativeness is a general quality, not one exclusively tied to probability sampling: "random and representative are neither synonymous nor necessarily interrelated terms"[8]. In other words, it is expected that even non-probabilistic samples may be more or less unrepresentative, and there may even be a point at which it can be said that non-probability sampling can generate sufficient, acceptable representativeness to allow for generalization. The problem is that in the absence of a mathematical calculation (possible only in probability sampling), it is very difficult to determine when this point is reached. For this reason, all non-probabilistic samples are considered unrepresentative. But this does not mean that the representativeness of the non-probability sample should not preoccupy us. Researchers should strive to increase representativeness (or decrease the unrepresentativeness) of the sample they are working with.

It should also be mentioned that what has been described above is the statistical approach to representativeness, but this is not the only possible approach. The statistical approach involves controlling the distribution of general population characteristics – gender, age, marital status, income,

and background. Statistical representativeness / unrepresentativeness varies depending on how well this distribution is controlled. As a side note, randomization ranks so highly among procedures that ensure a high degree of representativeness precisely because it excels in controlling variables – not only observable ones but also unobservable ones, which the researcher might not even think of considering. Returning to the main point, representativeness / unrepresentativeness of a sample can be determined not only in relation to the statistical distribution of socio-demographic characteristics, but also in relation to variations in the data collected from subjects (e.g., variations in the perceptions or experiences reported by our subjects). Controlling data variation leads to a different kind of representativeness. I have called it theoretical representativeness.

Therefore, if through sampling we aim to identify and include participants who generate variety in the data in a sample, it means we are seeking to increase the representativeness of the sample – even if we are not talking about statistical representativeness. In these cases, "representative samples are not predicted in advance but found, constructed and discovered gradually in the field"[9]. Smaling[10] argues that: "the purpose is not to cover statistical distribution but merely to cover the existing variation. The sample is not representative in a statistical manner, but rather it represents the variation. The researcher can aim at generalization based on covering the variation" (p. 7). Depending on how much of this variation is covered in the data, we have samples that are more or less representative / unrepresentative. Once the variation is sufficiently well-covered in the sample, it can be said that sampling is representative and the results can be generalized to the whole population: "covering more of the variation in qualitative different views will enhance the generalizability of the study"[11].

As the literature shows, the statistical representativeness of the sample (and implicitly the statistical generalization of research results) is ensured by probability sampling. Similarly, the theoretical representativeness of the sample (and theoretical generalization of results) is ensured by theoretical sampling (see [12]). In situations where data variation is not controlled (i.e., in unrepresentative sampling), it remains desirable for researchers to strive for an inventory of the most consistent diversity of subject / participant perspectives to increase theoretical representativeness (or decrease unrepresentativeness) of the sample.

3.2. Representative sampling

We are interested in representativeness because the results obtained from a representative sample can be extrapolated to the entire population from which it was drawn. This refers to the internal generalization

of research results, according to the distinction proposed by Maxwell[13]: "internal generalization refers to generalizing *within* the setting, group, or population studied, to persons, events, and activities that are not directly represented in the data collected; [...] external generalization, in contrast, refers to generalization *beyond* the person(s), setting, case, or time specifically studied, to other persons, settings, cases, or times" (p. 112).

Normally, studies with multiple units of investigation should prioritize the generalization of results, which means using representative sampling. If they do not, it means they aim to explore the phenomena under investigation. Without minimizing the importance of exploration, we must acknowledge that without generalization, knowledge is limited, has lower impact, and reduced utility: "no matter how interesting or insightful the research is, if it is not generalizable then it is not considered to be evidence that can be put to use"[14]. As Hans Reichenbach, a leading German philosopher of science, states: "the essence of knowledge is generalization." Exploration is only the beginning of knowledge, and: "generalizability of research findings is one of the key criterion researchers use to determine the value and usefulness of research"[15]. More on the importance of generalization for the advancement of knowledge in social sciences can be found in Scârneci-Domnișoru[16][12].

If we aim for internal generalization of results, we can perform two types of sampling in social research: probability sampling and theoretical sampling. In specialized literature, generalization from sample to population is strictly associated with probability statistical sampling and quantitative research (just as representativeness is strictly linked to randomization in literature). However, based on Maxwell's[13] idea that: "internal generalizations are not necessarily statistical" (p. 112), I showed in Scârneci-Domnișoru[12] that a theoretical sampling procedure exists that allows generalization from sample to population in qualitative studies.

In probability statistical sampling, we know before entering the field how many subjects the sample will include and what characteristics they will have. Their number is determined based on dispersion, probability, and margin of error, and the selection of subjects is made according to strict sampling rules: "samples [are] selected in accord with probability theory, typically involving some random-selection mechanism"[3]. Probability sampling can be simple random, systematic, stratified, cluster, or multi-stage. What these statistical sampling techniques have in common is that they relate to the population – size, heterogeneity, and its socio-demographic characteristics. In these cases, representativeness is statistical,

and it is calculated: “with a probability sampling strategy, we try to create an accurate representative sample that has mathematically predictable errors”[4].

We apply statistical sampling especially when we have to study large, heterogeneous populations, when we know their size and socio-demographic features, when we collect data with structured or standardized techniques, and when we aim to test hypotheses about populations with a specific profile. In such situations, we want to know how often a variant occurs in the population; our interest is mainly in testing theories.

Concerning theoretical sampling, the selection is made by the relevance of cases without knowing before entering the field how many participants will be needed or all the characteristics they will have. Processing, analyzing, and interpreting data will show what sort of participants need to be further selected, and saturation will show at what moment the selection of new participants can be stopped. Theoretical sampling can be guided by variation or by theory, and these two subtypes of sampling allow (just like the probability sampling techniques) generalization from sample to population (again, see Scârneci-Domnișoru[12], for a detailed description of the sampling procedure that allows for an internal generalization in qualitative studies). In this case, the representativeness is theoretical.

What theoretical techniques have in common is that they relate to data and the way in which they vary, or to theory and the way in which it is grounded in data. We apply theoretical sampling especially when we have smaller and more homogeneous populations, when we collect data with unstructured or non-standardized techniques, when we do not collect all data at once, at the beginning of research. In this situation, we are interested especially in the data variations and not how often they occur; we are interested in completing or discovering theories.

I have tried to summarize the types of sampling in Figure 1, focusing on those that allow result generalization. That figure also visualizes the synthesis of ideas presented below, describing the process of generalization depending on the type of sampling used.

It is possible to generalize results from sample to population not only through statistical sampling, but also through theoretical sampling, because there are, as Mitchell[17] shows: “two very different and even unconnected inferential processes – that of statistical inference, which makes a statement about the confidence we may have that the surface relationships observed in our sample will in fact occur in the parent population, and that of logical or scientific inference, which makes a statement about the confidence we may have that the theoretically necessary or logical connection among the features

observed in the sample pertain also to the parent population" (p. 207). There are authors who hold even the superiority of logical inference (Ragin[18] calls it "interpretive inference") to statistical inference – see Znaniecki (apud [8]) who describes the process as being "the true method of science and it is the superior method (because it discovers the causal relations of a phenomenon rather than only the probabilistic ones of co-occurrence)" (p. 195).

Internal generalization (that is, extending statistical and/or interpretive inferences to the entire population) is performed based on inductive reasoning, so it is a form of inductive generalization. Smaling[10] argues that there are three forms of inductive generalization: "statistical generalization, variation-based generalization and theory-carried generalization" (p. 2) and that each of these three types of generalization "is concerned with drawing conclusions pertaining to (a class of) subjects, cases, situations, etc., that have not been researched, based on a limited number of persons, cases, situations, etc., that have been researched. In all those cases, a generalization is made from research results to a population or to a scope belonging to a theory. In this sense, these forms of generalization can be seen as forms of inductive generalization: methods of arriving at general or universal propositions" (p. 8).

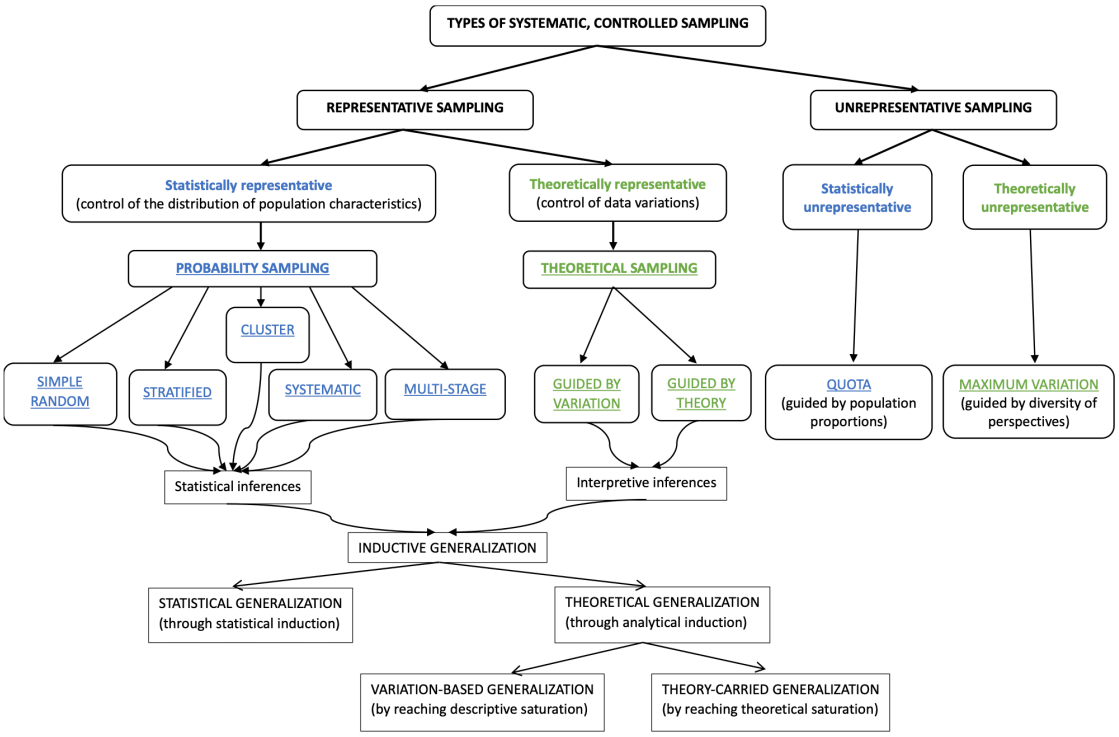

Figure 1. Types of sampling

doi.org/10.32388/9HVN52

Thus, inductive generalization can be statistical generalization (that being performed with probability statistical sampling, through a statistical induction process) or theoretical generalization (that being performed with theoretical sampling, through an analytical induction process). Theoretical generalization can be guided by variation or by theory. Variation-based generalization can be achieved in descriptive studies through reaching descriptive saturation, as Smaling[10] calls it. Theory-carried generalization can be achieved in explanatory studies through reaching theoretical saturation.

In the case of studies with multiple units of investigation, representative sampling is ideal. There is no comparison between research whose results are valid for the entire population of interest and one whose results are valid only for a few subjects / participants from whom data was collected. In fundamental research, representative sampling allows for relevant theory testing (on large, diverse populations rather than limited, local ones), for refining, modifying, or grounding new theories. These are essential contributions to the progress of knowledge in social sciences. Regarding applied research, representative sampling leads to an increase in the usefulness and impact of research results. Knowledge valid for whole populations enables the implementation of general or well-tailored solutions. Moreover, by providing reliable and useful information, researchers gain the trust and respect of beneficiaries of social knowledge, along with confidence in scientific research and science overall.

3.3. Unrepresentative sampling

Unfortunately, we cannot always perform representative sampling. Neuman[4] states that other types of sampling are “acceptable when probability sampling [I would say when representative sampling] is impossible, too costly, time consuming, or impractical” (p. 248). These are unrepresentative types of sampling. As already shown, at least in the qualitative methodological literature, the list of purposive (unrepresentative) sampling procedures is becoming longer and longer. In this context, the question arises: is any kind of subject / participant selection acceptable in scientific research?

In my opinion, scientific studies should abide by certain standards. There should be rules guiding all choices, procedures, and the knowledge production process in general. Sampling should not be an exception; just as there are rules for how we collect and analyze data, there must also be rules for selecting subjects / participants. We all agree that what we call scientific knowledge is different from common knowledge; the procedure of producing scientific knowledge also differs from the procedure of producing common knowledge; sampling in scientific research should also differ from the way we select information and informants when enriching our common knowledge. The procedure for selecting

subjects / participants in scientific projects should be different from asking for information from parents, friends, neighbors, or colleagues when we have a lack of common knowledge.

Therefore, even in simpler exploratory projects of phenomena without claims to generalize results, sampling should follow certain rules and not be done haphazardly. In other words, sampling should be systematic – organized, methodical, and carried out according to a pre-designed plan or method; the selection procedure should be rigorous and controlled. The main goal of systematization and control should be to increase the representativeness of the sample used. I believe that in scientific research, not just any level of unrepresentativeness should be accepted, and in every study, strategies to reduce unrepresentativeness (to increase representativeness) should be mandatory.

If, for example, we are interested in how citizens of Brașov decide whom to vote for in the Romanian presidential election and we go to the park to ask questions of whoever we find there, this should not be considered an acceptable sampling in scientific research (same goes for e-mailing the questionnaire to all our acquaintances, posting the link on our faculty's Facebook page, and so on). This would be non-systematic, uncontrolled sampling (because participant selection would follow no rules), and the proposed sample would be unacceptably unrepresentative (made up of participants very similar to one another), without any strategy in place to increase its representativeness.

Statistical representativeness can be increased by introducing elements of randomization. For example, if we find a row of benches in the park with many potential respondents, it would be better (from a representativeness standpoint) to use a selection method close to a random technique (e.g., say, every third person). Representativeness can also be increased by statistically controlling socio-demographics. This means that, among those on the benches, we should interview both men and women in equal proportions, and respondents should be of various ages and in different proportions (depending on the study's objectives and the characteristics of the investigated population), and so on.

Most types if sampling listed in the literature as statistically unrepresentative lack identifiable selection rules, do not aim to increase representativeness, and make no attempt to reduce the subjectivity in choosing participants. I believe that such methods cannot be considered usable sampling types in scientific research, because they are neither systematic nor controlled, and they generate a sample with an unreasonable degree of unrepresentativeness. This is the case, for example, with "convenience sampling" (i.e. accidental, availability, or haphazard sampling), where "our primary criteria for selecting cases are that they are easy to reach, convenient, or readily available"[4]. If this kind of sampling is

acceptable, can we still distinguish between the results of a scientific study that uses accidental sampling and a “journalistic survey” that presents the opinions of passersby interviewed by a street reporter?

In situations where statistically representative sampling cannot be performed, quota sampling is an acceptable alternative. In this type of sampling, even though selection is not randomized, a rule is followed when choosing subjects, and there is a concern to increase the representativeness of the sample: “units are selected into a sample on the basis of prespecified characteristics, so that the total sample will have the same distribution of characteristics assumed to exist in the population being studied”[3]. In quota sampling, researchers attempt to control certain socio-demographic characteristics, which leads to a sample with a reasonable degree of unrepresentativeness.

Theoretical representativeness can be increased by controlling the diversity of subjects / participants to increase variation in the newly collected data. For example, data collection in a park becomes acceptable (i.e., systematic, controlled, less subjective / accidental) if care is taken to interview not only elderly people present at noon in the park but also young people present in the evening, as well as mothers, fathers, and nannies who are out with children, and so on. In cases where theoretical sampling cannot be achieved, following variation in data to saturation, participants can at least be selected depending on how different they are from one another, on the idea that highly diverse participants can reveal significant data variations, thus increasing theoretical representativeness. This is known in the literature as maximum variation sampling. This sampling type controls diversity but does not verify the variation in data in the same way theoretical sampling does. Hence, maximum variation sampling does not generate a theoretically representative sample, but also not one with unreasonable unrepresentativeness.

In my opinion, purposive or purposeful sampling and many other sampling procedures grouped under this label are scientifically unacceptable because they are not systematic or controlled, and they have loose, unclear, or ambiguous definitions and instructions for participant selection (see [2]), referring only to the idea of choosing people who might provide useful information. No selection rules are specified, subjectivity in choice is implicit, and the resulting sample's unrepresentativeness is unacceptably high.

Therefore, I did not include in Figure 1 either convenience samples or purposeful samples as types usable in studies with claims of scientific research: of all the unrepresentative sampling procedures listed in the literature, I kept only quota sampling and maximum variation sampling. This does not mean that other unrepresentative acceptable types do not exist or could not be discovered. As long as the proposed sampling is systematic, its selection process is controlled and thoroughly described, in order to be assessed and as long as researchers explain the applied procedures for reducing subjectivity and prove

concern for reducing the unrepresentativeness of the resulting sample, the list of scientifically acceptable unrepresentative sampling procedures remains open to any process.

In conclusion, even when using unrepresentative sampling, we must limit subjectivity as much as possible in choosing subjects / participants in the study and control their selection using systematic procedures. A rigorous, objective exploration of the studied phenomenon—with the limiting of subjectivity in sampling—can only benefit the research, as it can propose research objectives or hypotheses closer to what can and deserves to be subsequently studied or tested. A careless exploration, on the other hand, with subjectivity in sampling, may propose conclusions easy to dismantle, unnecessary subsequent research, it may damage our credibility, and reduce public trust in researchers and social research.

There are also special research situations where the previously mentioned unrepresentative sampling procedures are harder to apply. For example, when subjects / participants are very few, hard to find, or difficult to convince to participate in the research (e.g., human trafficking victims, drug traffickers, or people with rare diseases). In such cases, no matter how much we want to apply selection rules, it is impossible or extremely difficult. In fact, the challenge is more about identifying subjects / participants than selecting among them. For instance, interviewing even one prostitute is very hard, and clearly, we cannot choose from hundreds whom we could have available. In this case, we can use “snowball sampling” as a procedure of subject identification (after interviewing one, we ask them to help us find others). Similarly, former political prisoners may be hard to identify, but the head of their association is likely to solve our problem; he / she can help us find possible participants just like a doctor might help us find patients with a certain profile. In such cases, we use what the literature calls “gatekeepers” (a person with direct access to potential subjects / participants helps us). In these special situations, we study those whom we can identify, sometimes without even thinking of the issue of the selection of subjects, because the number of potential study participants is anyway very small. Moreover, in these cases, the research universe has special features, namely it is usually more homogeneous, with fewer variations within it (both in population characteristics and in collected data).

Therefore, in my opinion, what is found in the literature under the names of snowball sampling, gatekeepers, or key informants are not sampling types, but rather strategies or techniques for identifying potential subjects. These should be used only in special cases where potential participants are hard to find, when they have features or combinations of features that make them hard to identify. Otherwise, the option for a procedure when the researcher leaves the selection of participants to someone else

(subjects or informants) is scientifically unacceptable, as, this way, the researcher loses control over the sampling process and the selection of participants. The less control there is, the more unrepresentative the sample becomes. The same happens with voluntary sampling (e.g., posting links to questionnaires on websites), where the researcher gives up the control of sampling, leaving the process of selection, to a large extent, to the subjects (which, in fact, becomes self-selection).

Thus, snowball sampling, informants, and voluntary sampling were not listed in Figure 1, among the sampling types that can be used in respectable research. This does not mean that snowball or informant procedures for the identification of subjects can never be used. They can be employed within recognized sampling processes already described. For example, in theoretical sampling, the snowball technique can be used in need to identify participants with certain traits (just like sampling techniques specific to case study can be used to select deviant cases for example). Similarly, informants can be used within quota sampling to help find possible subjects with specific socio-demographic traits, and the gatekeeper technique may be used in maximum variation sampling. Voluntary sampling may also be acceptable under certain conditions (e.g., if diversity control strategies are applied to the websites where the link to questions is posted).

I have excluded from the long list of unrepresentative types of sampling in the literature not only unsystematic ones and those that give up the control of sampling, but also those related to case study (i.e. case selection procedures - see deviant case, typical case, politically important case, and many others). Case studies and their selection procedures are not the subject of this article and will be detailed in another paper.

Returning to the main point: there are solutions, or we can develop new scientifically valid sampling solutions for each research context. But in every case, we must describe the selection process, justify the chosen sampling method (why it was used, why representative sampling could not be applied), and detail the ways to reduce subjectivity applied in choosing the subjects / participants in the study, showing our concern for increasing the sample's representativeness. Making sampling transparent and exposing it to public scrutiny allows for proper evaluation of the research results and could also increase the study's credibility.

Transparency enables research evaluators to examine the situation, understand and reproduce procedures, assess the soundness of decisions, and so on. Let me give an example: suppose I want to study the dissatisfactions of parents whose children are in a primary school, and I choose to conduct interviews. There are over 600 parents, so I cannot interview all of them – therefore, sampling is needed.

I go to the school on a weekday at noon, when the children are dismissed, and I find a few parents willing to be interviewed. I ask them to recommend other parents who could be interviewed, and after meeting them, I repeat the process. Thus, I apply what is known in the literature as snowball sampling, and after 8, 10, or 12 interviews, I decide I have reached saturation and formulate my conclusions. I have described in detail how I selected my subjects; it is already a big help in evaluating my results. By making the procedure transparent, an evaluator of my research can ask whether this was the best choice of the sample. Why opting for snowball sampling? How can this option be justified? Could the researcher have applied another sampling type?

An evaluator might rightfully say that my choice was inappropriate and that the results obtained are not very reliable or valuable. Here is the explanation: what kind of parents did I encounter at 12 o'clock in front of the school? Probably ones with certain characteristics – either unemployed or having flexible work schedules, some who do not send their children to after-school programs either because they cannot afford it or because they want to spend more time together with their children. What kind of parents did these participants then recommend for interviews? Probably others similar to themselves – parents they talk to every day in front of the school while waiting for their children, parents whose children are in the same class as theirs, and so on. It is very likely that, based on this type of sampling, I arrived at results that do not describe the dissatisfactions of parents in that school, but only those of a particular type of parents. Snowball sampling cannot be scientifically justified, neither in this case nor in similar ones. Why should I let the parents build my sample for me? How unrepresentative do the results become in doing so? A more suitable approach would have been maximum variation sampling, trying to capture a diversity of parents: parents of children in different grades, children who do or do not attend after-school programs, children who perform well or poorly in school; parents who are wealthy or less well-off, who live near or far from school, who have completed various levels of education, and so on. Or perhaps even theoretical, representative sampling could have worked.

In conclusion, unrepresentative sampling should always be a second choice, not the first. Choosing it should be acceptable only in special situations and only if explained and justified. Its application should be carried out carefully, with control maintained and with the elimination of subjectivity in the selection of study subjects / participants.

4. Combining sampling procedures

The typology of sampling procedures described in this article facilitates the understanding of the differences between the types of sampling discussed in the specialized literature. It clarifies the classification criteria and also sheds light on the issue of combining different types of subject / participant selection. Since I have reduced the number of sampling procedures acceptable in scientific research and classified them according to representativeness / unrepresentativeness as well as by the type of representativeness (statistical or theoretical) it is no longer complicated to choose a sampling procedure, nor are there dilemmas about what and how sampling methods can be combined.

Any researcher, regardless of the type of study they conduct, when faced with the selection of subjects / participants in studies involving multiple units of investigation, must make the following choices: (1) whether or not to use a representative sample, and (2) what type of representativeness they aim for. If they seek statistical representativeness, they can use probability sampling methods. If probability sampling cannot be applied, they can opt for quota - the statistically unrepresentative sampling. If the goal is theoretical representativeness, theoretical sampling should be used. If that is not feasible, maximum variation sampling can be implemented. As previously mentioned, other unrepresentative sampling procedures may also be applied, but only under certain conditions.

As for combining different types of sampling, the issue of their compatibility arises. When the sampling procedures used aim for the same type of representativeness, combination is possible. For example, as previously mentioned in this article, the use of randomization elements in statistically unrepresentative sampling is always welcome. Using these elements brings numerous benefits – the sample's unrepresentativeness decreases and the study's credibility increases. On the other hand, applying randomized selection criteria to research that seeks theoretical representativeness does not make much sense; it not only fails to offer benefits, but can even be detrimental. For example, if we are to interview a few individuals out of 50 members of an organization, it is better to select them based on how different their perspectives on the research topic might be, rather than choosing them randomly. If we randomly select 10 participants out of the 50, we will obtain lower theoretical representativeness than if we chose them based on the office they work in, job duties, hierarchical position and so on. The idea of "random purposeful" sampling in a small research universe only reduces the chances of achieving data diversity.

In mixed-method research, the same logic applies. For each component of the study the appropriate sampling type is used depending on the representativeness sought. For example: a probability sampling

in the quantitative component when surveying several hundred subjects, and a theoretical sampling in the qualitative component for conducting interviews with a few dozen participants.

In my opinion, what is presented in the literature as “stratified purposive sampling” or “purposive random sampling”, and other similar terms (see, for example, [19]), are merely labels for sampling procedures created using probabilistic logic, but which, at some point in the selection process, failed to follow the rules of random selection and became non-probabilistic. The “purposive” label inserted in these names does not indicate a combination of sampling procedures, it only eliminates their representative character.

5. Conclusion

The sampling typology proposed in this article emphasizes sampling that allows for generalization – controlled sampling based on rules – and places less focus on other types of sampling commonly mentioned in the specialized literature. For many of the latter, criteria for the selection of subjects / participants are very hard (or impossible) to identify, and this is unacceptable for scientific research, even if exploratory. These sampling techniques should be avoided or used only in exceptional situations, and the rule should be to attempt to eliminate subjectivity from the sampling process.

The first choice of any researcher should be representative sampling, followed by other methods, depending on the context. What truly matters is how appropriate the chosen sampling procedure is for the given situation, considering the aim of scientific research, the specific objectives of the study, the characteristics of the target population, and the information available about it.

The representativeness sought by researchers can be statistical or theoretical. A probability sample is statistically representative, while a quota sample is not. Theoretical sampling ensures theoretical representativeness, while maximum variation sampling does not. These are systematic sampling types, controlled by the researcher, guided by rules, and with reduced subjectivity in the selection of subjects / participants, ensuring an acceptable representativeness or unrepresentativeness of the sample. Other sampling procedures mentioned in the literature lack these characteristics.

It should be noted that this article only refers to studies with multiple units of investigation, and not to case studies. In case studies, the sampling issue has different characteristics due to the purpose of this specific research and to the type of generalization they aim for (external, not internal; analogical, not inductive).

A detailed presentation of the sampling procedure applied should be mandatory for any researcher. The selection of subjects / participants should be transparent, and decisions regarding selection should be justified. In this way, research studies can be properly assessed, and the value of the results can be more easily established. The obligation to describe and justify the sampling process would increase researchers' attention to the sampling procedure, foster greater responsibility in conducting research, and reduce the number of carelessly conducted studies.

Statements and Declarations

Acknowledgements

I would like to thank my colleagues Daniela Sorea and Carmen Buzea, whose reviews of my articles have become indispensable to me. I also thank Claudia Ciubancan, who, for several years, has translated my texts accurately and fluently into English.

References

1. ^*Onwuegbuzie AJ, Collins KMT (2007). "A typology of mixed methods sampling designs in social science re search." The Qualitative Report. 12(2):281–316. doi:10.46743/2160-3715/2007.1638.*
2. a, b*Gentles SJ, Charles C, Ploeg J, McKibbon KA (2015). "Sampling in qualitative research: Insights from an o verview of the methods literature." The Qualitative Report. 20(11):1772–1789. http://nsuworks.nova.edu/tqr/ vol20/iss11/5.*
3. a, b, c, d, e*Babbie E (2013). The Practice of Social Research. Wadsworth, Cengage Learning.*
4. a, b, c, d, e, f, g*Neuman L (2014). Social Research Methods: Qualitative and Quantitative Approaches. Pearson Education Limited.*
5. ^*Patton Q (2015). Qualitative Research and Evaluation Methods. Sage Publications.*
6. ^*Fricker RD (2017). "Sampling Methods for Online Surveys." In Fielding N, Lee RM, Blank G (Eds.), The Sage Handbook of Online Research Methods. Sage Publications. pp. 162-183.*
7. a, b, c*Rotariu T, Iluț P (2006). Sociological Survey and Opinion Poll. Polirom.*
8. a, b*Gobo G (2008). "Re-conceptualizing generalization: Old issues in a new frame." In Alasuutari P, Bickma n L, Brannen J (Eds.), The Sage handbook of social research methods. Sage Publications. pp. 193-213.*

9. ^Gobo G (2004). "Sampling, representativeness, and generalizability." In Seale C, Gobo G, Gubrium JF, Silver man D (Eds.), Qualitative research practice. Sage Publications. pp. 435–456.
10. a, b, cSmaling A (2003). "Inductive, analogical, and communicative generalization." International Journal of Qualitative Methods. 2(1). Article 5. Retrieved from http://www.ualberta.ca/~iiqm/backissues/2_1/html/smaling.html.
11. ^Larsson S (2009). "A pluralist view of generalization in qualitative research." International Journal of Research & Method in Education. 32(1):25–38. doi:10.1080/17437270902759931.
12. a, b, c, dScârneci-Domnișoru F (2024). "From sample to population generalization in qualitative research." The Qualitative Report. 29(8):2362–2383. https://nsuworks.nova.edu/tqr/vol29/iss8/14/.
13. a, bMaxwell JA (2021). "Why qualitative methods are necessary for generalization." Qualitative Psychology. 8(1):111–118. doi:10.1037/qup0000173.
14. ^Bloor M, Wood F (2006). Keywords in qualitative methods. Sage Publications. doi:10.4135/9781849209403.
15. ^Ercikan K (2009). "Limitations in sample-to-population generalizing." In Ercikan K, Roth W-M (Eds.), Generalizing from educational research. Routledge. pp. 211-234.
16. ^Scârneci-Domnișoru F (2023). "Data interpretation in social research: A guide to standardising research outcomes and outcome evaluation." Qeios Preprint. doi:10.32388/DELXFD.
17. ^Mitchell JC (1983). "Case and situation analysis." The Sociological Review. 31(2):187–211. doi:10.1111/j.1467-954X.1983.tb00387.x.
18. ^Ragin C (2023). Analytic Induction for Social Research. University of California Press.
19. ^Teddlie C, Yu F (2007). "Mixed Methods Sampling: A Typology with Examples." Journal of Mixed Methods Research. 1(1):77–100.

Declarations

Funding: No specific funding was received for this work.

Potential competing interests: No potential competing interests to declare.