The Sampling Conundrum in Qualitative Research: Can Saturation Help Alleviate the Controversy and Alleged Subjectivity in Sampling?

Favourate Y Sebele-Mpofu[1]

[1] National University of Science and Technology, Zimbabwe

Correspondence: Favourate Y Sebele-Mpofu, National University of Science and Technology, Zimbabwe.

Abstract

Sampling is one of the most controversial matters in qualitative research. Qualitative researchers have often been denounced for not giving adequate rationalisations for their sample size resolutions. This study aimed to provide an extensive review of sampling methods used in qualitative research and discuss the extent to which saturation might help alleviate the issues concerning these methods, sample size sufficiency and when to sample. The study specifically honed on the sampling adequacy (how big or how small should a sample be), the sampling techniques used and whether sample sizes should be delineated a priori, posteriori or during analysis. Having highlighted, the paradoxically nature of these aspects, through an overview of the sampling process, the researcher explored saturation as a tool to alleviate the challenges and the lack of objectivity in sampling in qualitative research. The overall findings were that, saturation does provide same degree of transparency and quality in sampling, but the concept is not immune to controversy, guidelines on how to apply it or achieve it remain foggy and contestable among researchers. Discussions are in most cases oversimplified and comparatively unknowledgeable. The answer to the research question, was that, what really constitutes an adequate sample size is only answerable within the context of the study, scientific paradigm, epistemological stance, ontological and methodological assumptions of the research conducted. Contextualisation of the mode of saturation adopted, clear articulation of the research methodology and transparent reporting of the whole process is key to enhance the role of saturation in alleviating subjectivity in sampling. This paper sought to make a contribution to the on-going methodological discourse on how qualitative researchers can justify their sampling decisions.

Keywords: qualitative research, conundrum, saturation, sample sizes, guidelines

1. Introduction

Sampling is one of the most challenging and contended ingredients in qualitative research (Sim, Saunders, Waterfield, & Kingstone, 2018). In consonance Vasileiou, Barnett, Thorpe, and Young (2018:1) state that "choosing a suitable sample size in qualitative is an area of conceptual debate and practical uncertainty". Disputation surrounds the selection of sample sizes, how to sample, should the samples be big or small and how big or how small (Moser & Korstjens, 2018; Mpofu, 2021a; Saunders et al., 2018). Qualitative researchers have often been denounced for failing to persuasively rationalise their sample selections or let alone justify them. At times, where they have attempted to offer reasoning for the decisions, the validations are naïve and comparatively unsubstantiated (Boddy, 2016). Frequently raised queries include whether sample sizes should be concluded a priori, through adjustable strategies such as sampling or a posteriori, during data analysis (Sim et al., 2018). Shedding more light on the sampling paradox, Vasileiou et al. (2018:1) advance that "sample size insufficiency was seen to threaten the validity and generalisability of studies' results, with the latter being frequently conceived in nomothetic terms"

Saturation has been increasingly proffered by scholars as a measure that can help resolve the issues of subjectivity in sampling, data collection and enrich qualitative research validity (Guest, Namey, & Chen, 2020; Sebele-Mpofu, 2020). Positivist scholars have questioned the credibility, quality and depth of qualitative researches, saturation is argued to be an important edict to help address this criticism (Guest, Namey, & McKenna, 2017; Morse, 2000, 2015; Van Rijnsoever, 2015). According to Morse (2015): 587, "saturation is the most frequently touted guarantee of qualitative rigor offered by authors to reviewers and readers". In affirmation Hancock, Amankwaa, Revell, and Mueller (2016:2124) describe it as "the qualitative research "gold standard" for quality research" and Guest, Bunce, and Johnson (2006:60) as the

instance where "no new information or themes" are evident from the collected data. In light of the above background information, this study discusses the sampling paradox in qualitative research by focusing on the sampling methods available to qualitative researchers, then further expounds on the techniques that can guide these researchers when choosing a sample (rule of the thumb, statistical formulae, conceptual models among others) (Malterud, Siersma, & Guassora, 2016; Sim et al., 2018). Each technique is evaluated for its cogency, pros and cons.   Lastly the study further explains whether and how saturation can assist in adaptively selecting adequate samples and sample sizes as well as alleviating the sampling quandary that affects qualitative researchers.

The research was motivated firstly, by the fact that, being a researcher in taxation (which has a quantitative essence) and tax policy, understanding people's views and opinions (qualitative) is always pivotal to effectively evaluating any policy (Mangioni & McKerchar, 2013; McKerchar, 2008), understand it from the perspectives of those who have experienced it. This is demonstrative of the mixed nature of taxation (quantitative and qualitative). This observation points to a glaring recognition that the interpretivist paradigm, constructivist epistemological view and the qualitative research (phenomenology, ethnography, grounded, content analysis and case studies) are key components of tax research. So how to improve its quality through appropriate sampling is important.   Secondly, the study was prompted by an interesting publication by Vasileiou et al. (2018) whose findings made it visibly evident that the doubts in the validity and quality of qualitative research are not only raised by positivist scholars, but qualitative research scholars as well. The researchers gave evidence how qualitative researchers often fail to justify their sample sizes or wrongly justify them (Francis et al., 2010; Mason, 2010) and in instances where they adequately justify them, they are apologetic for their sample sizes, show lack of confidence in their samples as well as their findings and at times question the completeness and depth of their own findings even after having justified how saturation point was attained (Vasileiou et al., 2018).   If qualitative researchers have no confidence in their own samples and findings, it would be difficult for others to find the research believable. The question is, why this seemingly lack of confidence in the samples and results especially the reference of failure to achieve statistically generalisability? Is it because there is a gap in the understanding of generalisability as it relates to qualitative research or perhaps a lack of comprehensive understanding of sampling decisions as they relate to validity of findings (not representational generalizability, but trustworthiness, credibility and transferability qualities of qualitative research) ? The study also sheds light on the saturation paradox in relation to sample size determination because despite the concept gaining prominence, guidelines on how to meaningfully apply it still remain vague, inconsistent and contestable among researchers (Guest et al., 2020; Morse, 2015; Morse, Lowery, & Steury, 2014). This paper is an extensive literature review of existing conceptual models and empirical studies on qualitative research such as Boddy (2016), Moser and Korstjens (2018) and Vasileiou et al. (2018), challenges in sampling and the emergence of saturation as an answer to some of these problems faced by qualitative researchers as discussed by researchers such as Hennink, Kaiser and Weber (2019).

2. The Sampling Conundrum in Qualitative Research.

As highlighted earlier in the study, sample size estimation has been a topic of "enduring discussions" (Vasileiou et al., 2018:2). Despite it being easy to select sample sizes with precision and certainty quantitative research using statistical formulae, qualitative researchers grapple with the complexities of sample size determination originating from the multiplicity in perspectives regarding: designs, theoretical, methodological, ontological, epistemological and ideological viewpoints present in the qualitative approach (Vasileiou et al., 2018). Sample size delineation raise many questions. How many participants? How many interviews? (Guest et al., 2006; Hennink, Kaiser, & Marconi, 2017) How many focus groups? How many participants per focus group (Guest et al., 2017; Hennink et al., 2019)? Baker and Edwards (2012), summarising various answers to the question from prominent qualitative research scholars deduce that the ideal rejoinder to the questions is "it depends". The articles from the various researchers argue that judgment is dependent upon a variety of factors. The decision hinges on factors which include: the nature of participants, quality of the pool (heterogeneity, homogeneousness and knowledge as well as competence) (Hammersley, 2015), minimum sample requirements, breadth and scope of the research and research questions, epistemological view, methodological assumptions, theoretical underpinnings and saturation (Baker & Edwards, 2012; Vasileiou et al., 2018). In a compatible opinion, Morse (2000) tables important aspects of consideration such as: the breadth of the study, intricacy of the topic, accessibility of participants, design of the study, data quality, structure of the questions (structured, semi-structured, unstructured, close ended or open ended). Malterud et al. (2016:1756) refer to the "information power" of the sample as a key cogitation in sample size determination. Information power depends on: "a) aim of the study b) sample specificity c) use of established theory d) quality of dialogue e) analysis strategy".   Through an assessment of various selected qualitative studies from three journals, Vasileiou et al. (2018) found out several justifications for the sample sizes given by the researchers and these included: saturation, pragmatic considerations, meeting sampling guidelines suggested by previous researchers, researchers' prior experience, richness and volume of data, further sampling to test for consistency in findings and the quality of data analysis. The emblematic deduction is that there is no one size fit all sampling (Sim et

al., 2018; Vasileiou et al., 2018), thus the need to contextualise to the study. In light of this conclusion, Vasileiou et al. (2018) advance that the comments that reviewers make on the adequacy and relevant samples sizes must take this into cognisance. Contradictions also surround the size of the sample (big or small) (Roy, Zvonkovic, Goldberg, Sharp, & LaRossa, 2015; Sim et al., 2018; Vasileiou et al., 2018). Sample sizes justification must always endeavor to bring an equilibrium with that of other components of data collection, to reduce the dominance of size over other important cogitations such as complexity of the topic and broadness or narrowness of research scope and questions as well as the quality of fieldwork (Morse, 2000; Roy et al., 2015). Emmel (2013:154) re-affirms Morse (2000)'s argument stating that "it is not the number of cases that matters, it is what you do with them that counts"

2.1 How to Sample in Qualitative Research

Random sampling is unfeasible with qualitative research. In affirmation Moser and Korstjens (2018) asseverate that "in qualitative research, you sample deliberately, not at random". The researchers suggest that qualitative samples are generally small but the smallness varies with studies. The sample sizes resolutions are informed by conceptual needs not essentially by representativeness (Moser & Korstjens, 2018). It is therefore imperative that researchers always describe fully and justify their sampling resolutions (Dworkin, 2012; Guetterman, 2015; Vasileiou et al., 2018). The predominantly used purposeful sampling strategies for qualitative research are theoretical, criterion, purposive, convenience and snowball sampling methods (Moser & Korstjens, 2018:10; Rosenthal, 2016:511). These are presented in Table 1 below. The participants are chosen because of their essential knowledge on the phenomenon.

Table 1. Qualitative Research Sampling Strategies and Their Explanations

Sampling Strategy	Explanation
Purposive Sampling	Choosing of participants guided by the researcher's personal judgement, considering the informative nature or "information power" of study participants. The choice is normally influenced by experience and expertise, institutional memory, participant knowledge and the relevance to the study, purpose of the study as well as the research questions.
Criterion	Participants are chosen relying on a pre-determined criteria of importance.
Theoretical	The selection of participants is influenced by emerging findings, this is to ensure adequacy in dealing with theoretical concepts that are fundamental to the research.
Convenience	Sampling of participants is dependent on availability and willingness to participate.
Snowballing	Participant choice is informed by referrals by initially selected participants. For example, the first participant refers you to another more informed potential participant and the next refers you to the next until the sample is adequate.
Maximum Variation	Selection of participants based on a broad range of variations in the backgrounds of these participants.
Extreme Case	Purposeful selection of the most uncommon cases.
Typical Case	Most typical and average participants are chosen.
Confirming and Disconfirming	Sampling that is aimed at supporting, corroborating or questioning of surfacing trends, perspectives or themes in the data

Source: Adapted from Moser and Korstjens (2018):10; Sebele-Mpofu (2020)

From the table above, it is clear that there are multiple ways that are available to qualitative researchers to select their samples and these have varying implications for sample size adequacy as a well as saturation attainment. For example, snowballing and purposive sample would ensure the researcher comes up with a study focused sample or a information power filled sample or one that is homogeneous and as established in submissions by Malterud et al. (2016) and Roy et al. (2015), smaller samples would be more appriopriate to achieve saturation. Maximum variation sampling on the other hand points to a diverse population suggesting most likely a need for bigger samples in order to reach satution point. Convinience sampling and saturation could point to a different status quo altogether. This points to the fact that researchers have to explain the sampling technique used , the rationale behind the choice, the appropriateness of the sample and its implication to the saturation concept in validating their sampling decisions.

2.2 Can the Conundrum be Resolved? Which One is the Adequate Sample Size, Big or Small?

Sampling techniques should be guided by the objective of collecting information that is reconcilable with the methodological angle adopted (Moser & Korstjens, 2018). Sample size sufficiency must be anchored on the "appropriateness of the composition and size" as these feed into the assessment of "quality and trustworthiness" in qualitative research (Vasileiou et al., 2018:2). Disagreements often arise among researchers regarding how big or how small a sample must be in order to be considered adequate (Guest et al., 2006; Roy et al., 2015), yet other researchers suggest that, the focus must not be on the numerical aspect of how big or small. Hammersley (2015) suggests that it is not about how many informants but which informants are included in the sample and the ability of the sample to generate

adequate data (information power) (Malterud et al., 2016) as well to provide comprehensive information or rich data (Vasileiou et al., 2018).   Arguments and varied suggestions have been tabled on what constitutes an adequate sample size in qualitative research. For example, Boddy (2016) adduces that one interview might be sufficient, Ritchie (2003) recommends not more than 50 interviews, Francis et al. (2010) suggest that researchers must have an initial working sample and a stopping criterion sample (because sampling is an on -going process influenced by decisions that can be made during the conduct of the research process such as the need for further sampling for consistency, validation and confirmatory intentions). Qualitative samples have been criticised for being too small (Roy et al., 2015), yet others suggest that they indeed must be small and there is absolutely nothing wrong with that as it allows for deeper investigation of the phenomenon under study (Dworkin, 2012; Guetterman, 2015; Rosenthal, 2016). Frequently raised questions include: "How larger does my sample have to be?"(Dworkin, 2012:1319). Other common questions include, "How large should a qualitative sample have to be to viewed as adequate?" (Rosenthal, 2016). Qualitative researchers have to fight the desire to reason about the qualitative samples from the quantitative stand point (generalisability of findings) (Guetterman, 2015). The main objective for sampling in qualitative research is "about balancing the need to obtain a rich experiential description from participants without sacrificing the equal representation of experiences across the population of possible participants". Saturation can be the ideal tool to strike the much needed balance (Rosenthal, 2016:511). Saturation can only alleviate sampling subjectivity if the concept is explained fully and contextualised to the research at hand, otherwise failure to articulate fully the type of sampling addressed, the rationale behind the choice and how it was reconciled with sample sizes adequate, will add more to the criticism of sampling decisions in qualitative research.

2.2.1 Are Small Samples Better or Deficient?

Another popular question is, how many qualitative interviews is enough? (Bryman, 2012). The determination of the adequate sample size is vital because the use of unnecessarily large or small samples has an ethical impact. For example, on the one hand, with unjustifiably large samples, why misspend resources and time (for both interviewer and interviewee)? On the other hand, unreasonably smaller samples may not be sufficiently informative, findings might exhibit idiosyncratic or individualistic data and fail to meet the qualitative research quality criteria (transferability, credibility and trustworthiness among others) and this may prove to be just a waste of time, effort and resources (Francis et al., 2010:1230). Some researchers question the cogency and credibility of small samples even the qualitative researcher themselves (Vasileiou et al., 2018) yet others are in full concord with using small samples (Guest et al., 2006; Roy et al., 2015). According to Roy et al. (2015:249), "Bigger is not better, but neither is small, the N itself is not sufficient to judge the quality of study and the depth of analyses to be carried out". O'reilly and Parker (2013) expostulate that the defensibility of the quality of qualitative research to a considerable degree relates to sampling adequacy that should provide depth and maximum opportunity for the transferability of findings. Sandelowski (1995) asserts that researchers must strike a balance between huge and small samples. Samples must be big enough to allow the emergence of novel and "richly textured understanding of the phenomenon", but small enough to enable an in-depth understanding and analysis of the same phenomenon. Morse (2000) advocates for a limited number of participants especially where in-depth interviews are used. Bowen (2008) suggest that adequacy (Sufficiency and appropriateness) should not be viewed in the context of numerical perspective on whether the sample brings on the table people with the right competences who best understand and represent the research subject area. Sample sizes must be selected in line with the objectives and design of the study. Smaller samples can result in large volumes of data, due to the depth of information, knowledgeability of informants, repeated data gathering or even the use of multiple methods (Roy et al., 2015). Malterud et al. (2016) re-affirming the potency of small samples, expostulate that when using "information power" as a guideline for sampling fundamental factors to be assessed include: the purpose or objective of the study (the more constricted the aim, the smaller the sample needed), specificity of the sample (the more specific and focused the sample to the study objective, the limited the number of participants required), theoretical underlying assumptions steering the study (the foggy these are the greater the sample size required and the clearer they are, the smaller the requisite sample), the quality of dialogue (the highly knowledgeable   or if these are experts in the area, the richer the quality of engagements, the richer or 'nutritional' richer the discussion, the smaller the sample size needed) and strategy. A research aimed at accomplishing a deeper evaluation and building a complete picture of the subject of exploration at hand will suffice with smaller sample sizes (Malterud et al., 2016:1757). The information power argued by the researchers though intuitively persuasive, it has elements of subjectivity. For example, the quality of dialogue is influenced by other things such as the mutual understanding or "chemistry" between the interviewer and interviewee, interviewer's skills and experience and sometimes the quality of findings is dependent on the "intensity of field work" and comprehensiveness of analysis (Vasileiou et al., 2018). Boddy (2016:426) insists that "sample sizes as low as one can be justified" and "…can provide a new, deep and nuanced understanding of previously unexplored phenomena" (Boddy, 2016:428). The researcher concludes that an unreasonably large sample can impede purposeful, consequential and timely research including data exploration.

Comparing larger samples to smaller samples, Vasileiou et al. (2018) expostulate that in their findings, they established that qualitative researchers suggest that larger samples are sufficient in terms of allowing for gathering of thick and rich data capable of boosting generalisability of findings to a larger population. An argument they find questionable as generalisability of results was found to refer to "representational generalisation" linked to bigger samples not the 'generalisability' in relation to quality of qualitative research (transferability, credibility, completeness and dependability of findings). Qualitative researchers appeared to ignore that small samples can provide credible findings which give a complete picture, offering both converging and diverging views that are all essential in boosting content validity (Vasileiou et al., 2018). Confirming that small sizes are adequate as long as they are properly justified and give adequate data to justify the findings and conclusions drawn from the study, Rosenthal (2016) postulates that generalisability is not the key pillar of qualitative research but the key objective is to create a complete and intense comprehension of the meaning behind the experiences, views and behaviour of participants to allow for transferable, dependable and truthful results. Low (2019) avows that large samples have a statistical logic, for example, the testing for correlation among variables, quantitative focus and not qualitative one, therefore as long as the small sample allows for gathering enough data to explain all pivotal element of the phenomenon under study, the sample is adequate.

Interestingly, Vasileiou et al. (2018) table that in the bulk of the qualitative studies they reviewed, researchers were often apologetic for having chosen small samples and pointed to them as a limitation in their section of 'limitations to the study', even in times where they claimed to have reached saturation. Others described their samples as "small and problematic". According to the researchers, very few studies accepted the adequacy of their small samples despite showing recognition that their findings might be criticised or misunderstood by quantitative scholars and other readers for the small samples, even fewer seemed to convincingly defend their sampling resolution's adequacy on the qualitative research quality grounds. These could be: the use of smaller samples in order to capture the complexity of a phenomenon in-depth (Boddy, 2016; Guest et al., 2006) or the information comprehensiveness (Vasileiou et al., 2018) or the intensity of analysis or field work (Roy et al., 2015) or information power (Malterud et al., 2016). Samples sizes are sufficient as long as they are proper rationalisations for their adoption. Qualitative researchers must believe themselves first, show confidence in their decisions before others can find their work believable!

2.2.2 Are Bigger Sample Sizes Better and Richer?

In answering the above question, Roy et al. (2015) portend that "Bigger is not better…". Data adequacy depends on an assortment of factors not just on the size of a sample. The notion that bigger samples are better and richer than small samples is an unfortunate misconception in qualitative research, prompting Vasileiou et al. (2018) to ask two intriguing questions, too small for who? Bigger samples, Promise or Peril? More data does not translate into more or enough information to give a complete and deeper picture of a phenomenon (Mason, 2010) and a bigger sample of participants who are information deficient and have no understanding of the phenomenon is equally useless (competence and knowledge matter not just how many?)(Hammersley, 2015; Malterud et al., 2016; Mpofu, 2021a). Arguing against large samples, Roy et al. (2015) point to the intricacy in data analysis that could compromise meaningful appraisal of the data gathered, adequate answering of the research questions and full contextualisation of data analysis and quotations. Mason (2010) posits that samples must be large enough to capture the necessary diverse experiences but not too big to be repetitious and waste resources unwarrantedly. Sampling can be guided by the epistemological perspective, aims and objectives of the study. Malterud et al. (2016) situate the need for bigger samples better, when they expound on their information power argument. They explain that the broader the objectives of the study, the larger the sample required, the less specific the sample to the research questions and focus of the study, the bigger the requisite sample and the poor the quality of dialogue, the greater the sample size needed. In short, the lower the information power of participants the bigger the sample required. Kuzel (1992) advance that the degree of homogeneity of the sample will affect the size, explicating that for homogeneous population 6 to 8 interview participants will suffice. A view shared by Guest et al. (2006) and Sim et al. (2018) who express that for homegeneous participants, small samples are ideal. Therefore for relatively uniform groups a sample of above 12 focus groups or more than 30 in-depth interviews would be rather too large and would need to be rationalised (Roy et al., 2015). Heterogeneous population will require larger samples (Malterud et al., 2016; Sim et al., 2018). It is important to always contextualise the decisions on sample sizes in order to avoid the misconceived need for large samples even where they are not necessary. As outlined by Roy et al. (2015:250), contrary to the assumption that larger samples will bring on board adequate data to answer research questions, "large samples may make it difficult to examine data in all their complexity, limiting the ability to probe data collection, develop emergent questions or contextualise quotes". Vasileiou et al. (2018) posit that, if the objective of the study is to explore and comprehend a wide variation within a certain phenomenon, a bigger heterogeneous sample may be ideal. Large samples may help discover new themes and unforeseen experiences. Roy et al. (2015) suggest that large samples might then require the rules applicable to quantitative research to be employed, for example, statistically representativeness of the samples, though they admit that at times smaller samples might lack opportunity for repetition of themes yet this is pivotal to theoretical

and thematic saturation.

2.2.3 What Is the Appropriate Answer to the Questions: Which One Is more Adequate? How Big or How Small?

In light of the above argument, perhaps the appropriate answer to the above questions is as rightly professed by Baker and Edwards (2012): "it depends". In consonance with the response, Dworkin (2012:1319) considers the response "vague (and actually reasonable)". Boddy (2016), stressing that this decision is not purely a straight forward one based on numerical issues alone, but one that needs to be well thought and informed, announces that "the issue of what constitutes an appropriate sample size in qualitative research is only answerable within the context and scientific paradigm of the research being conducted. In a constructivist or in-depth qualitative research a single case can be highly instructive". The sample size resolution is indeed contingent upon several factors as discussed in 2.1.1 and 2.1.2. Key among them the breadth of the study (Baker & Edwards, 2012), information power of participants (Malterud et al., 2016), methodological assumptions (Morse, 2000) and the analysis strategy (Vasileiou et al., 2018). According to Boddy (2016:429) "In qualitative research, the determination of sample size is contextual and partially dependent upon the scientific paradigm under which the investigation is taking place. Positivism oriented research will accordingly be characterised by bigger samples and "criticism for small samples might be justified"(Boddy, 2016: 429)., as compared to a purely interpretivist inclined one whose concern is the development of a deeper understanding as opposed to broadening the findings to allow for generalisability.   For a research with an outright interpretivist inclination, disapproval of small samples might need to have compelling rationalisation. The size of the sample also hinges on saturation. It is clear that in evaluating sample sizes for adequacy, researchers should not view adequacy in terms of the number of participants, but in terms of whether enough quality and relevant information can be extracted from the chosen sample and also whether that information enables the research to achieve the saturation defined by the study to have been targeted.

2.3 The When to Sample Puzzle: Should It Be Done a Priori, as an Adaptive Process, During the Research Process or During Analysis?

The enigma that surrounds sample sizes is equally evident also when it comes to the question, when is the appropriate time to sample? Is Baker and Edwards (2012)'s "it depends" rejoinder still applicable and the most practicable one. Qualitative researchers find themselves grappling with the when question of whether to sample in advance, before the research is begins, during the course of the research or be guided by saturation attainment or perhaps make a summative decision during analysis, in view of the divergent submissions by various researchers (Baker & Edwards, 2012; Blaikie, 2018; Boddy, 2016; Sim et al., 2018). As articulated by Boddy (2016:427) "a premeditated or in advance sample size selection approach is not wholly congruent with some of the principles of qualitative research", yet Blaikie (2018) asseverates that the increased evaluation of social science research by a variety of committees for proposal acceptance, funding needs and ethical clearance seem to point to the need to estimate samples a priori. Acknowledging the puzzle and the lack of prudence in sampling with finality a priori, Saunders et al. (2018), considers sampling a priori, erroneous and Sim et al (2018) characterise it as "illogical".   In affirmation, Blaikie (2018) describes such a decision as unfortunate and ill-advised. The researcher points to the fact that notwithstanding that it is possible to define a research problem, explicate the research purpose, raise research questions a priori to the actual research process implementation, these resolutions may be further developed and refined throughout the research process, therefore conclusively estimating sample sizes before the research kick starts is unsound. They explain their standpoint as follows:

"The characteristics of whatever is being sampled is usually unknown at the outset and in fact, the description of them maybe a study's primary purpose. The question of what the necessary application of the researcher's creativity in this process will produce, clearly cannot be predicted" (Blaikie, 2018:636)

  The views are also shared by Sim et al. (2018:619) who express that "determining qualitative sample size a priori is an inherently problematic approach especially in more interpretive models of qualitative research". Despite the observation, the researchers go on to table four possible ways of responding to the increasing requests for a priori sample size resolution and these are through: "rules of the thumb, conceptual models, numerical guidelines from empirical studies and statistical formulae" (Sim et al., 2018:619). Blaikie (2018) considers these "practical and sensible" under the circumstances. These are reviewed in detail and evaluated below for their potency, pros and cons. The question is whether apriori sampling, the best acceptable and compatible way to do in it qualitative research? Francis et al. (2010) calls for two types of sample, the initial criterion sample that progresses with the research process and context to a final criterion sample that takes into cognisance the research as an iterative process and saturation attainment needs. In concord, Moser and Korstjens (2018:10) expostulate that   in qualitative research "the sample will emerge during the study: based on further questions raised in the process of data collection and analysis, inclusion and exclusion criteria might be altered, or the sampling sites might be changed". There are clearly disagreements on the appropriate time to sample, as it emerges that the requirements of other stakeholders often indirectly compel researchers to sample a priori, yet those decisions come with a lot of problems that can be adequately answerable during the research process or analysis. It is important for

researchers to be clear on when they sampled and why, as well as how they mitigated the challenges raised by researchers in the review.

2.3.1 A Priori Sample Size Selection Methods and Their Cogency

Blaikie (2018) states that where researchers are confronted with the urgent and pragmatic requests by reviewers, ethics committees and funders to estimate sample size prior to research, the above suggestions by Sim et al. (2018) might be the best tools to resolve the conflict. Sim et al. (2018) avow that despite it being ill-advised sampling a priori might bring some standardisation and objectivity in sampling or approval of studies by ethical review committees.

   a) **Conceptual models**.

These models are normally guided by theoretical models and conceptual models that are contingent upon certain aspects of the study, such as perhaps the aim and objectives of the study, the guiding theoretical perspectives or theories, as well the anticipated nature of the analysis (perhaps qualitative or quantitative analysis). Malterud et al. (2016) states that "theoretical frameworks offer models and concepts that may explain the relations between different aspects of empirical data in a coherent way". Several conceptual models have been tabled by the researchers to guide estimation of samples (Baker & Edwards, 2012; Malterud et al., 2016; Morse, 2000; Vasileiou et al., 2018). According to Morse (2000), the scope of the study is one the vital factors (the narrow the scope the lower the sample size needed), as well the complexity of the research topic (the less complex it is, accordingly the less the sample size required) and the potency of the data (the poor the quality of that, the greater the need for a bigger sample size). Vasileiou et al. (2018) provide evidence that some qualitative researchers justified their sample size selection based sampling guide lines suggested by other researchers in line with the scope and nature of the research (conceptual models).(Malterud et al., 2016) propose a conceptual model based on what they term "the information power" of the sample. This was explained comprehensively in Section 2.1.1 of this paper. In short, a sample with high information power will suffice with lower numbers of participants whilst that with low information power will consequently need greater number of participants to achieve data adequacy (Malterud et al., 2016:1757). Morse (2015:588) assert that "Qualitative Samples are relatively small, they must be adequate (large enough for replication to occur and be noted) and appropriate (those interviewed must be experts in the phenomena of interest. Blaikie (2018) recommends a consideration of ontological assumptions and logics of enquiry (grouped in to inductive, deductive, abductive and retroductive logics as outlined in Blaikie (2007)). The researcher argues that sample sizes must fully reflect, the logics of enquiry, their ontological perspectives, their assumptions and their implications, failure of which discussions on sample sizes and adequacy will remained marred with controversy and uninformed. Sampling needs of research garnering for "data collection and theory testing" should be very distinct from one that is focusing on both data gathering and theory building (Blaikie, 2018). Adding merit to the argument, while acknowledging the diversity of samples and the use of saturation as a gold standard of quality in qualitative research, Guest et al. (2006:60) advance that interviews based on concepts and constructs that have been pre-determined by through existing theories or theoretical frameworks will attain data adequacy and saturation with small sample sizes. "Ensuring that there is enough data is a precursor to credible analysis and reporting" in any research (Marshall, Cardon, Poddar, & Fontenot, 2013:11). Sim et al. (2018) also underscore the need to consider philosophical, epistemological, ontological and methodological assumptions, their congruence as well as well as compatibility with analysis strategies to be employed. The researchers were also quick to apprise that generic determinants of the study by themselves are not enough, the how, to what extent an effectiveness or feasibility of questions to elicit enough information from the preliminary sample can only be fully adjudged during the study, hence these broad features can only be used as embryonic or "springboards" for initial sampling (Sim et al., 2018:628; Bryman, 2012:19).

   b) **Rules of the thumb**

Rules of the thumb entail choosing a sample with justification and guidelines suggested by previous researchers and their prior methodological considerations as well as the researcher's previous experience with regards to the data and volume required for studies of the same nature. The major criticism of these is the absence of transparency and persuasive reasoning (Sim et al., 2018). The similarity in studies is no guarantee that if the samples were adequate for one study they will be enough for the other. This assumption ignores the diversity and uniqueness of studies, their context and situational factors as these are context and time dependent. The samples can also vary because of the design of the studies (grounded, phenomenology, ethnographic, case study, content analysis) as well as whether in-depth interviews are used or focus groups. It also depends on the data collection instruments (questionnaires or interviews and in them also the nature of the questions, close ended or open ended). Creswell (2014) advance 3-10 interviewees in phenomenology, Kuzel (1992) 6-8 participants where the groups is homogenous and Boddy (2016) contends that   one case is enough in an interpretivist study seeking a deeper understanding of phenomena and in a wider population with diverse group one participant per each group is adequate. The researcher argues for a bigger sample where the research

has a positivism inclination.

c) Numerical guidelines deduced from empirical studies

Sim et al. (2018) and Vasileiou et al. (2018) suggest that these are the most commonly proffered as justifications for sample selection or where sampling has been tabled first as the guiding factor when sampling, saturation achievement was often rationalised on the basis of the sample sizes proposed by researchers such as Guest et al. (2006) and Hennink et al. (2017) on the number of interviews needed to achieve saturation or Hennink et al. (2019) and Guest et al. (2017) on the number of focus groups necessary to achieve saturation in qualitative research. The broadest criticism for these numerical guidelines is for overlooking conceptual, ontological, thematic aspects and other important aspects of qualitative researcher such as theoretical underpinnings, pragmatic and practical considerations such as time and resources as well as the quality of data analysis. Reiterating this concern, Vasileiou et al. (2018) suggests that where researchers want to rely on previous methodological studies and "sample size community norms", they must not adopt these blindly. They must appraise them for data adequacy, sampling sufficiency and the possibility of saturation accomplishment based on the characteristics that are "intrinsic to the research at hand" (Vasileiou et al., 2018): In consonance, Malterud et al. (2016) posit that, basing the methodological choice, sampling techniques, sizes and justification on the work of previous researchers in the subject area or similar studies context has two major shortcomings: (1) the presupposition of an untrue uniformity or consistency among studies even under similar methodological and comparable evaluative angles. (2) The likelihood of exaggerating or overemphasising the analogue in the studies to the point of overlooking their individuality and peculiarity that comes with the circumstances and the contexts (Malterud et al., 2016): (developed or developing country contexts, national variations or even the time. Sim et al. (2018) emphasise that "Ignoring the unique features of a particular study goes against the contextual orientation of most qualitative studies". At times researchers forget the shortcomings highlighted by the previous researchers or they fail to explain how they mitigated the impacts of the methodological limitations on their research. For example, in some cases previous researchers would have acknowledged the shortcomings and compromised usefulness of their methodology's application to future studies or even in for researches in which the context, topic and target population varies. It is without doubt that contextualisation is mandatory for any justification leaning on previous research methodology in order to preserve the uniqueness of studies and their settings

d) Statistical decisions and formulae

Statistical formulae has been tabled and used by a number of researchers in qualitative studies to estimate sample sizes especially with the motive of addressing saturation achievability concerns (Saunders et al., 2018; Sim et al., 2018). These researchers include: (1) Fugard and Potts (2015) who employ a binominal distribution to estimate the minimum acceptable sample size to achieve a predetermined confidence level given the frequency in occurrence of a particular theme across the focal population (2) Van Rijnsoever (2015) who utilises simulation as a guideline for selecting minimum sample sizes or "information sources" that a crucial in attaining saturation in qualitative research. The researcher works with three clear outlined situations, the random chance, minimal information and maximum information. What is key for the success of this model is the certainty on the number of codes as well the likelihood of observing those codes. (3)Tran, Porcher, Tran, and Ravaud (2017) applied mathematical models to establish the saturation point in qualitative research studies. Sim et al (2018) refers to the other researchers who adopted statistical models whose works were not reviewed for this study, Galvin (2015) on binominal distribution and Tran, Porcher, Falissard and Ravaud (2016) on Monte Carlo Simulations on open ended survey question answers. The statistical models towards attaining saturation can be criticised for the quantification of qualitative research and can be often challenging for qualitative researchers to employ and justify comprehensively.

2.3.2 Saturation and Adaptive Sample Size Selection

As highlighted by several researchers on the discussion on a priori sampling that such a move is groundless (Saunders et al., 2018), ill-advised or misguided (Blaikie, 2018:635), " illogical and…incompatible with the conceptual methodological notions underpinning qualitative research" (Sim et al., 2018:630):, not "wholly congruent with some of the principles of qualitative research"(Boddy, 2016:427) and that it must be used for estimating preliminary working samples that will change as the research progresses, taking into cognisance any necessary adjustments to the sample brought about by research question, data adequacy and saturation point considerations (Blaikie, 2018; Morse, 2000, 2015; Rosenthal, 2016). Saturation conceptualisation has been proffered by various researchers as probably the most reliable tool to address the sampling conundrum in qualitative research and perhaps to address the subjectivity in the findings and quality in qualitative research(Guest et al., 2020; Guetterman, 2015; Majid, Othman, Mohamad, & Lim, 2018).

In emphasising the criticalness of saturation in qualitative research, researchers have given the term varying description. For example, "the most frequently touted guarantee of qualitative rigor offered by authors to reviewers and readers" (Morse, 2015:587), "the gold standard by which purposive samples are determined"(Guest et al., 2006:60), "the

qualitative research 'gold standard' for quality research (Hancock et al., 2016:2125) and the point of "information redundancy" (Sandelowski, 2008:875) or "diminishing returns" from continued or further sampling (Rowlands, Waddell, & McKenna, 2016). Affirming the fundamental impact of saturation in qualitative research and perhaps reflecting on the intriguing description by the various researchers, Fusch and Ness (2015:1408) adduce that "failure to reach data saturation has an impact on the validity of the research conducted and hampers content validity. Vasileiou et al. (2018), when investigating how qualitative researchers described and justified their sample selection and sizes in qualitative research, found out that saturation topped the list of justifications, followed by pragmatic considerations. Saturation is considered as a reliable tool to progressively and adaptively estimate sample sizes in qualitative research but the concept is not without disputation. It is surrounded with a lot of controversy in relation to underlying assumptions (Malterud et al., 2016; Morse, 2000), "operationalisation and conceptualisation" (Saunders et al., 2018:1893), forms of saturation (Glaser & Strauss, 1967; Guest et al., 2020; Hennink et al., 2019; Saunders et al., 2018), methodological guidelines (Marshall et al., 2013), applicability (Constantinou, Georgiou, & Perdikogianni, 2017; Majid et al., 2018) and how it is described and communicated in qualitative research (Mason, 2010;Morse, 2015; Vasileiou et al., 2018). Saturation is considered a crucial edict and an "increasingly ubiquitous discourse" to help standardise and may be iron the hot issue of quality in qualitative research and the diversity in opinions and frameworks on sampling and sample sizes suggested by various researchers (O'reilly & Parker, 2013:190). On the other hand, Low (2019:131), censures the definition of saturation from the angle of "no new information" emerging as simply "a logical fallacy", arguing that new information is bound to continuously crop up as long further interviewing takes place. So what really is Saturation? Before discussing how saturation can be used to estimate sample sizes in qualitative researcher, Table 2 below gives a summary of the areas of controversy that researchers have highlighted.

2.3.3 Saturation Contention

The applicability of saturation is questioned by some researchers, Marshall et al. (2013:11) express that "Data saturation is an elusive concept and standard in qualitative research since few concrete guidelines exist". O'reilly and Parker (2013:190) query "the unquestioned acceptance of the concept of saturation and consider its plausibility and transferability across all qualitative approaches". Arguing from the angles of transparency and epistemological assumptions, the researchers argue that applying saturation as a blanket or all-encompassing "quality marker is inappropriate". Some of the controversies are tabulated in Table 2 below.

Table 2: Areas of Controversy on the Concept of Saturation

Area of Controversy/Dissension	Explanations	Researchers
Definitions	• The point where no additional data to develop properties of categories is being found. • The instance where "no additional codes are emerging" or where no additional insights arise from the data. • The juncture where there are increasing occurrences of the same, but no new ones. • Further coding becomes impracticable, no emerging information, codes or themes from additional interviews. • The point where there is enough information to depict the theoretical framework guiding study.	(Glaser & Strauss, 1967):61 (Hennink et al., 2017):2015;(Hennink et al., 2019) (Urquhart, 2012):194 (Fusch & Ness, 2015) (Saunders et al., 2018)
Methodological guidelines	Very minimal, vague, absent, unclear, controversial and varying. How to sample and when: purposive, snowballing, criterion or other qualitative sampling techniques or use statistical models (Comparative Method for theme saturation, simulation or binominal distribution) or conceptual models	(Guest et al., 2020; J. M. Morse, 2015) (Moser & Korstjens, 2018; Rosenthal, 2016; Sim et al., 2018; Vasileiou et al., 2018)
Application	• Diversity in qualitative research designs (phenomenology, ethnography, content analysis, meta-analysis, conversation analysis and grounded studies). No all-inclusive saturation point. • Definition of a theme "superficial at its best", Are the themes identified or developed, do we measure saturation on theme occurrences or the ability of the themes to enable meaningfully development of evaluative arguments. • What is actually saturated in particular? Is it themes, codes, data, meanings or theories?	(Blaikie, 2018; Fusch & Ness, 2015; O'reilly & Parker, 2013; Rosenthal, 2016) (Braun & Clarke, 2016, 2019; Low, 2019; Morse, 2015; Saunders et al., 2018; Van Rijnsoever, 2015) (Constantinou et al., 2017)
Forms	Theoretical, thematic, data, code and meaning saturation	(Glaser & Strauss, 1967; Guest et al., 2006; Guest et al., 2020; Morse, 2015;

		Rowlands et al., 2016; Saunders et al., 2018)
Attainment	• Measuring saturation only on the non-emergence of new codes is shallow, meaning saturation is appropriate and occurs when standpoints, divergent viewpoints, clear and detailed appreciation of information is evident in the data.	(Hennink et al., 2017; Hennink et al., 2019)
	• Achieved when phenomenon is evidently stronger, more clearer, stable and deeper as well as when findings intrigue and excite the reader	(Morse, 2015:588)
	• When there is some level of generalisability in findings.	(Boddy, 2016)
	• All the attributes and perspectives of concepts and constructs as well as their relatedness is fully explained and incorporated in their intricacy and diversity in data.	(Sandelowski, 2008:875)
	• All the peculiar angles that characterise, shape, explain and alter phenomenon data are completely captured and reflected in the data.	(Roy et al., 2015)
Reporting	Saturation not properly explained, articulated, communicated, described, reported or published in studies or at times just referred to without elaboration as to what, why, how and when answers	(Guetterman, 2015; Morse, 2015; Vasileiou et al., 2018)
Sample sizes	• How many interviews? • How many focus groups • What is the adequate sample size to reach saturation?	(Dworkin, 2012; Guest et al., 2006; Guest et al., 2017; Hennink et al., 2017; Hennink et al., 2019; Mangioni & McKerchar, 2013)

Source: Own compilation from numerous researchers

2.3.3.1 How to Apply Saturation in Estimating Sample Sizes

In light of the dissensions highlighted in Table 2 above and some that haven't been captured, such as the variation in qualitative data collection methods that can be used, their complexities as well underlying propositions, for example, questionnaires, interviews, focus groups, observations and literature review (Fusch & Ness, 2015; O'reilly & Parker, 2013; Saunders et al., 2018), challenges on the practicable adoption of the term are evident. More visible is the gap between theoretical suggestions and the practical application. (Marshall et al., 2013). As explained earlier, despite the contention, it continues to be topical and highly regarded as an importance measure in qualitative, so how and when to apply it to effectively determine samples in qualitative research continues to be imperative (Blaikie, 2018; Braun & Clarke, 2016, 2019; Morse et al., 2014). Table 3 below gives a summary of the forms of saturation and the point of times in the research processes when they can be applied as adapted from Saunders et al. (2018).

Table 3. Saturation Forms and Suggested Application points in the Research Process

Form	Suggested point of application	Justification	Researchers/studies
Thematic	Sampling and Data Analysis	A priori sampling can be done with pre-determined themes in mind so that data fully reflects them or these can emanate during analysis both on the angle of numerical and meaning, thus further or no further sampling needs can be judged at this stage thus determining sample size.	(Bowen, 2008; Constantinou et al., 2017; Urquhart, 2012).
Theoretical	Sampling and Data Analysis	Depending on whether the research aims for theory building or testing, sample sizes might differ and also a priori sampling can target participants with relevant information to build or test the theory. Where the theory constructs are pre-determined lower samples could suffice and larger samples for theory building.	(Blaikie, 2018; Glaser & Strauss, 1967; Guest et al., 2020; Morse, 2015)
Data	Data Collection and Data Analysis	Data replication and information redundancy or point of diminishing returns, further interviewing or data analysis produces no new data other than the one previously found.	(Guest et al., 2020)
Code	Sampling and Data Analysis	Pre-determination of codes can enable a priori sampling and determining them during analysis call for adaptive sample size judgement.	(Hancock et al., 2016; Hennink et al., 2017; Hennink et al., 2019)
Meaning	Throughout the conduct of the research (a priori sampling, data	Quality, relevant, thick and rich data can be collected when preliminary samples are appropriately selected to address meaning saturation, altered and finalised during data analysis when holes are poked in the data or the need for further	(Hennink et al., 2017; Hennink et al., 2019; Van Rijnsoever, 2015)

	gathering and analysis	sampling for confirmatory or further clarification is established	

Source: Own Compilation from Saunders et al. (2018) and other researchers.

3. Research Methodology

This research employed an empirical question in order to guide the methodology, guide the search and selection of relevant studies and to orient the subsequent analysis and synthesis of the findings from literature (Mpofu, 2021b). Several issues cause dissension on the quality and rigor among qualitative studies, but this study specifically sought to address sampling as one such issue. Consequently, the review focused on aspects that cause controversy in sampling such as what constitutes an adequate sample, how big or how small, when is the appropriate time to sample, a priori, posteriori or during data analysis and also on issues, how does the choice of sampling techniques (convenience, purposive and snowballing among others) affect saturation attainment. Literature review was used in the quest to answer the question, Can saturation help solve the sampling conundrum and alleviate subjectivity in sampling decisions? This question was crafted after reading through various papers on sampling and saturation (Saunders et al., 2018; Sim et al., 2018; Vasileiou et al., 2018) and especially on the recommendations for further research from these studies.

3.1 Literature Searches and Inclusion and Exclusion Criteria

Databases such as Scopus, Web of Science, Google Scholar and others are recommended for searches for related literature (Snyder, 2019; Wee and Banister, 2019). For this study, the preliminary studies such as the ones cited above had been found from the Google Scholar database after it had been systematically searched using the phrases "sampling and saturation in qualitative research", "how to address saturation when sampling in qualitative research", "sampling in qualitative research" and "sample size adequacy in qualitative research among others". The researcher's interest was provoked by these articles hence the construction of the research question. The researcher then searched more on the same database with phrases like "sampling controversy in qualitative research", the sampling paradox in qualitative research, "Saturation" and "saturation as a measure to alleviate subjectivity in sampling in qualitative research". The two searches raised 42 papers and only 20 papers were found relevant to work with, but bearing in mind that Wee and Banister (2016) proposed a sample range of between 30 to 100 papers in a review study, the researcher used "snowballing" (Jalali & Wohlin, 2012) or "citation mining" (Padron, 2018) to source for more publications. Snowballing was undertaken in its two forms: backward and forward snowballing. The former refers to selecting studies based on the reference list of the preliminary identified studies from data searches in the case of this study. The identified relevant studies that had been 'mined' from the reference list based on their titles were searched for in Google scholar and their abstracts read and the appropriate ones added to the study sample. The latter describes a situation where citations to the earlier identified papers were also assessed for relevance to the current study based on their titles, abstract and key words (Jalali & Wohlin, 2012; Mpofu 2021b;Wee & Banister, 2016).This was vital in enabling the researcher to further widen her knowledge by getting access to the recent works on the concept, by establishing who else is researching on the subject and what was their focus. Backward searching allowed the researcher to gain a deeper understanding of the origins of the saturation concept especially as it relates to the works of early researchers on it such as Morse, 1995, 2000 and Guest et al. (2006), thus expanding the researcher's knowledge on the phenomenon. Though the research to a greater extent reviewed studies that were more recent (Blaikie, 2018; Boddy, 2016; Guest et al., 2020; Morse, 2015), it was because, the saturation and sampling paradox is a topical subject that is being explored by contemporary researchers. A handful of old studies that were available were also used in the background of the study (introduction) in order to give a contextual setting for the study and highlight the origins of the concept by extracting them from the 'references of the references'. In total the study reviewed 50 papers.

3.2 Extraction of Information, Analysis and Synthesis

Transparency in the research process is key in any review process. Having identified the relevant studies for this review, these were reviewed, analysed and synthesised following a thematic analysis approach. The themes and relationships that have been identified earlier in presenting and discussing literature in Section 2 were used to discuss the findings in a chronological manner. Braun and Clarke (2006) recommend the use of thematic analysis in qualitative research as the best way to identify, analyse, synthesise and communicate results using themes and subthemes within textual information. The results were presented in the form of a series of tables showing the authors, the key discussion and areas of dissension or agreement. The idea was to use the summarised tables to collate data and present it in a clear, cohesive and concise manner (Mpofu, 2021b; Snyder, 2019). This made it easier to make the findings more transparent for readers to understand and compare the broader results and different views from different scholars with regards to sampling. For example, the various techniques used in sampling and the challenges they can impose on the saturation concept. Textual discussions were also used to buttress the data presentation and analysis.

4. Discussions and Recommendations on Adaptive Sample Estimation, Saturation and Researchers'

Recommendations

What is evidently clear is that there are diverse forms of saturation and researchers must accordingly begin by understanding their own research in terms of scope, context, breadth of the research question, the complexity of the topic and the research design to be used (Morse, 2000) and then re-visit the forms of saturation and consider which one is more applicable. The researchers can then decide on their sample sizes, whether larger or sample and perhaps begin with an a priori sample that seeks to take care of saturation considerations on a preliminary basis. Numerical guidelines and statistical models earlier discussed could be used to estimate these working samples but the key is to ensure that with proper consideration of the research, designs, their underlying presumptions, POEM assumptions the sample chosen will help achieve saturation. Other important factors could be the 'information power' of the sample suggested by Malterud et al. (2016) which relates the breadth of the study aim, exactitude or particularity of the sample, theoretical origins as well as the envisaged data analysis strategy as these would determine whether a larger or smaller sample is required to achieve saturation.   In shedding more light on the saturation and sample size selection paradox, Dworkin (2012) expostulate that saturation as a tool for estimating sample sizes should be thoughtfully evaluated as it leans on a variety of factors some of them that the researcher has no influence over:

"Some of these include: How homogeneous or heterogeneous is the population being studied? What is the selection criteria? How much money is in the budget to carry out the study? Are there key stratifiers (e.g. conceptual, demographic) that are critical for the in-depth understanding of the topic examined? What is the time line the researcher faces? How experienced is the researcher in being able to even determine when he or she has actually reached saturation? (Dworkin, 2012:1319).

Table 4 below outlines the different recommendations by different researchers, on how to comprehensively use saturation to approximate sample sizes in qualitative research.

Table 4. Different Recommendations by Various Researchers Saturation and Sample Size Selection

Studies	Recommendations
(Malterud et al., 2016; Sim et al., 2018)	Sampling to be considered an iterative and ongoing process that needs balancing and realignment throughout the research process. A priori sampling ill-advisable but can be considered a preliminary step in sample size determination.
(Constantinou et al., 2017)	Recommends the use of CoMeTS can be used in estimating sample sizes in homogeneous populations and that saturation point depends on the order of interview analysis, re-ordering and re-analysis of interviews may yield varying saturation points though the deviations is 2 to 3 interviews up.
(Blaikie, 2018)	Researchers must be explicit on the ontological suppositions and the logic of enquiry adopted ad how sampling adequacy and saturation fit in.
(Fusch & Ness, 2015; Saunders et al., 2018)	Saturation should be clearly "conceptualised and operationalised, forms identified and defined and the appropriate form applied must be fully reported as to the why, how and when aspects. Researchers must be clear in defining saturation and how it was achieved.
(Francis et al., 2010; Marshall et al., 2013; Mason, 2010; Vasileiou et al., 2018)	The development of evidence driven guidelines for estimating sample sizes as researchers fail to account for saturation or other just describe it without detailed explanations. Researchers must strive to report saturation in light of the context and uniqueness of their studies. Saturation should be given a blanket application.

Source: Author's compilation from various sources.

5. Conclusions

The researcher sought to explicate the controversy is sample size selection in qualitative research and to expound on how saturation can be used to progressively estimate samples and alleviate the subjectivity in the whole process. The researcher also sought to enrich the discussion on sample size selection, in view of the concerns of such discussions being superficial and unenlightened as suggested by various researchers. Through a comprehensive literature review of publications on qualitative sample selection and saturation the paper gave a contextual background to the sample size conundrum and explicated how saturation can aid in its resolution .The researcher concludes that researchers must estimate sample sizes in the context of their studies, rationalise them in light of their specificity and uniqueness and accordingly articulate, report and publish saturation in the setting of their studies. In evaluating sample sizes, reviewers and funders must pay attention to the background and conditions of the study and not to denounce sample sizes with preconceived notions. There is neither a one size fit all saturation or a universal sample size or sample size selection approach. Justification will vary in relation a variety of factors and these need to be understood before criticism can be unleashed. Lastly qualitative researchers have to believe themselves and convince themselves first on the appropriateness

and adequacy of their samples before others can be convinced and find the work believable. In order to be believable, qualitative researchers should contextualise, explain their sample choices, rationalise these on qualitative grounds and report them clearly without being apologetic for their choices. Being a review study, the paper had some limitations. As highlighted in the methodology in section 3, the paper employed both database search and snowballing as methods of identifying the studies to be included in the review, database searches allowed for a wider pool of papers to be accessed but it was time-consuming and demanding especially the formulation and reformulation of search phrases as at times it was vital to broaden them or narrow them where necessary. Snowballing on the other hand provided the study with more focused papers as these were 'mined' from the reference list, backwards and forward being guided by their tittles, but considering titles were considered when mining the papers, there is a likelihood that some relevant papers with irrelevant keywords in their tittles could have been ignored. Overall the combination of the two techniques was meant to strike a balance and increase the reliability and validity of the study. Further studies could investigate the topic using more critiquing methods such as critical reviews or quantitative approaches such as meta-analysis. It could also be interesting to explore the saturation-sampling paradox using an empirical study or focusing the sampling techniques available for use by qualitative researchers and how saturation can be employed in relation to each technique.

References

Baker, S. E., & Edwards, R. (2012). How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research.

Blaikie, N. (2007). *Approaches to social enquiry: Advancing knowledge*: Polity.

Blaikie, N. (2018). Confounding issues related to determining sample size in qualitative research. *International Journal of Social Research Methodology, 21*(5), 635-641. https://doi.org/10.1080/13645579.2018.1454644

Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal*. https://doi.org/10.1108/QMR-06-2016-0053

Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: a research note. *Qualitative research, 8*(1), 137-152. https://doi.org/10.1177/1468794107085301

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2016). (Mis) conceptualising themes, thematic analysis, and other problems with Fugard and Potts'(2015) sample-size tool for thematic analysis. *International Journal of Social Research Methodology, 19*(6), 739-743. https://doi.org/10.1080/13645579.2016.1195588

Braun, V., & Clarke, V. (2019). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, 1-16. https://doi.org/10.1080/2159676X.2019.1704846

Bryman, A. (2012). How many qualitative interviews is enough. *How many qualitative interviews is enough*, 18-20.

Constantinou, C. S., Georgiou, M., & Perdikogianni, M. (2017). A comparative method for themes saturation (CoMeTS) in qualitative interviews. *Qualitative research, 17*(5), 571-588. https://doi.org/10.1177/1468794116686650

Creswell, J. W. (2014). *A concise introduction to mixed methods research*. Sage Publications.

Dworkin, S. L. (2012). Sample size policy for qualitative studies using in-depth interviews: Springer. https://doi.org/10.1007/s10508-012-0016-6

Emmel, N. (2013). *Sampling and choosing cases in qualitative studies using in-depth research: A realist approach.* London: Sage. https://doi.org/10.4135/9781473913882

Francis, J. J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M. P., & Grimshaw, J. M. (2010). What is an adequate sample size? Operationalising data saturation for theory-based interview studies. *Psychology and Health, 25*(10), 1229-1245. https://doi.org/10.1080/08870440903194015

Fugard, A. J., & Potts, H. W. (2015). Supporting thinking on sample sizes for thematic analyses: a quantitative tool. *International Journal of Social Research Methodology, 18*(6), 669-684. https://doi.org/10.1080/13645579.2015.1005453

Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The qualitative report, 20*(9), 1408. https://doi.org/10.46743/2160-3715/2015.2281

Glaser, B., & Strauss, A. (1967). Thediscovery of grounded theory. *Strategies for Qualitative Research. Mill Valley: The Sociology Press*.

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal, 26*(2), 91-108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field methods, 18*(1), 59-82. https://doi.org/10.1177/1525822X05279903

Guest, G., Namey, E., & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *PloS one, 15*(5), e0232076. https://doi.org/10.1371/journal.pone.0232076

Guest, G., Namey, E., & McKenna, K. (2017). How many focus groups are enough? Building an evidence base for nonprobability sample sizes. *Field methods, 29*(1), 3-22. https://doi.org/10.1177/1525822X16639015

Guetterman, T. (2015). Descriptions of sampling practices within five approaches to qualitative research in education and the health sciences.

Hammersley, M. (2015). Sampling and thematic analysis: a response to Fugard and Potts. *International Journal of Social Research Methodology, 18*(6), 687-688. https://doi.org/10.1080/13645579.2015.1005456

Hancock, M. E., Amankwaa, L., Revell, M. A., & Mueller, D. (2016). Focus group data saturation: A new approach to data analysis. *The qualitative report, 21*(11), 2124. https://doi.org/10.46743/2160-3715/2016.2330

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: how many interviews are enough? *Qualitative health research, 27*(4), 591-608. https://doi.org/10.1177/1049732316665344

Hennink, M. M., Kaiser, B. N., & Weber, M. B. (2019). What influences saturation? Estimating sample sizes in focus group research. *Qualitative health research, 29*(10), 1483-1496. https://doi.org/10.1177/1049732318821692

Jalali, S., & Wohlin, C. (2012). *Systematic literature studies: database searches vs. backward snowballing.* Paper presented at the Proceedings of the 2012 ACM-IEEE international symposium on empirical software engineering and measurement. https://doi.org/10.1145/2372251.2372257

Kuzel, A. J. (1992). Sampling in Qualitative Research. In B. F Crabtree & W. L. Miller, *Doing qualitative research* (1st ed.). Sage, Thousad Oaks.

Low, J. (2019). A pragmatic definition of the concept of theoretical saturation. *Sociological Focus, 52*(2), 131-139. https://doi.org/10.1080/00380237.2018.1544514

Majid, M. A. A., Othman, M., Mohamad, S. F., & Lim, S. A. H. (2018). Achieving data saturation: evidence from a qualitative study of job satisfaction. *Social and Management Research Journal, 15*(2), 66-77. https://doi.org/10.24191/smrj.v15i2.4972

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: guided by information power. *Qualitative health research, 26*(13), 1753-1760. https://doi.org/10.1177/1049732315617444

Mangioni, V., & McKerchar, M. (2013). Strengthening the validity and reliability of the focus group as a method in tax research. *eJournal of Tax Research, 11*(2), 176.

Marshall, B., Cardon, P., Poddar, A., & Fontenot, R. (2013). Does sample size matter in qualitative research? A review of qualitative interviews in IS research. *Journal of Computer Information Systems, 54*(1), 11-22. https://doi.org/10.1080/08874417.2013.11645667

Mason, J. (2006). Mixing methods in a qualitatively driven way. *Qualitative research, 6*(1), 9-25. https://doi.org/10.1177/1468794106058866

Mason, M. (2010). *Sample size and saturation in PhD studies using qualitative interviews.* Paper presented at the Forum qualitative Sozialforschung/Forum: qualitative social research.

McKerchar, M. A. (2008). *Philosophical paradigms, inquiry strategies and knowledge claims: applying the principles of research design and conduct to taxation.*

Morse, J. M. (2000). *Determining sample size: Sage Publications Sage CA: Thousand Oaks, CA.*

Morse, J. M. (2015). *Data were saturated: Sage Publications Sage CA: Los Angeles, CA.*

Morse, W. C., Lowery, D. R., & Steury, T. (2014). Exploring saturation of themes and spatial locations in qualitative public participation geographic information systems research. *Society & Natural Resources, 27*(5), 557-571. https://doi.org/10.1080/08941920.2014.888791

Moser, A., & Korstjens, I. (2018). Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. *European Journal of General Practice, 24*(1), 9-18. https://doi.org/10.1080/13814788.2017.1375091

Mpofu, F. Y. (2021b). Review Articles: A Critical Review of the Pitfalls and Guidelines to effectictively conducting and

reporting reviews. *Technium Social Sciences Journal, 18,* 550.

Mpofu, F. Y. (2021a). Addressing the Saturation Attainment Controversy: Evidence from the Qualitative Research on Assessing the Feasibility of Informal Sector Taxation in Zimbabwe. *Technium Social Sciences Journal, 19,* 607-630.

O'reilly, M., & Parker, N. (2013). 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative research, 13*(2), 190-197. https://doi.org/10.1177/1468794112446106

Padron, K. (2018). *The literature Review.* Florida Atlantic University Libraries.

Rosenthal, M. (2016). Qualitative research methods: Why, when, and how to conduct interviews and focus groups in pharmacy research. *Currents in pharmacy teaching and learning, 8*(4), 509-516. https://doi.org/10.1016/j.cptl.2016.03.021

Rowlands, T., Waddell, N., & McKenna, B. (2016). Are we there yet? A technique to determine theoretical saturation. *Journal of Computer Information Systems, 56*(1), 40-47. https://doi.org/10.1080/08874417.2015.11645799

Roy, K., Zvonkovic, A., Goldberg, A., Sharp, E., & LaRossa, R. (2015). Sampling richness and qualitative integrity: Challenges for research with families. *Journal of Marriage and Family, 77*(1), 243-260. https://doi.org/10.1111/jomf.12147

Sandelowski, M. (1995). Sample Size in qualitative research. *Research in Nursing and Health, 18,* 179-183. https://doi.org/10.1002/nur.4770180211

Sandelowski, M. (2008). Theoretical Saturation. In Given, L. M. (ed) *The Sage Encyclopedia of Qualitative Research Methods* (2), 875-876. Sage, Thousand Oaks.

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., . . . Jinks, C. (2018). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity, 52*(4), 1893-1907. https://doi.org/10.1007/s11135-017-0574-8

Sebele-Mpofu, F. Y. (2020). Saturation controversy in qualitative research: Complexities and Underlying assumptions. A literature review. *Cogent Social Sciences, 6*(1), 1838706. https://doi.org/10.1080/23311886.2020.1838706

Sim, J., Saunders, B., Waterfield, J., & Kingstone, T. (2018). Can sample size in qualitative research be determined a priori? *International Journal of Social Research Methodology, 21*(5), 619-634. https://doi.org/10.1080/13645579.2018.1454643

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333-339. https://doi.org/10.1016/j.jbusres.2019.07.039

Tran, V. T., Porcher, R., Tran, V. C., & Ravaud, P. (2017). Predicting data saturation in qualitative surveys with mathematical models from ecological research. *Journal of clinical epidemiology, 82*, 71-78. e72. https://doi.org/10.1016/j.jclinepi.2016.10.001

Urquhart, C. (2012). *Grounded theory for qualitative research: A practical guide*: Sage. https://doi.org/10.4135/9781526402196

Van Rijnsoever, F. (2015). (I Can't Get No) Saturation: A Simulation and Guidelines for Minimum Sample Sizes in Qualitative Research. *Innovation Studies Utrecht (ISU) Working Paper Series, 15*(05), 1-25. https://doi.org/10.5465/ambpp.2016.12040abstract

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC medical research methodology, 18*(1), 148. https://doi.org/10.1186/s12874-018-0594-7

Wee, B. V., & Banister, D. (2016). How to write a literature review paper? *Transport Reviews, 36*(2), 278-288. https://doi.org/10.1080/01441647.2015.1065456