

Qualitative Research J

A practical example of using theoretical sampling throughout a grounded theory study.

Journal:	Qualitative Research Journal
Manuscript ID	QRJ-07-2019-0059
Manuscript Type:	Research Paper
Keywords:	Data selection, Grounded theory, Qualitative research methods, Theoretical sampling

SCHOLARONE™ Manuscripts

A practical example of using theoretical sampling throughout a grounded theory study

Abstract

Purpose: A discussion of the practical use of theoretical sampling as a method for selecting data that provides a richer and deeper understanding of the phenomenon being investigated.

Design/methodology/approach: Theoretical sampling is a well-known method in grounded theory studies to seek additional data based on concepts developed from initial data analysis. This method involves following where the data has led to expand and refine the evolving theory during the analytical process. However, there is a dearth of information detailing the practical steps needed to undertake theoretical sampling.

Findings: The authors used the theoretical sampling method in their study in four ways: asking additional interview questions and/or widening the scope of existing interview questions; recruiting participants with additional diversity of attributes within the same group; adding a new group of participants; and expanding research settings.

Originality/value: Theoretical sampling is a valuable and practical method for the purpose of addressing gaps in the data in qualitative research. When using theoretical sampling, it is essential to consider potential strategies for countering challenges that may arise. Practical recommendations are offered on the use of theoretical sampling during data analysis, for the purpose of achieving theoretical integration.

Keywords: grounded theory; data selection; qualitative research methods; theoretical sampling.

Introduction

Theoretical sampling is an essential feature in a grounded theory study, but there is little practical direction on how to implement this method effectively (Ennis et al., 2015, Neill, 2012, Slatyer et al., 2015). McCrae and Purssell (2016) found that the use of theoretical sampling is not consistent in grounded theory research and the process of how to implement theoretical sampling is often invisible in research publications. The authors advocate that it should be used, however warn that novice researchers can be misled if the process has not been adequately implemented or explained by the researchers. Inappropriately using theoretical sampling may lead researchers to collect 'unnecessary and conceptually thin data' (Charmaz, 2014). This paper provides an in-depth description of the application of theoretical sampling, as a method for collecting data, using the authors' grounded theory study as an exemplar. Although the study was conducted in Indonesia, the paper would be of benefit to all international researchers using or wishing to use theoretical sampling.

What is theoretical sampling and why is it important?

Identifying what data to collect can be a challenge for researchers undertaking grounded theory studies. Corbin and Strauss (2008) emphasise that additional data to be collected is based on the analysis of previous data. As data is analysed, questions arise and potential gaps in the data and/or the evolving grounded theory are identified. Theoretical sampling can help address these gaps and expand upon emerging concepts (Charmaz, 2014). During the process of data collection, the emerging theory guides the process of theoretical sampling by directing the researcher as to what data to collect to obtain further information (Glaser and Strauss, 1967). This process guides the researcher with not only 'who to talk to' but also 'where to go' and 'what to ask', to help address gaps in information.

Using theoretical sampling means the researcher is being responsive to the data and is, therefore, flexible to explore the depth of the concepts derived from the data and then decide what concepts will be sought next (Corbin and Strauss, 2008). During the process of data analysis, theoretical sensitivity, or 'insight into data', evolves (Corbin and Strauss, 2008). Being theoretically sensitive, the researcher is aware of, and able to show how the concepts are grounded in the data and are relevant to the evolving theory. Thus the researcher can triage if the concepts are important or not (Birks and Mills, 2015).

Theoretical sampling facilitates the delineation of a category that has properties, dimensions, variations and relationships (Corbin and Strauss, 2008). All collected data are

subjected to constant comparative analysis until saturation is achieved (Glaser, 1978) or when new codes are developed that do not contribute to an already established category (Birks and Mills, 2015). Constantly comparing the conceptual level of data during the analysis guides theoretical sampling, because it allows the researchers to raise questions and thus directs the ongoing data collection. The new data generated provides a richer and deeper understanding of the phenomenon, by comparing incident to incident to establish similarities and differences (Birks and Mills, 2015).

Theoretical sampling is used to explain and refine categories by building properties of the categories, not to randomly select populations or their representative distributions (Breckenridge and Jones, 2009, Bryant and Charmaz, 2007). Theoretical sampling helps to illuminate and make clear the variations, properties, dimensions and relationships between codes and categories for the purpose of developing a credible and authentic theory. A general rule that grounded theory researchers apply to test the efficacy of categories derived from the data is to implement theoretical sampling. "...[C]onducting theoretical sampling depends on having already identified a category" and theoretical sampling allows us to refine our theoretical categories (Charmaz, 2014). However, according to Birks and Mills (2015), as the concepts can emerge during early data collection and analysis, then theoretical sampling can occur even from the initial interview or data set. The researcher continues to seek further data driven by theoretical sampling, until all concepts are well defined and explained, categories are saturated and a theory explicated (Corbin and Strauss, 2008).

Overview of the research study

This grounded theory study generated a theory about the process of how Indonesian people learn about their diabetes. With the philosophical stances of symbolic interactionism and constructivism, a grounded theory methodology was used to answer the research question: "How do people living with diabetes mellitus in Indonesia learn about their disease?" The study was conducted in West Kalimantan, Indonesia and received ethics approvals from James Cook University, Australia and Universitas Indonesia. The theoretical sampling process was implemented throughout the study, as can be seen in Figure 1.

Figure 1, which has been informed by (Corbin and Strauss, 2008) and Charmaz (2014), provides a diagrammatic representation of the theoretical sampling method used in this study. There were three phases of data collection conducted over 16 months. During this time a total of 26 participants were interviewed across seven settings. The initial phase involved recruitment

of seven participants using purposive sampling, to provide insights that allowed exploration of the phenomenon of interest (Creswell, 2013, Palinkas et al., 2015).

PLEASE INSERT FIGURE 1 HERE.

Figure 1 shows that the use of theoretical sampling occurred during the first phase of data generation when the researchers undertook data collection/generation and analysis concurrently. In this phase, the researchers used four ways of deciding what other data to seek. These four ways will be explained later in this paper. The second phase of data generation was decided after the researchers analysed the entire data set from the first phase. Likewise, theoretical sampling was used during the second phase of data generation and when analysing the entire data set from this phase. Similar to the first phase, memos and concept maps helped the researchers in making decisions about seeking further data. Lastly, the third phase of data generation occurred. The selection of participants for the third data generation phase was guided by theoretical sampling in the second phase. However, the aim of theoretical sampling in the third phase was to confirm a storyline, as theoretical saturation was considered to have been achieved, and this premise was being evaluated by these final participants. A storyline was used to facilitate the refinement of the theory from this study. The storyline and concept map were distributed to participants in the third phase and participants were asked for feedback. They were asked to comment upon aspects of the theory such as resonance, cohesiveness and completeness.

The theory identified by this research study was entitled 'Learning, Choosing, and Acting: Self-management of Diabetes in Indonesia' (Author et al., 2019). The theory consists of five major categories that explain how people in Indonesia learn about their disease. The categories are: seeking and receiving diabetes-related information; processing received information; responding to recommendations; appraising the results; and sharing with others. A full explication of the theory can be found in Author et al. (2019).

How was theoretical sampling useful in this study?

Theoretical sampling helped to direct data collection in four main ways: (1) in asking additional interview questions and/or widening the scope of existing interview questions; (2) in recruiting participants with additional diversity of attributes within the same group; (3) in adding new groups of participants; and (4) in expanding research settings. Four examples will be shared, although there were many to choose from, to illustrate how the authors employed theoretical sampling in their research study. When mentioning the participants, pseudonyms have been

used throughout this paper.

Asking additional interview questions and/or widening the scope of existing interview questions

The first example is in relation to how theoretical sampling was used to ask additional interview questions. New interview questions were added, and/or the existing questions were expanded and improved upon following the constant comparative analysis of each interview. For instance, following the analysis of an interview with a nurse academic, based on his responses about his experiences of providing health education to people living with diabetes, a gap in the data was identified. The gap was pertaining to how people who were recipients of diabetic health education were responding to the information – for example, *did they use the information to make changes in their lives?* Also, a gap in the data existed about what materials or educational aids worked best to assist people with diabetes to understand their disease. Theoretical sampling helped to improve the initial question, which asked 'Could you explain your experiences in providing health education for people living with diabetes?' to include two follow- up questions, which were: 'How do you think people with diabetes respond after you provide information to them?'; and 'Could you tell me about what health education materials you use to help people understand their diabetes?' These questions were then asked when interviewing the next participants (i.e. clinicians).

Theoretical sampling helped with modifying and refining the interview questions and interviewing style that was adopted for forthcoming interviews (Glaser, 1978). The researchers were cognisant of Corbin and Strauss (Corbin and Strauss, 2008, Corbin and Strauss, 2015) who explained, what has been found from previous analysis should drive the questions for future interviews. Ongoing addition and widening the scope of interview questions occurred from the analysis of the very first interview, as well as during the analysis of all interviews throughout this current research study. Adapting interview questions was also employed in a study by Butler et al. (2018) as an approach to theoretical sampling by adding new interview questions or eliminating inappropriate questions in order to gain a thorough picture of the categories and their relationship to each other.

Recruiting participants with an additional diversity of attributes within the same group

The second example describes how after analysing one of the interviews, the principal researcher identified a gap in the data pertaining to a particular type of person that needed to be interviewed; this 'type' of person was not included in phase one of the study, when purposive

sampling was used. Charmaz (2014) claims that "initial sampling in grounded theory gets you started but theoretical sampling guides where you go". Data from an interview and a memo written after the interview, will be used to illustrate how theoretical sampling helped fill this gap.

During the constant comparative analysis of data, theoretical sampling guided decisions to interview people with diabetes who had varying characteristics, such as complications associated with their diabetes (i.e. diabetic foot ulcers). Also, people with diabetes based on such attributes as where they lived or received care were identified to be recruited. For example, participants were recruited from both rural and urban geographical areas as well as from different healthcare facilities: outpatient settings; public health centres; inpatient hospital settings; a private wound care clinic and nursing educational institutions.

The illustration below shows how theoretical sampling was used to recruit participants with additional attributes. The following quote is from Betty (pseudonym), who has been living with diabetes for 14 years. When she was sitting in the clinic's waiting room, another person started telling her about complications that could happen because of diabetes. Betty then started asking others if they had any complications.

I get this sweet urine disease [diabetes]. "Sweet urine disease will impact to anywhere", he said...can have wound [ulcer], can go kidney disease, can be the liver, having a large stomach. It can impact the knees and can result in rheumatics. (Betty: G1P4)

Being responsive to the data made the researchers realise, that out of all of the people interviewed so far in their study, not one of them had any known complications from their diabetes. This was a gap in the data. Below is an example of being responsive to the data that was recorded from the concept map created on 26th August 2016 (Figure 2) and the memo written on 5th October 2016. The thick black arrow in Figure 2 highlighted our deliberations about recruiting participants with diabetes who have secondary complications. The authors also used memos to guide their theorising, as evidenced in the following excerpt.

PLEASE INSERT FIGURE 2 HERE.

People living with diabetes that had developed complications could add valuable insight into how they engaged in education to promote their health and wellbeing. Questions were raised, such as were there specific resources available that helped people with complications and if so, what were they and how did people access them? Questions that highlight whether

people with diabetes were aware of their risk for developing complications and how to prevent complications were not specifically asked, but examples of associated questions asked of participants with complications secondary to diabetes included: Regarding the complication you have (e.g. diabetic foot ulcer), how did you find the information to manage it?; How do you manage your complications at home?; What do you want to know about your complications?; Do you know what do when you develop a complication (e.g. ulcer)? Are you worried about your complications?; Do you know how to prevent further complications developing?

Theoretical sampling made it apparent that people living with diabetes, who had developed complications, such as foot ulcers, eye sight problems or kidney disease, needed to be interviewed. Below is the memo written by the principal researcher after Betty's interview.

A person with diabetes can learn from others such as fellows or friends. While waiting in the waiting room, the patient actively asked other fellow patients sitting beside her. She looked at what happened to the others. She also asked what happened to them. They told her what happened. Here, she learned the complications of diabetes such as ulcers and kidney disease. She was also told by her fellow patients that diabetes is not curable (the prognosis). The patient learned about complications from other patients who already had complications. Does it worry her? Yes or no? How powerful was the proximity of patients with the complications for making them become aware or perhaps worry about their disease? Does it make them change their behaviour and act toward promoting health and wellbeing? This person has no complication related to diabetes such as a diabetic foot ulcer. I think a person with diabetes complication needs to be sought and interviewed. (Principal Researcher, Memo #49)

Adding new groups of participants

Following the first phase of data collection and analysis, a number of additional groups of participants and settings were identified to be interviewed. Additional groups of participants included: families of people with diabetes; other types of healthcare professionals, such as nursing students; and non-professional healthcare providers, such as an exercise instructor and *kader* (lay health worker). The additional information obtained from these new groups of participants helped to achieve data saturation and theoretical integration.

The following example of a participant excerpt and a concept map show how theoretical sampling was used to add a new group of participants, family members of people living with diabetes. Family members were not included as data sources in the purposive sampling done in the first phase of data collection. However, in the first phase, some participants mentioned family members as people who were making a contribution to the diabetes self-management of their relative living with diabetes.

Danur (pseudonym) is a nurse clinician who works at the hospital. He cares for people with diabetes who are experiencing the secondary complication of a diabetic foot ulcer. Below is the excerpt from Danur who mentioned the importance of family members in the management of diabetes.

For patients coming from the rural area, we [usually] communicate [the care] with the family member. Therefore, by involving the family [in the care], the patient can pay attention [concentrate] as the family accompanying him/her. Support from the family members is influential [important] for patients on long term medications who are not making progress. (Danur: G1P3,9)

The following diagram (Figure 3) is the sample of a concept map (created on 6th September 2016) that illustrated the need to add a further data source - family members.

PLEASE INSERT FIGURE 3 HERE.

Expanding research settings

The final example of how theoretical sampling drives data collection/generation describes how the researchers identified further settings (and thus further participants within those settings) to include in their study. In the original phase of the study, where purposive sampling was employed, the following health professionals were interviewed: a nurse academic; clinical nurses; pharmacist; dietician and a specialist doctor. The two nurses that were interviewed in phase one of the study, both worked in public outpatient clinics. It came to light during the interview with Cahyo (pseudonym), the pharmacist, that people with diabetes, who were admitted to hospital, were receiving education about their diabetes while in hospital.

Both my colleague and I provide education for patients in inpatient care. In this room here, we provide education and information about medicines to patients...and provide counselling if the patients ask for counselling. We provide counselling in inpatient care in the patient's room. The room is more private. When the patients plan their discharge, the pharmacist assistant will call the pharmacist to provide education. (Cahyo: G1P6)

Therefore, in phase two of the study nurses were selected from inpatient hospital settings and also from a private wound care clinic. This was to address the gap in the data as to what type of education was being provided by nurses in acute care facilities, such as a hospital, and whether there was a difference in the education being provided between the different healthcare facilities. The following memo was written by the principal researcher in response to the above excerpt from the pharmacist.

According to this clinician (the pharmacist), health education is provided to people in inpatient and outpatient settings. He said people discharging from the hospital are given health education. I have not interviewed people with diabetes who are hospitalised. I think they need to be asked about the health education they had. It is also important to interview a nurse clinician who takes care of hospitalised people with diabetes. Is there any differences or similarities in the education provided in outpatient and inpatient settings? What information do they usually deliver to these people? What questions do people with diabetes usually ask the nurse related to their disease? (Principal Researcher, Memo #40)

The use of theoretical sampling in this study was similar to what Butler et al. (2018) have applied, in that theoretical sampling was used for introducing new research sites, adapting interview questions and seeking new participant characteristics. However, our paper has explained theoretical sampling in a more practical way by showing how it is linked to and supported by analytical tools such as memos and concept maps. These tools facilitated the authors in identifying gaps as well as a decision process about how to seek further data.

Implications for practice

Theoretical sampling increased the awareness of the authors in this study, by allowing the authors to ask additional interview questions, specifically select additional participants, and expand research sites that provided them with more information to better understand the phenomenon the authors were researching. This deeper and more accurate understanding of the process, was then translated into a theory that explicates how people with diabetes in Indonesia learn about their disease, therefore providing the authors with an opportunity to make recommendations aimed at improving this process. However, theoretical sampling is not without its challenges. Two such challenges the authors faced were gaining access to this new data and having to obtain relevant ethical approval with each new phase, as new participants and settings were identified.

An example of having to gain access to new data came in the form of recruiting family members of people living with diabetes, to be interviewed. Theoretical sampling identified that family members played an important part in educating and supporting their loved ones who had diabetes. However, recruiting family members from which to generate this further data proved challenging. One strategy the authors implemented was asking previous and existing participants as well as key stakeholders if they knew of any family members that would be willing to be interviewed. This sampling method is commonly known as snowball sampling (Marshall, 1996). Sadler et al. (2010) and Faugier and Sargeant (1997) regard using social networks and the snowball method of sampling to find 'hidden' or 'hard-to find' participants,

as an effective method to help with further recruitment. Therefore, our strategy in overcoming this challenge was to use snowball sampling.

The other significant challenge was having to obtain ethical approval with each of the different data collection/generation phases of the research. Grounded theory studies are inductive (Birks and Mills, 2015), which means that the initial planning of the study may not cover all contingencies. It is important to state this feature in the initial ethics application. As the study evolved, it was necessary to submit amendments to the original ethics application to both the Universities' Human Research Ethics Committees, to cover further data collection, such as involving additional types of participants and additional sites. Two amendments were submitted; one for the second data generation phase and another for the third phase of data generation. The second and third phases of data generation did not commence until after each of the ethics amendments were approved. These additional steps need to be considered when creating project timelines and planning field trips.

Conclusion

Theoretical sampling is a distinctive feature of grounded theory studies. This paper has demonstrated some practical examples on how theoretical sampling was used in the authors' study to gain a better and deeper understanding of the topic area being researched. The paper can be a guideline for both higher degree students and their supervisors when selecting grounded theory as the research design to help them in further data selection. Also, the explanation of applying theoretical sampling in this paper is practical, feasible and useful in facilitating other researchers globally to effectively conduct a grounded theory study.

Notes

- 1. An in-depth description of theoretical sampling as a method for selecting further data using a grounded theory study as an example. A practical guideline of using theoretical sampling in four ways for qualitative researchers who are considering employing theoretical sampling.
- 2. When stepping from first phase of data generation to the second phase and the third phase, the ethics amendments were approved before the collecting of data (i.e. interviews) commenced.
- 3. The use of storyline in the third phase of data generation was the Indonesian language version translated from the English version.

Acknowledgement

The authors would like to thank all participants who willingly shared their experiences and stories, without which this research would not have been possible. Further acknowledgement details are withheld for the peer review process.

References

- Author and Others. (2019) Details Withheld for Peer Review.
- Birks, M. & Mills, J. 2015. Grounded theory: A practical guide, Thousand Oaks, CA, SAGE.
- Breckenridge, J. & Jones, D. 2009. Demystifying theoretical sampling in grounded theory research. *Grounded Theory Review*, 8, 113-126.
- Bryant, A. & Charmaz, K. 2007. *The SAGE handbook of grounded theory*, Los Angeles, SAGE Publications.
- Butler, A. E., Copnell, B. & Hall, H. 2018. The development of theoretical sampling in practice. *Collegian*, 25, 561-566.
- Charmaz, K. 2014. Constructing grounded theory, London, SAGE.
- Corbin, J. M. & Strauss, A. L. 2008. *Basics of qualitative research: Techniques and procedures for developing grounded theory,* Los Angeles, Sage Publications, Inc.
- Corbin, J. M. & Strauss, A. L. 2015. *Basics of qualitative research: Techniques and procedures for developing grounded theory,* Los Angeles, SAGE.
- Creswell, J. W. 2013. *Qualitative inquiry and research design: Choosing among five approaches*, Thousand Oaks, Calif, SAGE Publications.
- Ennis, G., Happell, B. & Reid-Searl, K. 2015. Enabling professional development in mental health nursing: the role of clinical leadership: Clinical leadership in mental health nursing. *Journal of Psychiatric and Mental Health Nursing*, 22, 616-622.
- Faugier, J. & Sargeant, M. 1997. Sampling hard to reach populations. *Journal of Advanced Nursing*, 26, 790-797.
- Glaser, B. G. 1978. *Theoretical sensitivity: Advances in the methodology of grounded theory,* Mill Valley, CA, Sociology Press.
- Glaser, B. G. & Strauss, A. L. 1967. *The discovery of grounded theory: Strategies for qualitative research*, Hawthorne, NY, Aldine de Gruyter.
- Marshall, M. N. 1996. Sampling for qualitative research. Family Practice, 13, 522-525.
- Mccrae, N. & Purssell, E. 2016. Is it really theoretical? A review of sampling in grounded theory studies in nursing journals. *Journal of Advanced Nursing*, 72, 2284-2293.

- Neill, M. 2012. Graduate-entry nursing students' journeys to registered nursing. *Nurse Education in Practice*, 12, 89-94.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N. & Hoagwood, K. 2015. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Administration and Policy in Mental Health and Mental Health Services Research*, 42, 533-544.
- Sadler, G. R., Lee, H.-C., Lim, R. S.-H. & Fullerton, J. 2010. Research Article: Recruitment of hard-to-reach population subgroups via adaptations of the snowball sampling strategy: Hard-to-reach populations. *Nursing & Health Sciences*, 12, 369-374.
- Slatyer, S., Williams, A. M. & Michael, R. 2015. Seeking empowerment to comfort patients in severe pain: A grounded theory study of the nurse's perspective. *International* Journal of Nursing Studies, 52, 229-239.

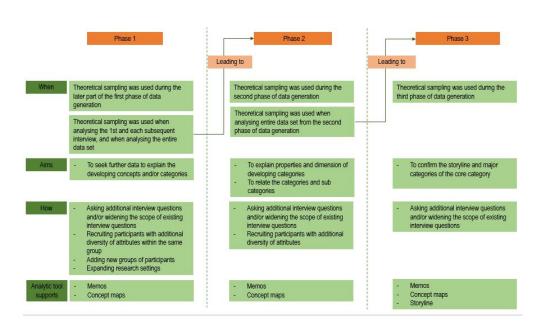


Figure 1. Using theoretical sampling method during three phases of the study 75x43mm (300 x 300 DPI)

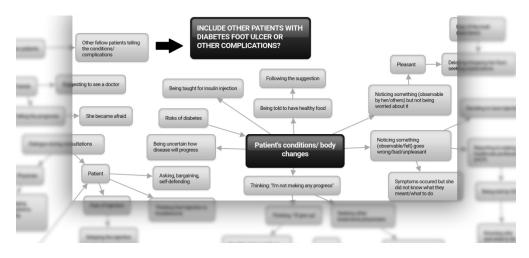


Figure 2. A concept map illustrating data from a person with diabetes (participant #2) $213x97mm (300 \times 300 \text{ DPI})$

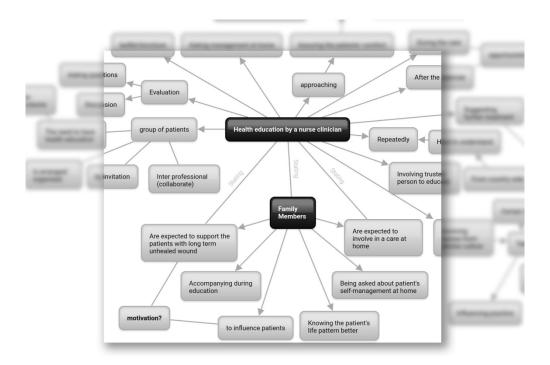


Figure 3. A concept map illustrating data from the nurse clinician (participant #4) $170 \times 115 \text{mm} \ (300 \times 300 \ \text{DPI})$