

# Sample Sizes for 10 Types of Qualitative Data Analysis: An Integrative Review, Empirical Guidance, and Next Steps

International Journal of Qualitative Methods

Volume 23: 1–14

© The Author(s) 2024

DOI: 10.1177/16094069241296206

[journals.sagepub.com/home/ijq](https://journals.sagepub.com/home/ijq)**Amber Wutich<sup>1</sup>, Melissa Beresford<sup>2</sup> , and H. Russell Bernard<sup>3</sup>**

## Abstract

There has been a recent explosion of articles on minimum sample sizes needed for analyzing qualitative data. The purpose of this integrated review is to examine this literature for 10 types of qualitative data analysis (5 types of saturation and 5 common methods). Building on established reviews and expanding to new methods, our findings extract the following sample size guidelines: theme saturation (9 interviews; 4 focus groups), meaning saturation (24 interviews; 8 focus groups), theoretical saturation (20–30+ interviews), metatheme saturation (20–40 interviews per site), and saturation in salience (10 exhaustive free lists); two methods where power analysis determines sample size: classical content analysis (statistical power analysis) and qualitative content analysis (information power); and three methods with little or no sample size guidance: reflexive thematic analysis, schema analysis, and ethnography (current guidance indicates 50–81 data documents or 20–30 interviews may be adequate). Our review highlights areas in which the extant literature does not provide sufficient sample size guidance—not because it is epistemologically flawed, but because it is not yet comprehensive and nuanced enough. To address this, we conclude by proposing ways researchers can navigate and contribute to the complex literature on sample size estimates.

## Keywords

saturation, sampling, sample size

## Introduction

There has been a recent explosion of articles on minimum sample sizes needed for analyzing qualitative data. Now is an excellent time for taking stock of the empirical literature on sample sizes for qualitative data analysis—and to critically assess what future research is needed to advance knowledge of minimum sample sizes. In this integrative review, we examine the literature on minimum sample size estimation for 10 types of qualitative data analysis: 5 types of saturation and 5 common methods. To begin, we draw on the large and fast-growing literature on saturation. We then discuss qualitative data analysis methods for which sample size estimation is rarely discussed and poorly understood. In doing so, we generate a taxonomy, conceptual framework, and research agenda for understanding minimum sample size that includes and goes beyond the literature on saturation. Our overall goal is to assess what we know, reflect on the extent that it helps us move forward, and anticipate what we will need to know in the future.

## Background

In 2006, [Guest and colleagues \(2006\)](#) published an article on estimating the minimum number of interviews that researchers need for studies based on qualitative data. The main finding of that article was striking: Theme<sup>1</sup> saturation—defined as identifying around 90% of themes in a set of qualitative interviews—could be reached within 6–12 interviews, an important addition to the literature on sample<sup>2</sup> size estimation for nonprobability samples. Advances in this field help researchers more accurately predict the

<sup>1</sup>School of Human Evolution & Social Change, Arizona State University, USA

<sup>2</sup>Department of Anthropology, San José State University, USA

<sup>3</sup>Institute for Social Science Research, Arizona State University, USA

### Corresponding Author:

Melissa Beresford, Department of Anthropology, San José State University, Clark Hall, #469, San Jose, CA 95112, USA.

Email: [melissa.beresford@sjsu.edu](mailto:melissa.beresford@sjsu.edu)



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons

Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use,

reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE

and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

sample sizes needed for research designs that use qualitative data, to better forecast the time and cost of field research, and to justify sample sizes in grant proposals.

Much of the work in this field prior to Guest et al.'s (2006) study focused on methods and parameters for estimating sample size *a priori* (Boddy, 2016; Malterud et al., 2016; Morse, 1995, 2000; Onwuegbuzie, 2003; Onwuegbuzie & Leech, 2007; Sim et al., 2018). Guest's team took a different approach: They analyzed data from interviews they collected in Ghana and Nigeria to determine *ex post facto* the minimum units of data needed to reach data saturation (or, as it is now called: theme saturation). This empirical approach was foreshadowed by Morgan et al. (2002) and builds on a long tradition in statistics, beginning with Condorcet's jury theorem (1785) and addressing the so-called wisdom of crowds (Dawes et al., 1977; Galton, 1907; Surowiecki, 2005).

With the impact of the finding from Guest et al. (28,000+ citations as of 2024), the empirical approach to determining data saturation is now the most common one (e.g., Constantinou et al., 2017; Francis et al., 2010), as seen in Hennink and Kaiser's (2022) influential systematic review of the saturation literature. Complementary approaches use statistical modeling or simulations to estimate the size of samples required to identify common and rare themes in purposive samples (Fugard & Potts, 2015; Galvin, 2015; Lowe et al., 2018; Tran et al., 2017; van Rijnsouwer, 2017) or to find examples of behaviors that occur with unknown or rare frequency (Altmann, 1974; Babbage, 1835; Bernard & Killworth, 1993; Gross, 1984; Tippett, 1935). Meta-analyses of the peer-reviewed research on sample sizes for qualitative data are also widely cited in the sample size estimation literature (Carlsen & Glenton, 2011; Marshall et al., 2013; Mason, 2010; Sobal, 2001; Thomson, 2011).

As consensus around the empirical approach solidified, debates in the literature on data saturation and sample size estimation shifted, centering on what "saturation" means and how to study it (Aldiabat & Le Navenc, 2018; Daher, 2023; Daniel, 2019; Hennink et al., 2017; Leese et al., 2021; Marshall et al., 2013; Mason, 2010; Naeem et al., 2024; Rahimi, 2024; Saunders et al., 2018; Sebele-Mpofu, 2020). Independent lines of analysis arose around five saturation concepts: *theme saturation*, *meaning saturation*, *theoretical saturation*, *metatheme saturation*, and *saturation in salience*. Each of these produces different estimates of minimum sample sizes (Hennink & Kaiser, 2022). The literature has focused less on *why* different kinds of saturation matter—i.e., what are the applications of these findings? Largely lost is the fact that theme saturation, the most commonly studied saturation concept (Blaikie, 2018; Vasileiou et al., 2018), has some significant limitations. For example, determining the minimum sample size for identifying themes does not necessarily mean that a researcher has enough data to do other kinds of analysis.

Additionally, many methodologists have indicated concern about widespread misunderstanding and misuse of guidelines for minimal sample sizes (Baker & Edwards, 2017; Blaikie, 2018; Bowen, 2008; Braun & Clarke, 2021; LaDonna et al., 2021; Marshall et al., 2013; Mason, 2010; Morse, 1995; Naeem et al.,

2024; O'Reilly & Parker, 2013; Sebele-Mpofu, 2020; Sim et al., 2018; Tight, 2023; Trotter, 2012; Vasileiou et al., 2018; Weller et al., 2018). Braun and Clarke (2021, p. 202) go so far as to argue that scholars should resist or reject neo-positivist-empiricist framings for data analysis (including the use of the saturation concept), and that it is inappropriate or impossible for qualitative researchers to make *a priori* estimates of sample sizes for qualitative research (Braun et al., 2022; Braun & Clarke, 2016, 2022b). We, however, do not think that misuses of the sample size estimation literature are occurring because it is epistemologically flawed, but rather because it is not yet comprehensive and nuanced enough. Like Onwuegbuzie and Collins (2007), we advocate addressing potential misuse and misapplication of minimal sample size guidelines with *more* empirical research.

## Approach

In this integrative review, we discuss sample size guidelines for 10 types of qualitative data analysis: 5 types of saturation and 5 common methods. We intentionally selected these 10 types of analysis from very different traditions and epistemologies in research with qualitative data. In doing so, we add a novel taxonomy and conceptual framework for non-methodologists to help navigate the complex literature on sample size estimation. In addition, we provide guidance for methodologists on how they can move this literature forward, including areas where existing research on minimum sample size estimation is insufficient or nonexistent. For each approach we review, we discuss its use, the minimum sample size recommendations from empirical research, and next steps for scholars seeking to advance this literature.

We followed well-established methods for conducting integrative literature reviews (Cronin & George, 2023; Torracco, 2005; Whitemore & Knafel, 2005). Unlike systematic reviews, integrative reviews are designed to go beyond systematic searches to discover and integrate literature that are not currently in conversation. We began with systematic searches of each of the 10 types of qualitative data analysis ("theme saturation," "data saturation," "metatheme saturation," "saturation in salience," "theoretical saturation," "reflective thematic analysis," "ethnography," "schema analysis," "qualitative content analysis," "classical content analysis") plus all possible variants of the search terms "saturation," "sample size estimation," and "minimum sample size." The results here characterize the results by method, as some searches yielded hundreds of articles with well-established methods reviews while other method searches yielded one or no articles. There are known weaknesses to systematic review methods, and systematic review methods may not be appropriate for reviews that combine concepts (Haddaway et al., 2020; Uttley et al., 2023). To ensure nothing was missed; we corresponded with expert scholars and added to the relevant literature. The purpose of an integrative review is to generate a taxonomy, conceptual framework, and research agenda (Torraco, 2005), and our approach reflects this.

We present the results of our integrative review in several sections. First, we discuss the five main saturation methods, explaining how each estimate for minimum sample size fits into a qualitative research design. Next, we discuss five exemplars of qualitative analysis, chosen to illustrate a range of approaches to sample size estimation. We begin with three examples of qualitative methods that currently have little or no sample size guidance—and explain what can be done to address this. Finally, we introduce two methods where the sample size for collecting qualitative data can be determined based on statistical power analysis or information power. The overall goal, as shown in the tables and figure below, is

to assess what we know, reflect on the extent that it helps us move different kinds of research designs forward, and anticipate what needs to be done to advance the field.

## Sample Size Estimates for 5 Types of Saturation

In this section, we discuss five types of saturation and how they are used to determine sample size adequacy. In doing so, we draw on the foundational contribution of [Hennink and Kaiser's \(2022\)](#) review paper, as well as a range of primary

**Table 1.** Five Types of Saturation: Empirical Recommendations for Minimal Sample Sizes.

Saturation	Purpose	Field(s) of Origin	Min. Sample Size	Key Citations	Next Steps: What's Needed	Next Steps: Example Research Questions
Theme (Data) saturation	Establish the number of interviews needed to identify a high number of themes	nursing, public health, anthropology	Interviews: 9 (mean: 12–13) Focus groups: 4 (mean: 5–6)	<a href="#">Hennink and Kaiser (2021)</a> , <a href="#">Hennink et al. (2017)</a> , and <a href="#">Guest et al. (2006, 2017)</a>	Empirical research on the effects of research topic and interview design on theme saturation.	When is theme saturation reached with interviews on different topics (e.g., health, class, race, gender, environment)?
Meaning saturation	Establish the number of interviews needed to capture all variance of meaning within themes	public health, allied health fields, medical anthropology	Interviews: 24 Focus groups: 8	<a href="#">Hennink and Kaiser (2022)</a> and <a href="#">Hennink et al. (2017, 2019)</a>	Conceptual research on how to define & measure meaning saturation. Empirical research on effects of study design on meaning saturation.	When is meaning saturation reached with outsider versus insider interviewers (e.g., matched language, gender, class & ethnic identity)?
Metatheme saturation	Establish the number of interviews needed to identify all cross-site themes in a cross-cultural or multi-sited dataset, as well as common within-site variants of each theme	anthropology, cross-cultural research	Interviews: 20–40 per site	<a href="#">Hagaman and Wutich (2017)</a>	Conceptual research on how to define & measure metatheme saturation. Empirical research on effects of research topic, interview design, and probing on metatheme saturation.	When is metatheme saturation reached in heterogenous versus homogenous cross-cultural samples?
Saturation in salience	Establish the number of interviews needed to identify the items or themes that are the most culturally important to interviewees	anthropology, medicine	Free lists: 10, if exhaustive	<a href="#">Weller et al. (2018)</a>	Empirical tests on how different styles of probing and prompting affect saturation in salience.	When is saturation in salience reached with exhaustive versus minimal probing?
Theoretical saturation	Establish the number of interviews needed to identify all major themes, their meanings and variants, and the inter-relationships among them	sociology	Interviews: 20–30+	<a href="#">Thomson (2011)</a>	Need meta-analysis of grounded theory studies that establishes sample sizes in studies that rigorously document theoretical sampling and saturation.	When is theoretical saturation researched in grounded theory designs that conduct theoretical sampling in batches versus selecting respondents one-by-one?

sources. While determining saturation is not an analytic outcome, we explain how the data processing that supports the assessment of saturation can yield results that are useful at different stages of research based on qualitative data (Table 1).

### Theme (Data) Saturation

Theme (or, data) saturation generally refers to identifying a high percentage of the themes that exist within a corpus of data. This concept was introduced initially as “information redundancy” (Lincoln & Guba, 1985), was advanced by Guest and colleagues (2006) as “data saturation,” and was widely adopted as “theme saturation” in later studies (Hennink & Kaiser, 2022). As such, theme saturation is the most commonly used saturation concept in the literature on estimating the minimum sample size for “qualitative research,” where that term is a euphemism for “statistically non-representative.”

**Uses.** Theme saturation is determined by identifying themes in a data set. If a researcher’s analytic goal is simply to produce a list of themes, then assessing theme saturation may be helpful. On a practical level, this kind of theme identification is very useful for codebook development or as an early step in a mixed-methods research design (e.g., to inform survey design). It is difficult, however, for a researcher to make a theoretical contribution based solely on such analyses because simply listing themes rarely provides enough information to advance a body of knowledge.

**Minimum Sample Size.** Nine interviews or four focus groups (Guest et al., 2006, 2017; Hennink & Kaiser, 2022), with the caveats that: (1) focus group size matters and (2) individual interviews often yield more data than focus groups (Fern, 1982; Schuster et al., 2023), especially for interviewers who are experienced in probing techniques (about which more below). Recent work suggests more interviews may be needed to reach saturation if the interviews are web-based, with estimates ranging from 30 to 67 web-based interviews (Squire et al., 2024). (But note: For developing surveys with cognitive interviewing, a minimum sample size of 55–75 is needed to identify the majority of high-impact survey problems [Blair & Conrad, 2011].)

**Next Steps.** Buckley’s (2022) paper proposes 10 easy-to-follow steps for theme saturation. And Mwita (2022) found that theme saturation can be affected by length (in minutes) of data collection, the use of mixed methods, and the relevance of respondents selected. Beyond this, there is value in understanding the range of sample sizes needed for theme saturation with different research topics (e.g., health, environment). For example, people are more forthcoming in focus groups on sensitive health issues, but less so for politically sensitive environmental or justice topics (Schuster et al., 2023). As well, bounded domains (“tell me about the routes you took when you evacuated from the flood”) may have fewer themes, while

unbounded domains (“tell me about your experience in walking from Guatemala to the U.S. border”) may have many themes (Weller et al., 2018).

The effects of these observations on theme saturation are not yet well tested empirically. More research is also needed with different purposive sampling designs (e.g., typical vs. extreme cases) and different data elicitation strategies (i.e., beyond individual interviews and focus groups). For example, further empirical research is needed on the number of interviews necessary to reach saturation for web-based interviews, dyadic/couple interviews and third-party-present interviews (Boeije, 2004), family history and life history interviews (Nelson, 2010), stakeholder interviews (Beresford et al., 2020), Photovoice interviews (Wang & Burris, 1997), and arts-based elicitation (Bagnoli, 2009). Such research will make valuable incremental contributions.

### Meaning Saturation

The concept of meaning saturation was introduced by Hennink et al. (2017) and refers to capturing all of the variance of meaning within a theme (Hennink et al., 2017, 2019). While theme saturation is about theme *presence*, meaning saturation identifies the point at which the researcher can explain the *depth, content, nuance, and variability* of that theme.

**Uses.** Unlike theme saturation and theoretical saturation, meaning saturation aligns more clearly with commonly used approaches to descriptive theme analysis. As a result, meaning saturation is a useful and widely applicable saturation concept for thematic research with qualitative data.

**Minimum Sample Size.** 24 interviews or eight focus groups (Hennink et al., 2017, 2019; Hennink & Kaiser, 2022).

**Next Steps.** Meaning saturation has not been the focus of as much methodological research as theme saturation and theoretical saturation. What’s needed next: more conceptual work on defining and measuring meaning saturation and more empirical research on the effects of probing and sampling strategy (such as homogeneous or maximum variability sampling) on meaning saturation. Because meaning saturation is likely the most useful saturation concept for theme analysis, this is where the most research is needed and where the greatest advances are to be made.

### Metatheme Saturation

Metatheme analysis was introduced as a type of cross-cultural metaphor analysis (Josephides, 1991), and later developed as a mixed-methods approach that derives overarching themes from the analysis of qualitative data using quantitative semantic network analysis of words or codes (Barnett & Danowski, 1992; Nolan & Ryan, 2000; Schnegg & Bernard, 1996; Wutich et al., 2014). Recent research has

established a rigorous qualitative approach to metatheme analysis, which uses qualitative comparisons of themes for identifying and comparing overarching themes from multi-sited or cross-cultural datasets (Beresford, 2021; du Bray et al., 2022; Schuster et al., 2022; SturtzSreetharan et al., 2021; Trainer et al., 2022; Wutich et al., 2021).

**Uses.** Metatheme saturation supports the identification of minimum sample sizes needed to create a list of the cross-site themes in a study, as well as to identify common within-site variants of each theme.

**Minimum Sample Size.** 20–40 interviews per site. Hagaman and Wutich (2017) found that all metathemes could be identified once within 20 interviews in each site, but that it took as many as 39 interviews to reliably identify metathemes (ensuring that each metatheme occurred at least three times per site) in each of their four cross-cultural research sites.

**Next Steps.** Empirical research is needed to examine how metatheme saturation varies depending on: research topic, number of research sites, data elicitation, and homogeneity/heterogeneity of purposive samples. Hagaman and Wutich (2017) found metatheme saturation was reached in as few as 18 interviews in a more heterogeneous sample, but it took up to 35–39 interviews to reach metatheme saturation in more homogeneous samples. This finding regarding metatheme saturation in heterogeneous versus homogeneous samples needs replication.

### Saturation in Salience

The concept of saturation in salience was introduced by Weller et al. (2018). It is commonly applied in the analysis of free-list data, a deceptively simple method. The analysis of free-list data was developed in cognitive psychology to examine how people hold on to, organize, and retrieve knowledge about things, like plants, animals, and occupations (Bousfield, 1953; Magaña & Norman, 1980) or about their own emotions (Yik & Chen, 2023). It was picked up by, and is a staple of, cognitive anthropology (Borgatti, 1994; Dengah et al., 2021; Romney & D'Andrade, 1964); it is now widely used in ethnobotany, consumer behavior, food science, and marketing.

**Uses.** Saturation in salience is used to identify the items or themes in a set of qualitative interviews that are the most culturally important to interviewees. Salience can be estimated for any set of qualitative data (e.g., illness narratives, migration histories, etc.), and can be used to identify the terms and topics needed to develop closed-ended survey questions in a mixed-method research design (Weller & Romney, 1988).

**Minimum Sample Size.** Ten free lists can produce 95% of salient items, if elicitation is exhaustive, with appropriate probing (Weller et al., 2018).

**Next Steps.** More research is needed on how different styles and amounts of probing in various kinds of open-ended interviews (free lists, life histories, etc.) impact when saturation in salience is reached (Brewer, 2002; Tran et al., 2017). Weller et al. (2018) found that extensive probing can yield more themes with fewer interviews and, conversely, saturation can be reached quickly with superficial interviewing using fewer probes. This may lead researchers to conclude falsely that the superficial (small-*n*, low-probe) data collection and the resulting theme analysis was adequate: a dangerous proposition.

### Theoretical Saturation

The concept of theoretical saturation originates in grounded theory and is the point when “no additional data” are found to develop the analytic categories’ properties (Glaser & Strauss, 2017 [1967]:61). In a grounded theory study, theoretical saturation is achieved through an iterative process of theoretical sampling (new respondents are selected to advance the development of a theory), data collection, coding, comparison, and model building until all concepts and linkages among them are described (Aldiabat & Le Navenc, 2018; Breckenridge & Jones, 2009; Charmaz, 2006; Glaser & Strauss, 2017 [1967]; Morse, 2004; Strauss & Corbin, 1990).

**Uses.** Theoretical saturation occurs when all of the themes identified are fully understood in all their variants, and inter-relations among themes are established. Theoretical saturation indicates a “stopping point” for sampling but can only be determined through an iterative process of data collection and analysis.

**Minimum Sample Size.** 20–30+ interviews recommended. Theoretical sample size cannot be determined *a priori* (Charmaz, 2006; Conlon et al., 2020; Glaser & Strauss, 2017 [1967]). However, meta-analyses provide evidence that samples sizes around 20–30+ have been sufficient to claim theoretical saturation across a range of grounded theory studies (Mason, 2010; Thomson, 2011).

**Next Steps.** Procedures for reaching theoretical saturation are extensively documented, though some grounded theorists have argued that sufficiently detailed documentation is lacking (Breckenridge & Jones, 2009; McCrae & Purssell, 2016). There are relatively few empirical studies that unpack the mechanics of reaching theoretical saturation and its implications for sample size adequacy (Butler et al., 2018; Kearney et al., 1995; Saunders et al., 2018; Tuckett, 2004). More empirical and comparative research is needed to determine how researcher expertise and research topic affect theoretical saturation (Morse, 2000; Thomson, 2011).

## Sample Size Estimates for 5 Types of Qualitative Data Analysis

In this section, we discuss five types of qualitative data analysis. For each, we review the state of the literature guiding sample size estimation. In most cases, there are no clear recommendations but there may be some calculations recommended based on a study's characteristics. In some cases, however, no guidance is available. These five methods are rarely put into comparison in discussions of saturation and

sample size adequacy, and this part of the review provides novel guidance for scholars working in research designs in which saturation is not appropriate (Table 2).

### Reflexive Thematic Analysis

Reflexive thematic analysis is an iterative process to identify themes in data and reflexively interpret thematic meanings (Braun & Clarke, 2022a, 2022b). The approach stems from constructivism: It assumes that the world, as we perceive it, is

**Table 2.** Five Types of Qualitative Data Analysis: Empirical Recommendations for Minimal Sample Sizes.

Qualitative Data Analysis	Purpose	Field(s) of Origin	Min. Sample Size	Key Citations	Next Steps: What's Needed	Next Steps: Example Research Questions
Reflexive thematic analysis	Iterative process to identify themes in data & reflexively interpret thematic meanings.	interdisciplinary, psychology	None, but use PRICE model to establish saturation	Braun and Clarke (2021) and Naeem et al. (2024)	Meta-analysis of reflexive thematic analysis studies could determine sample sizes that researchers deemed adequate.	What sample sizes are presented as adequate, and with what justification, in a meta-analysis of reflexive thematic analysis studies?
Ethnography	Describe culture, including its content and structure.	anthropology, sociology	Based on theme (data) saturation: 50-81 data documents (from interviews & observations) Based on cultural consensus analysis: 20-30 interviews	DiStefano and Yang (2024), Weller (2007), and Small (2009)	Meta-analysis of sample sizes used in ethnography.	What sample sizes are presented as adequate, and with what justification, in a meta-analysis of ethnographic studies?
Schema analysis	Identify mental or cultural models.	psychology, linguistics, cognitive anthropology	None.	Quinn (2005a, 2005b)	Meta-analysis of sample sizes used in schema analysis.	What sample sizes are presented as adequate, and with what justification, in a meta-analysis of schema analyses?
Qualitative content analysis	Systematic approach for categorizing and interpreting themes and contexts in qualitative data.	communication, sociology, psychology	Can be determined by information power.	Mayring (2000), Onwuegbuzie and Leech (2007), and Malterud et al. (2016)	More conceptual and empirical work is needed.	How can information power be calculated? How do different calculation methods impact (1) estimated sample sizes and (2) sample size adequacy for qualitative content analysis studies?
Classical content analysis	Systematic approach for coding and comparing qualitative data.	communication, political science	Can be determined by power analysis.	Krippendorff (2013) and Weller (2015)	None.	None.

a construction of human meaning and experiences that can never be objectively known. In other words, themes do not exist independently from the researcher who is identifying and interpreting them—and each researcher’s knowledge, biases, and positionality become central to their identification and interpretation of themes.

**Uses.** The analytic purpose of reflexive thematic analysis is to represent “situated and contextual interpretations of data” rather than to assume that inherent meaning can be discovered in a data set (Braun & Clarke, 2021).

**Minimum Sample Size.** None. Braun and Clarke (2022a) argue it is impossible to estimate sample size *a priori* for reflexive thematic analysis, but that an upper and lower limit for a sample range should be determined pragmatically based on the research goals and the time and resources available for the research. While some have argued that the concept of saturation is incompatible with reflexive thematic analysis, Naeem et al. (2024) have proposed a PRICE model—using steps for Perspectivation, Recapitulation, Integration, Crystallization, and Edification—to reach saturation for thematic analysis.

**Next Steps.** Reflexive thematic analysis is a widely used and cited method, and many researchers are affected by the absence of techniques to estimate sample size and to assess sample quality and adequacy. This puts these researchers at a disadvantage in competitions for scientific grant funding and during peer review in scientific journals and academic institutions. The PRICE model (Naeem et al., 2024) is a good first step toward theorizing the role of saturation in thematic analysis. To move forward, meta-analysis of reflexive thematic analysis studies could establish the “stopping points” at which researchers deemed sample sizes to be adequate.

## Ethnography

Ethnography is an umbrella term for the many methods used by social scientists to analyze and describe a culture (Ballesteros and Winthereik 2021; Ruth et al., 2022). While there is little consensus about what counts (or doesn’t count) as an ethnographic method, most writers on the subject agree on the importance of interviewing and participant-observation, a time-intensive data collection method in which researchers experience the phenomena under study. Here we follow Bernard (1994, 2017) in taking a big tent, mixed-methods approach to ethnography, which encompasses qualitative and quantitative approaches.

**Uses.** Ethnography is a research approach for analyzing and describing culture, defined as shared norms, knowledge, and behaviors. Different approaches focus on describing cultural components and areas of consensus or dissensus.

**Minimum Sample Size.** Guidance on sample sizes for studies based on ethnography comes from recent work by DiStefano and Yang (2024) on saturation in a mixed-methods ethnographic study with observations and interviews, and from 40 years of research on the cultural consensus model (Romney et al., 1986).

DiStefano and Yang (2024) found that it took 50 data documents (field notes from 33 observations and transcripts from 17 interviews) to identify 80% of the of the 2273 themes present in their study. To identify 90% of the themes, it took 81 data documents (field notes from 55 observations and transcripts from 26 interviews).

The consensus model is based on the observation that the more people agree about the content of a topic, the fewer people you need to understand and describe that topic. So, to discover the names of the days of the week in English, a sample size of four respondents is enough, but for research on lesser known or more disputed cultural topics, a larger sample size would be necessary. Overall, the cultural consensus literature indicates that 20–30 systematic interviews on a single cultural topic should be sufficient to adequately understand and describe that topic (Weller, 2007; Weller & Romney, 1988). Importantly, the cultural consensus method applies to any set of structured questions, such as questionnaire surveys and other formal interviews. It does not apply to free-lists, pile-sorts, or unstructured interviewing.

**Next Steps.** Naroll (1970) found that anthropologists who stayed in the field for at least a year were more likely to report on sensitive issues like witchcraft, sexuality, political feuds, and so forth, and that anthropologists who spoke the local language were more likely to report data about witchcraft than were those who didn’t. Empirical research is needed to confirm this finding on sensitive topics and to determine adequate sample sizes for a wider range of ethnographic methods—e.g., the number of key informants, unstructured interviews, or time in participant-observation necessary to describe a given cultural phenomenon. Meta-analysis would be a good first step for researchers looking to contribute to the literature on minimum sample sizes for ethnography.

## Schema Analysis

Schema analysis is a method for identifying mental models, cognitive simplifications, or cultural scripts people use to understand the world (D’Andrade, 1991, 1995; Rice 1980; Rumelhart, 1984; Schank & Abelson, 1977; Strauss & Quinn, 1997). This approach uses methods derived from anthropology, psychology, linguistics, and sociology. Quinn’s foundational work detailed the mechanics of schema analysis, which are designed to extract the underlying cultural rules for behavior, timing, order, and the like (Quinn, 1987, 1997; 2005a; 2005b, 2018).



corpus of data (Krippendorff, 2013). For such samples, power analysis is the appropriate technique for determining minimum sample size. Weller's (2015) accessible approach enables non-statisticians to calculate sample size using a power analysis.

*Next Steps.* None. The literature provides sufficient guidance for scholars working in a classical content analysis tradition. This is largely settled science.

### Qualitative Content Analysis

Originating in communications (Mayring, 2000, 2019, 2022), qualitative content analysis is a rule-based procedural approach with reliability checks. It was developed originally as a more interpretive and contextual approach to classical content analysis, though many variants have evolved (e.g., Hsieh & Shannon, 2005; Kuckartz, 2014).

*Uses.* Qualitative content analysis is a systematic approach to categorizing and interpreting themes and contexts in qualitative data.

*Minimum Sample Size.* Sample size for qualitative content analyses—including within-case and cross-case comparisons—can be calculated using an information power approach (Malterud et al., 2016, 2021). This is a set of guidelines for considering the breadth of research aims, size of the study population, availability of prior theory to help define the research, quality of data available, and the overall analysis strategy. Thus, larger samples are needed for broader aims, larger study populations, more exploratory projects, lower-quality data, and more cross-case designs. For example, a comparison of men and women would require double the sample size of a study of women or men alone.

*Next Steps.* The information power approach calls for a uniquely qualitative approach to power analysis (Onwuegbuzie & Leech, 2007), but there are not yet clear guidelines for how to calculate target samples. More conceptual and empirical work is needed.

### Discussion and Conclusions

Our goal here has been to take stock of current research on minimum sample sizes for 10 types of qualitative data analysis: 5 types of saturation and 5 common methods. We find a trend toward over-reliance on theme saturation (see Figure 1) and inadequate attention to the literature on meaning saturation, metatheme saturation, theoretical saturation, and saturation in salience. We have also identified three general directions for continued research on *when and why* different sampling approaches, respondent selection, and sample sizes are necessary to better clarify minimum sample sizes needed for the different types of saturation and analyses of qualitative data.

First, we need more conceptual work to better define and operationalize theme saturation, meaning saturation, saturation in salience, and metatheme saturation, and more empirical research on all kinds of saturation. For example, on the conceptual level: What are the implications of defining saturation exhaustively versus prioritizing attention to the most salient, common, or repeated meanings? On the empirical level: What are the implications of setting theme saturation at 80% versus 90% of themes or metatheme saturation at 3 versus 5 repetitions of meaning?

Second, more meta-analysis is needed to amass empirical guidance for sample sizes used in theoretical saturation (for grounded theory), reflexive thematic analysis, ethnography, and schema analysis. While these methods have few or no techniques available to empirically test sample size adequacy, meta-analysis of existing studies with adequate sample sizes helps identify priorities for research.

Third, more work is needed on how different sample sizes are best suited for different types of data (e.g., free lists vs. narratives vs. field notes), different research questions, and different stages of analysis.

There is not a single answer to “how many interviews are enough?” ... other than: “it depends” (Baker & Edwards, 2012). It depends on the number and prevalence of themes in the population, the percentage of total themes a researcher needs to identify, and the way respondents are selected (Lowe et al., 2018; van Rijnsoever, 2017; Weller et al., 2018). It depends on how much probing and prompting is done, and with what skill (Fern, 1982; Tran et al., 2017; Weller et al., 2018). And it depends on the type and quality of qualitative data being analyzed, the analytical goal of the research, and the epistemological assumptions of the researcher(s) (see Figure 1) (Morse, 1995, 2000). Empirical work that further clarifies these things will support the continued use of qualitative data in social science.

### Acknowledgements

We thank our students (2022-2024) in the NSF Cultural Anthropology Methods Program, ASU's Institute for Social Science Research workshops on Qualitative Data Analysis: The Basics, Grounded Theory, and Content Analysis; ASU course ASB 500 (Qualitative Data Analysis); and SJSU course ANTH 234 (Advanced Research Methods: Qualitative Data Analysis) for their help piloting earlier versions of Figure 1 and workshopping these arguments.

### Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded under the US National Science Foundation Cultural Anthropology Program grant (Award SBE-2017491) to the NSF

Cultural Anthropology Methods Program and an NSF CAREER grant to Melissa Beresford (Award BCS-2143766).

## ORCID iD

Melissa Beresford  <https://orcid.org/0000-0002-5707-3943>

## Notes

1. We define “theme” as a construct of meaning that is linked to expressions found in qualitative data (Ryan & Bernard, 2003).
2. We define “sample” as a subset of people, documents, or other units of analysis that yield data for analysis.

## References

- Aldiabat, K. M., & Le Navenec, C. L. (2018). Data saturation: The mysterious step in grounded theory methodology. *The Qualitative Report*, 23, 245–261. <https://doi.org/10.46743/2160-3715/2018.2994>
- Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, 49(3), 227–267. <https://doi.org/10.1163/156853974x00534>
- Babbage, C. (1835). *On the economy of machinery and manufactures* (4th ed.). Charles Knight.
- Bagnoli, A. (2009). Beyond the standard interview: The use of graphic elicitation and arts-based methods. *Qualitative Research*, 9(5), 547–570. <https://doi.org/10.1177/1468794109343625>
- Baker, S. E., & Edwards, R. (2012). How many qualitative interviews is enough? *National Center for Research Methods*. <https://eprints.ncrm.ac.uk/2273/> (accessed February 5, 2024).
- Baker, S. E., & Edwards, R. (2017). How many qualitative interviews is enough? Expert voices and early career reflections on sampling and cases in qualitative research. *National Centre for Research Methods Review Paper*.
- Ballesteros, A., & Winthereik, B. R. (Eds.) (2021). *Experimenting with ethnography: A companion to analysis*. Duke University Press.
- Barnett, G. A., & Danowski, J. A. (1992). The structure of communication: A network analysis of the international communication association. *Human Communication Research*, 19(2): 264–285. <https://doi.org/10.1111/j.1468-2958.1992.tb00302.x>
- Beresford, M. (2021). Rethinking entrepreneurship through distribution: Distributive relations and the reproduction of racialized inequality among South African entrepreneurs. *Journal of the Royal Anthropological Institute*, 27(1):108–127. <https://doi.org/10.1111/1467-9655.13432>
- Beresford, M., Jones, J. L., Bausch, J. C., Williams, C. F., Wutich, A., Porter, S., Quimby, B., Eaton, W. M., & Brasier, K. J. (2020). Third-party effects in stakeholder interviews. *International Journal of Qualitative Methods*, 19, Article 1609406920966482. <https://doi.org/10.1177/1609406920966482>
- Bernard, H. R. (1994). Methods belong to all of us. In *Assessing cultural anthropology* (pp. 168–179). McGraw-Hill.
- Bernard, H. R. (2017). *Research methods in anthropology: Qualitative and quantitative approaches*. Rowman & Littlefield.
- Bernard, H. R., & Killworth, P. D. (1993). Sampling in time allocation research. *Ethnology*, 32(2), 207. <https://doi.org/10.2307/3773773>
- 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>
- Blair, J., & Conrad, F. G. (2011). Sample size for cognitive interview pretesting. *Public Opinion Quarterly*, 75(4), 636–658. <https://doi.org/10.1093/poq/nfr035>
- Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal*, 19(4), 426–432. <https://doi.org/10.1108/qmr-06-2016-0053>
- Boeije, H. R. (2004). And then there were three: Self-presentational styles and the presence of the partner as a third person in the interview. *Field Methods*, 16(1), 3–22. <https://doi.org/10.1177/1525822x03259228>
- Borgatti, S. P. (1994). Cultural domain analysis. *Journal of Quantitative Anthropology*, 4, 261–278.
- Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *The Journal of General Psychology*, 49(2), 229–240. <https://doi.org/10.1080/00221309.1953.9710088>
- 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. (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. (2021). 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*, 13(2), 201–216. <https://doi.org/10.1080/2159676x.2019.1704846>
- Braun, V., & Clarke, V. (2022a). Toward good practice in thematic analysis: Avoiding common problems and becoming a knowing researcher. *International Journal of Transgender Health*, 24, 1–6. <https://doi.org/10.1080/26895269.2022.2129597>
- Braun, V., & Clarke, V. (2022b). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, 9(1), 3–26. <https://doi.org/10.1037/qup0000196>
- Braun, V., Clarke, V., & Hayfield, N. (2022). “A starting point for your journey, not a map”: Nikki Hayfield in conversation with Virginia Braun and Victoria Clarke about thematic analysis. *Qualitative Research in Psychology*, 19(2), 424–445. <https://doi.org/10.1080/14780887.2019.1670765>
- Breckenridge, J., & Jones, D. (2009). Demystifying theoretical sampling in grounded theory research. *Grounded Theory Review*, 8, 113–126.
- Brewer, D. D. (2002). Supplementary interviewing techniques to maximize output in free listing tasks. *Field Methods*, 14(1), 108–118. <https://doi.org/10.1177/1525822x02014001007>
- Buckley, R. (2022). Ten steps for specifying saturation in qualitative research. *Social Science & Medicine*, 309, 115217. <https://doi.org/10.1016/j.socscimed.2022.115217>

- Butler, A. E., Copnell, B., & Hall, H. (2018). The development of theoretical sampling in practice. *Collegian*, 25(5), 561–566. <https://doi.org/10.1016/j.colegn.2018.01.002>
- Carlsen, B., & Glenton, C. (2011). What about N? A methodological study of sample-size reporting in focus group studies. *BMC Medical Research Methodology*, 11, 1–10. <https://doi.org/10.1186/1471-2288-11-26>
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.
- Condorcet, M. D. (1785). Essai sur l'application de l'analyse à la probabilité des décisions. *Rendues à la pluralité des voix*. <https://gallica.bnf.fr/ark:/12148/bpt6k417181> (accessed February 17, 2023).
- Conlon, C., Timonen, V., Elliott-O'Dare, C., O'Keeffe, S., & Foley, G. (2020). Confused about theoretical sampling? Engaging theoretical sampling in diverse grounded theory studies. *Qualitative Health Research*, 30(6), 947–959. <https://doi.org/10.1177/1049732319899139>
- Constantinou, C. S., Georgiou, M., & Perdikiogianni, M. (2017). A comparative method for themes saturation (CoMeTS) in qualitative interviews. *Qualitative Research*, 17(5), 571–588. <https://doi.org/10.1177/1468794116686650>
- Cronin, M. A., & George, E. (2023). The why and how of the integrative review. *Organizational Research Methods*, 26(1), 168–192. <https://doi.org/10.1177/1094428120935507>
- Daher, W. (2023). Saturation in qualitative educational technology research. *Education Sciences*, 13(2), 98. <https://doi.org/10.3390/educsci13020098>
- D'Andrade, R. (1991). The identification of schemas in naturalistic data. In M. J. Horowitz (Ed.), *Person schemas and maladaptive interpersonal patterns* (pp. 279–301). The University of Chicago Press.
- D'Andrade, R. (2005). *Some methods for studying cultural cognitive structures*. Palgrave Macmillan US.
- D'Andrade, R. G. (1995) *The development of cognitive anthropology*. Cambridge University Press.
- Daniel, B. K. (2019). Student experience of the maximum variation framework for determining sample size in qualitative research. In *Proceedings of the 18th European conference on research methodology for business and management studies*. Wits School of Business.
- Dawes, R. M., McTavish, J., & Shaklee, H. (1977). Behavior, communication, and assumptions about other people's behavior in a commons dilemma situation. *Journal of Personality and Social Psychology*, 35, 1–11. <https://doi.org/10.1037//0022-3514.35.1.1>
- Dengah, H. F. II, Snodgrass, J. G., Polzer, E. R., & Nixon, W. C. (2021). *Systematic methods for analyzing culture: A practical guide*. Routledge.
- DiStefano, A. S., & Yang, J. S. (2024). Sample size and saturation: A three-phase method for ethnographic research with multiple qualitative data sources. *Field Methods*, 36(2), 145–159. <https://doi.org/10.1177/1525822x231194515>
- du Bray, M. V., Stotts, R., Wutich, A., & Brewis, A. (2022). Cross-cultural perceptions of (in) equity in ecosystem service access: A preliminary comparison of emergent themes in four sites. *Human Ecology*, 50(4), 793–798. <https://doi.org/10.1007/s10745-022-00334-4>
- Fern, E. F. (1982). The use of focus groups for idea generation: The effects of group size, acquaintanceship, and moderator on response quantity and quality. *Journal of Marketing Research*, 19, 1–13. <https://doi.org/10.1177/002224378201900101>
- 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>
- Galton, F. (1907). One vote, one value. *Nature*, 75(1948), 414. <https://doi.org/10.1038/075414a0>
- Galvin, R. (2015). How many interviews are enough? Do qualitative interviews in building energy consumption research produce reliable knowledge? *Journal of Building Engineering*, 1, 2–12. <https://doi.org/10.1016/j.jobee.2014.12.001>
- Garro, L. C. (2000). Remembering what one knows and the construction of the past: A comparison of cultural consensus theory and cultural schema theory. *Ethos*, 28(3), 275–319. <https://doi.org/10.1525/eth.2000.28.3.275>
- Glaser, B. G., & Strauss, A. L. (2017 [1967]). *The discovery of grounded theory: Strategies for qualitative research*. Routledge.
- Gross, D. R. (1984). Time allocation: A tool for the study of cultural behavior. *Annual Review of Anthropology*, 13(1), 519–558. <https://doi.org/10.1146/annurev.anthro.13.1.519>
- 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., & McKenna, K. (2017). How many focus groups are enough? Building an evidence base for non-probability sample sizes. *Field Methods*, 29(1), 3–22. <https://doi.org/10.1177/1525822x16639015>
- Haddaway, N. R., Bethel, A., Dicks, L. V., Koricheva, J., Macura, B., Petrokofsky, G., Pullin, A. S., Savilaakso, S., & Stewart, G. B. (2020). Eight problems with literature reviews and how to fix them. *Nature Ecology & Evolution*, 4(12), 1582–1589. <https://doi.org/10.1038/s41559-020-01295-x>
- Hagaman, A. K., & Wutich, A. (2017). How many interviews are enough to identify metathemes in multisited and cross-cultural research? Another perspective on Guest, Bunce, and Johnson's (2006) landmark study. *Field Methods*, 29(1), 23–41. <https://doi.org/10.1177/1525822x16640447>
- Hennink, M. M., & Kaiser, B. N. (2022). Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social Science & Medicine*, 292, 114523. <https://doi.org/10.1016/j.socscimed.2021.114523>
- 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>
- Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>
- Josephides, L. (1991) Metaphors, metathemes, and the construction of sociality: A critique of the new melanesian ethnography. *MAN*, 26(1), 145–161. <https://doi.org/10.2307/2803479>
- Kearney, M. H., Murphy, S., Irwin, K., & Rosenbaum, M. (1995). Salvaging self: A grounded theory of pregnancy on crack cocaine. *Nursing Research*, 44(4), 208–213.
- Krippendorff, K. (2013). *Content analysis: An introduction to its methodology*. Sage.
- Kuckartz, U. (2014). *Qualitative text analysis. A guide to methods, practice & using software*. Sage.
- LaDonna, K. A., Artino, A. R. Jr., & Balmer, D. F. (2021). Beyond the guise of saturation: Rigor and qualitative interview data. *Journal of Graduate Medical Education*, 13(5), 607–611. <https://doi.org/10.4300/JGME-D-21-00752.1>
- Leese, J., Li, L. C., Nimmon, L., Townsend, A. F., & Backman, C. L. (2021). Moving beyond “until saturation was reached”: Critically examining how saturation is used and reported in qualitative research. *Arthritis Care & Research*, 73(9), 1225–1227. <https://doi.org/10.1002/acr.24600>
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Sage.
- Lowe, A., Norris, A. C., Farris, A. J., & Babbage, D. R. (2018). Quantifying thematic saturation in qualitative data analysis. *Field Methods*, 30(3), 191–207. <https://doi.org/10.1177/1525822x17749386>
- Magaña, J. R., & Norman, D. K. (1980). Methodological inquiry into elicitation procedures: Cognitive mapping and free listing. *Perceptual and Motor Skills*, 51(3), 931–934. <https://doi.org/10.2466/pms.1980.51.3.931>
- Malterud, K., Siersma, V., & Guassora, A. D. (2021). Information power: Sample content and size in qualitative studies. In P. M. Camic (Ed.), *Qualitative research in psychology: Expanding perspectives in methodology and design* (pp. 67–81). American Psychological Association.
- 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>
- 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, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 11. <https://doi.org/10.17169/fqs-11.3.1428> (accessed February 5, 2024).
- Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1, 20. <https://doi.org/10.17169/fqs-1.2.1089>
- Mayring, P. (2019) Qualitative content analysis: Demarcation, varieties, developments. *Forum: Qualitative Social Research*, 20, 1–26. <https://doi.org/10.17169/fqs-20.3.3343>
- Mayring, P. (2022). Qualitative content analysis: A step-by-step guide. *Qualitative content analysis*. Sage.
- 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(10), 2284–2293. <https://doi.org/10.1111/jan.12986>
- Morgan, M. G., Fischhoff, B., Bostrom, A., & Atman, C. J. (2002) *Risk communication: A mental models approach*. Cambridge University Press.
- Morse, J. M. (1995) The significance of saturation. *Qualitative Health Research*, 5(2), 147–149. <https://doi.org/10.1177/104973239500500201>
- Morse, J. M. (2000). Determining sample size. *Qualitative Health Research*, 10(1), 3–5. <https://doi.org/10.1177/104973200129118183>
- Morse, J. M. (2004). Constructing qualitatively derived theory: Concept construction and concept typologies. *Qualitative Health Research*, 14(10), 1387–1395. <https://doi.org/10.1177/1049732304269676>
- Mwita, K. (2022). Factors influencing data saturation in qualitative studies. *International Journal of Research in Business and Social Science*, 11(4), 414–420. <https://doi.org/10.20525/ijrbs.v11i4.1776>
- Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2024). Demystification and actualisation of data saturation in qualitative research through thematic analysis. *International Journal of Qualitative Methods*, 23, Article 16094069241229777. <https://doi.org/10.1177/16094069241229777>
- Naroll, R. (1970). What have we learned from cross-cultural surveys? 1. *American Anthropologist*, 72(6), 1227–1288. <https://doi.org/10.1525/aa.1970.72.6.02a00030>
- Nelson, I. A. (2010). From quantitative to qualitative: Adapting the life history calendar method. *Field Methods*, 22(4), 413–428. <https://doi.org/10.1177/1525822x10379793>
- Nolan, J. M., & Ryan, G. W. (2000). Fear and loathing at the cineplex: Gender differences in descriptions and perceptions of slasher films. *Sex Roles*, 42(1/2), 39–56. <https://doi.org/10.1023/a:1007080110663>
- Onwuegbuzie, A. J. (2003). Effect sizes in qualitative research: A prolegomenon. *Quality and Quantity*, 37(4), 393–409. <https://doi.org/10.1023/a:1027379223537>
- Onwuegbuzie, A. J., & Collins, K. M. (2007). A typology of mixed methods sampling designs in social science research. *Qualitative Report*, 12, 281–316. <https://www.nova.edu/ssss/QR/QR12-2/onwuegbuzie2.pdf>
- Onwuegbuzie, A. J., & Leech, N. (2007). A call for qualitative power analyses. *Quality & Quantity*, 41(1), 105–121. <https://doi.org/10.1007/s11135-005-1098-1>
- 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>

- Quinn, N. (1987). Convergent evidence for a cultural model of American marriage. In D. Holland, & N. Quinn (Eds.), *Cultural models in language and thought* (pp. 173–192). Cambridge University Press.
- Quinn, N. (1997). Research on shared task solutions. In C. Strauss, & N. Quinn (Eds.), *A cognitive theory of cultural meaning* (pp. 137–188). Cambridge University Press.
- Quinn, N. (2005a). How to reconstruct schemas people share, from what they say. In N. Quinn (Ed.), *Finding culture in talk: A collection of methods* (pp. 35–81). Palgrave Macmillan.
- Quinn, N. (2005b) Introduction. In N. Quinn (Ed.), *Finding culture in talk: A collection of methods* (pp. 1–34). Palgrave Macmillan.
- Quinn, N. (2018). An anthropologist's view of American marriage: Limitations of the tool kit theory of culture. In N. Quinn (Ed.), *Advances in culture theory from psychological anthropology* (pp. 139–84). Palgrave Macmillan.
- Rahimi, S. (2024). Saturation in qualitative research: An evolutionary concept analysis. *International Journal of Nursing Studies Advances*, 6, Article 100174. <https://doi.org/10.1016/j.ijnsa.2024.100174>
- Rice, E. (1980). On cultural schemata. *American Ethnologist*, 7(1), 152–171. <https://doi.org/10.1525/ae.1980.7.1.02a00090>
- Romney, A. K., & D'Andrade, R. G. (1964). Cognitive aspects of English kin terms. *American Anthropologist*, 66(3), 146–170. <https://doi.org/10.1525/aa.1964.66.3.02a00870>
- Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, 88(2), 313–338. <https://doi.org/10.1525/aa.1986.88.2.02a00020>
- Rumelhart, D. (1984). Schemata and the cognitive system. In R. Wyer, & T. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 161–188). Lawrence Erlbaum Associates.
- Ruth, A., Mayfour, K., Hardin, J., Sangaramoorthy, T., Wutich, A., Bernard, H. R., Brewis, A., Beresford, M., SturtzSreetharan, C., Jones Brayboy, B. M., Dengah, H. F., Gravlee, C. C., Guest, G., Harper, K., Mahdavi, P., Mattison, S. M., Moritz, M., Negrón, R., Piperata, B. A., & , ... Zarger, R.. (2022). Teaching ethnographic methods: The state of the art. *Human Organization*, 81(4), 401–412.
- Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods*, 15(1), 85–109. <https://doi.org/10.1177/1525822x02239569>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & 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>
- Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Erlbaum.
- Schnegg, M., & Bernard, H. R. (1996). Words as actors: A method for doing semantic network analysis. *CAM Journal*, 8(2), 7–10. <https://doi.org/10.1177/1525822x960080020601>
- Schuster, R. C., Brewis, A., Wutich, A., Safi, C., Vanrespaille, T. E., Bowen, G., SturtzSreetharan, C., McDaniel, A., & Ochandarena, P. (2023). Individual interviews versus focus groups for evaluations of international development programs: Systematic testing of method performance to elicit sensitive information in a justice study in Haiti. *Evaluation and Program Planning*, 97, Article 102208. <https://doi.org/10.1016/j.evalproplan.2022.102208>
- Schuster, R. C., Wutich, A., Brewis, A., & SturtzSreetharan, C. (2022). Ethnographic methods in cross-cultural research. In P. Liamputtong (Ed.), *Handbook of qualitative cross-cultural research methods: A social science perspective*, (Vol. 273). ElgaronRhine.com.
- Sebele-Mpofu, F. Y. (2020). Saturation controversy in qualitative research: Complexities and underlying assumptions. A literature review. *Cogent Social Sciences*, 6(1), Article 838706. <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>
- Small, M. L. (2009). How many cases do I need? On science and the logic of case selection in field-based research. *Ethnography*, 10(1), 5–38. <https://doi.org/10.1177/1466138108099586>
- Sobal, J. (2001). Sample extensiveness in qualitative nutrition education research. *Journal of Nutrition Education*, 33(4), 184–192. [https://doi.org/10.1016/s1499-4046\(06\)60030-4](https://doi.org/10.1016/s1499-4046(06)60030-4)
- Squire, C. M., Giombi, K. C., Rupert, D. J., Amoozegar, J., & Williams, P. (2024). Determining an appropriate sample size for qualitative interviews to achieve true and near code saturation: Secondary analysis of data. *Journal of Medical Internet Research*, 26, Article e52998. <https://doi.org/10.2196/52998>
- Strauss, A., & Corbin, J. (1990). *Basics of qualitative research*. Sage.
- Strauss, C., & Quinn, N. (1997). *A cognitive theory of cultural meaning*. Cambridge University Press.
- SturtzSreetharan, C., Brewis, A., Hardin, J., Trainer, S., & Wutich, A. (2021). *Fat in four cultures: A global ethnography of weight*. University of Toronto Press.
- Surowiecki, J. (2005). *The wisdom of crowds*. Anchor Books.
- Thomson, S. B. (2011). Sample size and grounded theory. *Journal of Administration and Governance*, 5, 45–52. <https://ssrn.com/abstract=3037218>
- Tight, M. (2023). Saturation: An overworked and misunderstood concept? *Qualitative Inquiry*, 30(7), 577–583. <https://doi.org/10.1177/10778004231183948>
- Tippett, L. H. C. (1935). 3—statistical methods in textile research. Part 3—a snap-reading method of making time-studies of machines and operatives in factory surveys. *Journal of the Textile Institute Transactions*, 26(2), T51–T70. <https://doi.org/10.1080/19447023508661637>
- Torraco, R. J. (2005). Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review*, 4(3), 356–367. <https://doi.org/10.1177/1534484305278283>

- Trainer, S., SturtzSreetharan, C., Wutich, A., Brewis, A., & Hardin, J. (2022). Fat is all my fault: Globalized metathemes of body self-blame. *Medical Anthropology Quarterly*, 36(1), 5–26. <https://doi.org/10.1111/maq.12687>
- 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. <https://doi.org/10.1016/j.jclinepi.2016.10.001>
- Trotter, R. T. 2nd (2012). Qualitative research sample design and sample size: Resolving and unresolved issues and inferential imperatives. *Preventive Medicine*, 55(5), 398–400. <https://doi.org/10.1016/j.ypmed.2012.07.003>
- Tuckett, A. (2004). Qualitative research sampling-the very real complexities. *Nurse Researcher*, 12(1), 47–61. <https://doi.org/10.7748/nr2004.07.12.1.47.c5930>
- Uttley, L., Quintana, D. S., Montgomery, P., Carroll, C., Page, M. J., Falzon, L., Sutton, A., & Moher, D. (2023). The problems with systematic reviews: A living systematic review. *Journal of Clinical Epidemiology*, 156, 30–41. <https://doi.org/10.1016/j.jclinepi.2023.01.011>
- van Rijnssoever, F. J. (2017). (I can't get no) saturation: A simulation and guidelines for sample sizes in qualitative research. *PloS One*, 12(7), Article e0181689. <https://doi.org/10.1371/journal.pone.0181689>
- 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, 148–218. <https://doi.org/10.1186/s12874-018-0594-7>
- Wang, C., & Burris, M. A. (1997). Photovoice: Concept, methodology, and use for participatory needs assessment. *Health Education & Behavior*, 24(3), 369–387. <https://doi.org/10.1177/109019819702400309>
- Weller, S. C. (2007). Cultural consensus theory: Applications and frequently asked questions. *Field Methods*, 19(4), 339–368. <https://doi.org/10.1177/1525822x07303502>
- Weller, S. C. (2015). Sample size estimation: The easy way. *Field Methods*, 27(4), 333–347. <https://doi.org/10.1177/1525822x14530086>
- Weller, S. C., Johnson, J. C., & Dressler, W. W. (2023). Validating cultural models with cultural consensus theory. *Qualitative Psychology*, 10(2), 359–369. <https://doi.org/10.1037/qup0000253>
- Weller, S. C., & Romney, A. K. (1988). *Systematic data collection* (Vol. 10). Sage.
- Weller, S. C., Vickers, B., Bernard, H. R., Blackburn, A. M., Borgatti, S., Gravlee, C. C., & Johnson, J. C. (2018). Open-ended interview questions and saturation. *PloS One*, 13(6), Article e0198606. <https://doi.org/10.1371/journal.pone.0198606>
- Whittemore, R., & Knafl, K. (2005). The integrative review: Updated methodology. *Journal of Advanced Nursing*, 52(5), 546–553. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>
- Wilcox, D. F. (1900). The American newspaper: A study in social psychology. *The Annals of the American Academy of Political and Social Science*, 16(1), 56–92. <https://doi.org/10.1177/000271620001600104>
- Wutich, A., Beresford, M., SturtzSreetharan, C., Brewis, A., Trainer, S., & Hardin, J. (2021). Metatheme analysis: A qualitative method for cross-cultural research. *International Journal of Qualitative Methods*, 20, Article 16094069211019907. <https://doi.org/10.1177/16094069211019907>
- Wutich, A., Bernard, H. R., & Ryan, G. (2014). Text analysis. In: *Handbook of methods in cultural anthropology*. Rowman & Littlefield.
- Yik, M., & Chen, C. Z. (2023). Unraveling Chinese talk about emotion. *Frontiers in Psychology*, 14, Article 1157863. <https://doi.org/10.3389/fpsyg.2023.1157863>