

Revisiting Sample Size and Saturation Dilemma in Qualitative Research: A Philosophical and Practical Reflection on Reflexive Sufficiency

International Journal of Qualitative Methods

Volume 25: 1–11

© The Author(s) 2026

DOI: 10.1177/16094069261446123

journals.sagepub.com/home/ijq

Tauhid Hossain Khan¹, Md. Shawn Sheikh², and Ellen MacEachen³

Abstract

For both seasoned and novice qualitative researchers, the common question arises: “Is the sample size large enough?” This question is relevant to the purpose, philosophy, epistemology, and ontology behind qualitative research. Unlike quantitative studies, the purpose of qualitative research is not counting opinions or people but rather exploring diverse experiences. The ultimate aim of qualitative research is not to test a hypothesis; hence, samples are often intentionally diverse, aiming for depth and richness of data rather than statistical representativeness. The number of participants depends on the nature of the topic and the resources available. As well, whatever sampling approach is adopted, a researcher should know which strategy will solicit the best information for their study. The question of how many qualitative interviews are enough is posed not only by positivist researchers but also by those using interpretive approaches. In this context, we have seen seasoned qualitative researchers with dissenting views on, “How many interviews are enough?” Overall, the ultimate aim of sampling for qualitative research is to provide a convincing analytical narrative based on richness, complexity, and detail rather than on numerical logic. Related to sample size, qualitative researchers also face questions about how they can understand when their data are saturated. However, we ask: is it necessary to have data saturation for qualitative research studies? Or is it the researcher themselves who will decide how many data they need to complete their study? In our paper, we will reflect on the debates around data saturation in relation to the rigor of qualitative research primarily on the ground of its philosophical-theoretical argumentation rather than an empirical qualitative study. We propose reflexive sufficiency, a transparent researcher-guided approach, in order to determine data adequacy that upholds methodological rigor while honoring the nuanced, contingent, and profoundly human nature of the phenomena under study.

Keywords

qualitative research, sampling, saturation, rigor, human experiences, reflection, non-reductionism

Received: 10 August 2025; Revised: 9 April 2026; Accepted: 10 April 2026

Introduction

In qualitative research, the determination of sample size remains a contentious and critical issue which influences the robustness and credibility of research findings. Unlike quantitative research, which usually needs large samples for statistical

¹Department of Sociology, Jagannath University, Dhaka, Bangladesh

²General Education Department (GED), University of Liberal Arts Bangladesh (ULAB), Dhaka, Bangladesh

³School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada

Corresponding author:

Tauhid Hossain Khan, Department of Sociology, Jagannath University, 9-10, Chittaranjan Avenue, Dhaka, Bangladesh-1100.

Email: th3khan@uwaterloo.ca



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>).

accuracy, qualitative research focuses more on understanding detailed and complex human experiences, which implies that researchers prioritize depth over the number of participants (Malterud et al., 2016). Qualitative research can provide logical inference, such as induction, deduction, and abduction (Danermark et al., 2019; Khan & MacEachen, 2022) and what some people call “naturalistic generalization” (Ruddin, 2006), as well as what some others call “bottom-up” generalization (Borgstede & Scholz, 2021). Instead of focusing on numbers or counting how many people agree or disagree with something, qualitative research digs deeper into human and organizational experiences and meanings given to situations (Khan & MacEachen, 2021). The main difference between qualitative and quantitative research is not just about methods – the goals also differ. Qualitative research focuses on understanding how people make sense of situations, while quantitative research looks for patterns that can be measured statistically (Cope, 2013). As Patton (1990) puts it, qualitative research focuses on depth and context rather than numbers and wide-scale trends. For example, a qualitative study might look at how a small group of teachers experience remote learning. Instead of counting how many agree or disagree, the researcher would explore *why* they experience it that way, their struggles, beliefs, and personal stories. The goal is not to say, “All teachers think this,” but to understand how people differently experience the same situation and why. Morse (2015) explains that qualitative findings may not apply everywhere, but they can still be meaningful for similar groups. For instance, a study about how single mothers in a rural community cope with childcare challenges might reveal that strong family support and informal neighborhood networks help them manage. While these findings may not apply to single mothers in big cities (where support systems differ), they could still resonate with other close-knit rural communities facing similar struggles. In short, qualitative research helps us deeply understand people’s experiences by focusing on context, processes, emotions, and personal stories. This approach gives a fuller picture of human and organizational behavior, especially in fields like social research, where individual perspectives matter.

The concept of data saturation, which refers to the point at which no new information emerges from data collection, serves as a guiding principle for many qualitative researchers in determining sample size (Francis et al., 2010; Guest et al., 2020). However, the way saturation is applied differs between studies, leading to disagreements about whether small samples are sufficient for analytic conclusions (Hennink et al., 2016; Young & Casey, 2018). Some researchers believe that small samples can reveal key themes and patterns (Young & Casey, 2018) while other researchers believe that small samples can limit the validity and representativeness of results (Omona, 2013). Moreover, the lack of consensus on sample size determination becomes even more complex due to varying standards in academic fields and journal requirements, which influences how researchers justify their sample sizes (Vasileiou et al., 2018). To improve the quality of qualitative studies, researchers must clearly explain why their sample size is sufficient (Vasileiou et al., 2018).

Empirical studies of qualitative practice guide our critique and discussion of saturation. For instance, Carpenter-Song et al. (2020) conducted 30 interviews for their study on mobile health interventions despite identifying thematic saturation after 12–18 interviews, demonstrating an explicit prioritization of a comprehensive understanding over a saturation benchmark. This example is not anomalous but indicative of a common practice where researcher judgment actively overrides a technical point of saturation, a tension well-documented in methodological analyses (Hennink et al., 2016; Saunders et al., 2017). Thus, our philosophical deconstruction addresses an active methodological dilemma, questioning the coherence of a concept that is routinely subordinated to the researcher’s discretion in practice. Various alternative concepts have been proposed to navigate this dilemma, such as ‘information power’ (Malterud et al., 2016), ‘theoretical sufficiency’ (Charmaz, 2017), ‘meaning saturation’ (Hennink et al., 2016), and ‘analytical adequacy’ (Nelson, 2017). These approaches shift focus from a static endpoint to a dynamic assessment of data richness in relation to study aims. Our argument builds upon this lineage, proposing a framework of reflexive sufficiency that places the researcher’s transparent interpretive judgment at the center of determining data adequacy.

The Philosophical Underpinnings of the Saturation Debate

The debate over sample size and data saturation in qualitative research is often framed in methodological terms, but its roots run far deeper. At its core, this debate challenges the very nature of knowledge, meaning, and reality. Plato’s ancient insight that objective value (Forms) is accessed only through subjective experience casts doubt on the concept of data saturation (Plato, 2000). Declaring saturation (“we have captured all relevant truths”) commits a category error: it conflates the researcher’s limited synthesis with the phenomenon’s boundless essence. This Platonic concern with the relationship between subjective experience and objective truth finds a modern echo and a deeper existential grounding in Heidegger’s philosophy. Heidegger’s notion of *Dasein* (being-in-the-world) highlights that human existence is always contextual and interpretative (Heidegger, 1967). If reality is inherently dynamic and contextual, how can researchers claim to have reached a point where no new insights emerge? Hermeneutics, the study of interpretation, further challenges saturation by emphasizing that understanding is always partial, shaped by the researcher’s perspective and historical context (Gadamer, 1975). There is no final, complete interpretation, only an ongoing dialogue with the data. When a researcher declares saturation, they are not

arriving at an objective endpoint but imposing a closure on an inherently open-ended process. Ultimately, deciding on sample size in qualitative research involves careful thought about the specific research context and the study's goals (Palinkas et al., 2013). The researcher determines the study's reality and data collection endpoint, interpreting both spoken and silent discourses to construct findings (Khan & MacEachen, 2022).

The argument surrounding sample size, such as the number of interviews, in qualitative research is diverse and encompasses theoretical, methodological, and practical dimensions. By employing a critical appraisal framework, this paper integrates theoretical, philosophical, and practical perspectives to deconstruct the limitations of saturation as a methodological benchmark. As researchers strive to balance the demands of rigor and richness in their studies, the ongoing exploration of sample size adequacy remains a pivotal aspect of qualitative research methodology. The subsequent sections of the paper explore the philosophical problem of data saturation, the foundational principles of qualitative research, the interpretive nature of determining "enough" data, the theoretical and practical flaws of saturation, and the fluidity of language and meaning, culminating in a formal inductive argument against saturation. The paper concludes by urging a paradigm shift toward prioritizing narrative richness and researcher autonomy over rigid external standards.

Saturation Unwrapped: Where 'Enough' is an Interpretive Judgement, Not a Statistic

Saturation is a key concept in qualitative research and is often seen as a benchmark for ensuring methodological rigor (Guest et al., 2020). The concept of 'saturation' is not monolithic; its application varies significantly across qualitative traditions. Key distinctions include code saturation (the point where no new thematic codes are identified), meaning saturation (the achievement of depth and nuance in understanding those codes), and theoretical saturation (the development of a sufficiently robust theoretical model, as in grounded theory) (Guest et al., 2020; Hennink et al., 2016). This plurality is often collapsed in practice, leading to inconsistent application. The paper addresses the core assumption common to all these types: that there exists a discernible point of informational completeness which refers to the stage in data collection where no new insights or information are uncovered, signaling that the sample size is sufficient for the study's objectives.

Discussions about saturation frequently intersect with debates on the broader legitimacy of qualitative methodologies. This ongoing struggle for recognition highlights the importance of employing rigorous methods and carefully documenting saturation to enhance the credibility of qualitative research findings. However, operationalizing saturation is not uniform across methodologies. For instance, Guest et al. (2020) propose a straightforward method for evaluating thematic saturation, which is especially useful for researchers conducting inductive thematic analyses. Their approach emphasizes the use of a well-structured codebook to streamline the saturation process. Similarly, Francis et al. (2010) argue that saturation should align with the theoretical framework i.e. theoretical saturation, underpinning the research, ensuring that data collection supports the development of strong, well-founded theories. Also, the concept of *information power* has been suggested as an alternative to traditional saturation metrics. This approach argues that the adequacy of a sample size depends on factors such as the clarity of the research aim, the specificity of the sample, the theoretical framework, the quality of the dialogue during data collection, and the chosen analytical strategy (Malterud et al., 2016). This idea complements Sale's assertion that saturation should not be treated as a rigid endpoint but rather as a flexible concept that adapts to the research context and objectives (Sale, 2008).

Code and meaning saturation are influenced by several factors that researchers must carefully evaluate when determining sample sizes for qualitative studies (Hennink et al., 2016). According to Hennink et al. (2016), saturation is not merely a declaration but a methodological process requiring thorough consideration and documentation. They argue that saturation cannot be predetermined and must be assessed during the research process, which complicates the task of establishing a sample size a priori (Hennink et al., 2016). They stress that understanding the parameters that affect saturation is crucial for enhancing the rigor of qualitative research. Another aspect of the saturation debate is the quality of the data collected. Achieving saturation is less about hitting a numerical target for interviews and more about applying a systematic and thoughtful approach to research design and data interpretation (Mays & Pope, 2000). Tutar (2023) also highlights the challenges of determining optimal sample sizes, noting that both small and large samples can have drawbacks depending on the research context. Thus, saturation is a helpful concept but offers limited practical guidance for estimating sample size in advance of a research study (Guest et al., 2006).

Moreover, critically evaluating qualitative research often involves scrutinizing how saturation is achieved and reported. Frameworks like the Consolidated Criteria for Reporting Qualitative Research (COREQ) and the Critical Appraisal Skills Programme (CASP) provide valuable tools for assessing methodological rigor, including how well saturation is addressed (Tong et al., 2007). These tools encourage researchers to reflect on their sampling strategies and demonstrate the extent to which saturation has been achieved, thus promoting transparency and strengthening the credibility of their findings. Critical appraisal plays a vital role in validating the research process and the conclusions derived from qualitative data.

Failing to achieve saturation can have significant consequences. Insufficient saturation can weaken the quality of qualitative research, leading to incomplete or biased interpretations of the data (Saunders et al., 2017). Engaging participants actively in the research process and involving multiple researchers in data analysis can help mitigate bias, further reinforcing the robustness of the findings. Saturation also plays a critical role in assessing the overall quality of qualitative research. The evaluation of qualitative evidence often includes examining how saturation is defined and achieved within a study's context (Williamson, 2009). As a multifaceted concept, saturation serves as a marker of methodological rigor, influenced by research objectives, sample characteristics, and data collection strategies.

However, the question is, why should external benchmarks dictate sample size? If a researcher considers it necessary to conduct more interviews to fully grasp a topic, they should continue, regardless of saturation. This is particularly relevant in qualitative research where the richness of data is paramount. As stated earlier, Carpenter-Song et al. (2020), in their study on mobile health interventions, found that saturation was typically achieved after 12 to 18 interviews. However, they opted for a larger sample size of 30 to ensure a comprehensive understanding of their topic. It shows that researcher discretion matters more than preset assumptions of saturation.

The ongoing discussion around saturation underscores the need for clear definitions and robust methodologies to ensure qualitative research withstands scrutiny and contributes meaningfully to various fields. But here, in this study, the purpose is to disregard saturation as an external benchmark and the authority of the researcher themselves and their ability to decide when to stop gathering data.

Sample Size in Qualitative Research: What is Sufficient Data?

The question 'how many interviews are enough?' is ultimately unanswerable with a universal number because it is the wrong question for qualitative inquiry. The more pertinent question is: How does a researcher know they have gathered sufficient data for their specific analytical and interpretive goals? Answering this question requires shifting the focus from seeking an external numeric threshold to articulating an internal, iterative process of judgment. Therefore, our critique of saturation challenges 'saturation' as a primary proxy used to signal 'enough'. In the following sections, the paper argues that adequacy emerges not from saturation but from transparent, reflexive engagement with the data, where the researcher determines when the depth and richness of data align with the study's purpose. Unlike quantitative research, where the focus is on statistical representation and generalizability, qualitative research emphasizes the depth and richness of data obtained from participants (Gaskell, 2000). It seeks to understand the complexity and context of the issues studied (Gaskell, 2000). This perspective is echoed by Kuzel (1992), who highlights that qualitative research is concerned with the richness of information rather than the sheer number of participants. Of importance, the question, "*How many interviews are enough?*" does not have a simple answer, as it depends on various factors, such as the nature of the research questions and objectives, topic complexity, the resources available, population diversity, and the chosen method of analysis (Baker & Edwards, 2012; Gaskell, 2000). Although some researchers argue for a representative sample, qualitative research often emphasizes depth over breadth (Boddy, 2016).

Tutar et al. (2023) points out that different qualitative designs require different sample sizes. For example, oral history research might rely on a single participant, while phenomenological or grounded theory studies often involve 5 to 15 participants, or even more for more complex phenomena. This variability highlights the need to align the sample size with the research design and objectives. Here, Tutar et al. (2023) highlights that different qualitative designs require different sample sizes, but the researcher's goals should be the deciding factor, not just methodological conventions.

In health research, sample size decisions are often shaped by disciplinary norms rather than the number of interviews needed for the research question and context (Vasileiou et al., 2018). This indicates a need for qualitative researchers to be transparent about their sample size decisions, as it is a critical marker of research quality (Vasileiou et al., 2018). Participant diversity is also essential, as it provides a broader understanding of the phenomenon being studied (Khan & MacEachen, 2021). This approach aligns with the findings of (Young & Casey, 2018), who demonstrated that even small samples in qualitative research can reveal meaningful themes and codes, proving that quality insights do not require large sample sizes. Furthermore, the quality of the data is as important as the number of interviews. Roulston (2010) suggests that qualitative researchers should prioritize collecting rich, detailed data over conducting numerous superficial interviews. High-quality interviews can provide deeper insights into the research question, making a smaller, carefully selected sample more valuable than a larger sample.

The iterative nature of qualitative research also affects sample size decisions. As Ayasreh et al. (2022) note, researchers often assess their data collection process continuously, adapting their approach based on emerging information. This flexibility allows them to refine interview questions and techniques, ensuring they capture the most relevant data. Consequently, the final sample size often evolves during the research process rather than being fixed from the outset.

McGrath et al. (2018) highlight that qualitative interviews provide in-depth insights into participants' unique experiences, which can vary greatly depending on the context and population. This variability suggests that the number of interviews required may differ based on the specific research focus and the heterogeneity of the participant group. For instance, Brayda and Boyce (2014) argue for the importance of culturally sensitive approaches in qualitative research, noting that the depth and richness of data obtained can influence the number of interviews needed to fully understand the subject matter. While empirical research and mathematical models have informed sample size estimation, there is no universal formula for determining what is adequate.

Practical constraints, however, often necessitate preliminary estimations. Researchers face institutional demands, such as grant proposals that require justification for projected sample sizes. In response to these pressures, scholars frequently cite specific ranges or numbers based on factors like population homogeneity, research design, or anticipated saturation points (Hagaman & Wutich, 2016). While practical requirements like funding applications make these estimates necessary, they remain context-dependent approximations rather than universal standards. Crucially, they should not supersede the core qualitative principle that sufficient sample size emerges iteratively through the research process, guided by the researcher's judgment, depth of data, and evolving research objectives (Hennink et al., 2019).

Above all, the question of how many qualitative interviews are enough is complex and context dependent. The right number of interviews depends on the situation, including using benchmarks to estimate sample size for a grant proposal. However, the sample size ultimately depends on what the researcher aims to discover, rather than arbitrary benchmarks. In all, we argue that researchers should adopt a reflexive approach to determining sufficiency, as outlined in Figure 1, below:

This framework shifts the demonstration of rigor from claiming a methodological endpoint to illustrating a conscientious process. It also replaces standards conventions with informed interpretive judgment. This judgment is not arbitrary but is cultivated through deep engagement with the data, theoretical knowledge, and reflexive practice. It parallels the analytical process in methods like Reflexive Thematic Analysis, where the researcher's active role in interpreting patterns is central to the rigor of the inquiry (Braun & Clarke, 2021).

Words Without Measure: The Arbitrary Endgame of Saturation

The concept of data saturation is philosophically problematic because it assumes an objective threshold that can never truly be reached. Since the meaning of a narrative cannot be measured objectively, a researcher's analysis of a participant's experiences remains inherently interpretive (Frost et al., 2010). There are no fixed standards to determine whether qualitative data is "complete" or "sufficient", terms like "good" or "bad" when assessing a story are based on interpreted judgements

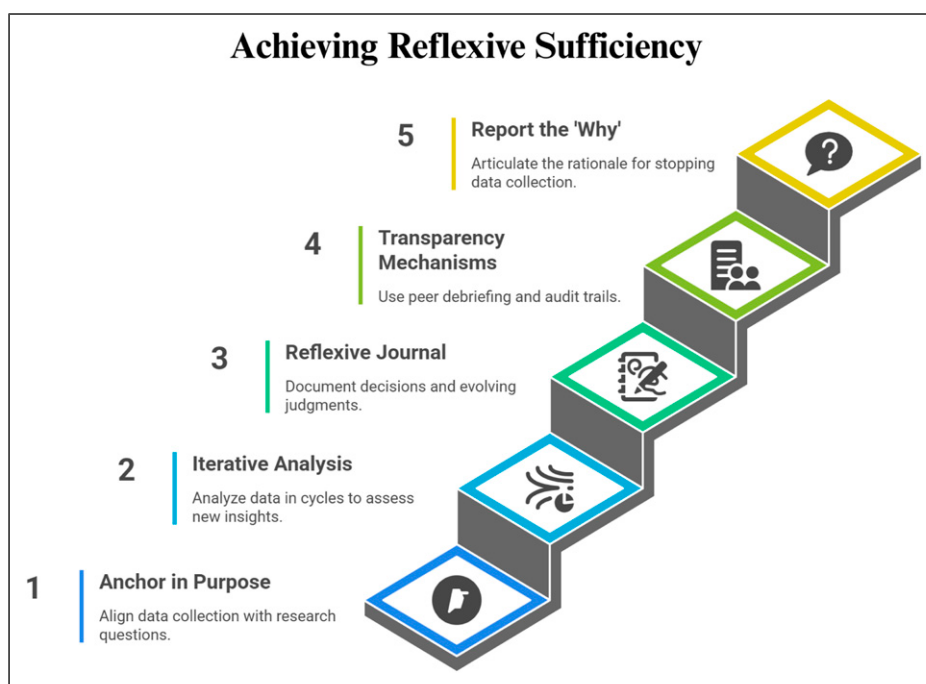


Figure 1. Steps of Achieving Reflexive Sufficiency

rather than objective scales. Consequently, declaring saturation to justify ending data collection is an arbitrary decision (Saunders et al., 2017) that is contingent on the researcher's judgment. One may choose to conduct one hundred interviews or stop after fifty, but this decision reflects the researcher's discretion rather than an empirically verifiable endpoint. (Hennink et al., 2016; Saunders et al., 2017). For instance, Oscar Lewis's *The Children of Sánchez* is based on the life histories of a single family in Mexico City (Bartlett et al., 1961). Despite focusing on just five individuals, the work was widely accepted as representative of the urban poor (Bartlett et al., 1961) and it was used to support his "culture of poverty" thesis. Another example is Sigmund Freud's study of "Anna O.," which was instrumental in the development of psychoanalysis (Breuer, 1893). Although based on a single patient, the case significantly influenced the trajectory of psychological theory and practice (Breuer, 1893). These examples show that researchers often rely on their own theoretical frameworks, interpretive depth, and personal judgment rather than any fixed rule to determine when enough data has been collected.

Against the backdrop of researcher discretion evident in the examples mentioned above, some might appeal to phenomenology. They might argue that Husserl's phenomenology, particularly the practice of *bracketing* (epoché), could mitigate subjectivity by setting aside the researcher's preconceptions (Husserl et al., 2013). However, this approach is philosophically untenable. Preconceptions and biases are intrinsic to one's identity, as they shape how a researcher perceives and interprets experiences. To fully eliminate them would be to erase the researcher's existential framework (Sartre, 2007). As Sartre posits, *existence precedes essence*: identity is not preordained but is constructed through lived experience (Sartre, 2007). Thus, stripping away a researcher's preconceptions would remove the very framework through which meaning is derived, leaving their analysis devoid of perspective, and therefore, of real significance.

Therefore, attempting to achieve saturation through bracketing does not resolve subjectivity but instead complicates it further. Without objective criteria, which further demands the objective meaning of qualitative data, saturation remains an ambiguous and ultimately indefensible concept in qualitative research.

Meaning of Language Differs From Generation to Generation

In his early work, *Tractatus Logico-Philosophicus*, Ludwig Wittgenstein argued that words have fixed, singular meanings (Wittgenstein, 2013). He believed that many philosophical problems arose from misunderstandings in language, and if each word could be assigned a precise, unchanging meaning, such confusion could be eliminated. For example, if a mason says "brick", it would strictly refer to the physical object—nothing more, nothing less (Wittgenstein, 2013).

But later, in *Philosophical Investigations*, he changed his mind. He realized that words can have many meanings depending on the situation (Wittgenstein, 1968). For instance, when a mason says "brick" to a worker, it might mean, "Bring me the brick," not just the object itself. Or the term "care" may signify medical treatment in a hospital, emotional support in a family, or bureaucratic responsibility in policy discussions. He states that language is a public tool to express our private life and that meaning depends on how words are used in real life (Wittgenstein, 1968).

Why is this important? It matters because it helps us understand why *data saturation* (the point where no new information is found in research) is hard to achieve. If we follow Wittgenstein's early idea, that words have fixed meanings, then researchers could easily find patterns in interviews and reach data saturation. But his later view shows that words change meaning based on context, making it nearly impossible to fully saturate data because meanings keep shifting depending on the context (Mistler & Mandelkorn, 2008).

Another issue is generational differences in language. For example, the word *pastime* might mean reading books or playing outside for millennials, while Gen-Z may think of video games or social media. Even if two people use the same word, their understanding of it could be completely different (Majid et al., 2018). Similarly, imagine a wife losing a game of Ludo and saying, "You don't love me" to her husband. Does she really mean her husband doesn't love her, or is she just joking? The meaning depends on tone, relationship, and context (Mistler & Mandelkorn, 2008). Further, the contextual fluidity of language across different social and temporal settings should be considered. For instance, the meaning of a concept like 'community' or 'privacy' can vary significantly between generations, cultural groups, and professional domains (Majid et al., 2018). This variation underscores that, even when the same word is used, the underlying concept and its experiential significance may differ, complicating any claim that a fixed set of data can fully 'capture' its meaning.

This shows that language is not fixed; it changes with time, culture, and situation. So, in research, reaching data saturation is very difficult because every interview can bring new meanings. Wittgenstein's later philosophy teaches us that words are flexible, and that is why research must always be open to new interpretations and contingent upon the judgements of the researcher.

A Part of Painting is Not the Whole Painting

In qualitative research, the aim is to deeply understand a specific social reality within its unique context, not to statistically generalize findings to a wider population (Hennink et al., 2016). When we explore the lives of people in a particular place; say, the fishermen on a remote island, we are uncovering a part of the larger picture, much like a section of a painting (Malterud et al., 2016). The concept of data saturation, often used in qualitative studies, suggests that data collection can stop once no new information is emerging. While this is commonly accepted as a way to validate sample size, it unintentionally leads to the idea that the insights gathered from a few participants represent the entire community (Guest et al., 2020). But this is misleading. Just as one corner of a painting cannot capture the full meaning or emotion of the whole artwork, one part of a community's experience cannot reflect the full complexity of that society (Guest et al., 2020). For example, interviewing ten fishermen and finding repeated patterns in their responses might suggest saturation, but it misses the stories of women in the fishing community, local traders, or young people aspiring to leave the island (Rahimi & Khatooni, 2024). Similarly, research on classroom challenges in one school may stop when themes repeat, but this does not mean the findings apply to all schools (Rahimi & Khatooni, 2024). Just as a painting is more than its individual parts, qualitative research offers a deep but partial perspective, one that should not be mistaken for the whole truth. Therefore, saturation should not dictate when research ends. Instead, it should be the researcher's judgment, guided by the focus of the study and the depth of insight desired, that determines when enough data has been collected. In this sense, the researcher is like an artist who decides when the painting is finished, not based on rules, but on vision, purpose, and what part of the picture they aim to show.

Formal Argument Against the Data Saturation

Given the arguments stated above, it is possible to structure an inductive argument to demonstrate why data saturation is unreliable and is a vague concept for determining the sample size of a study. An inductive approach is particularly necessary in social science framework formation as it encourages a shift away from rigid hypothetico-deductive models towards a more flexible, exploratory model that allows new theories to emerge based on observed and interpreted data (Locke, 2007). This helps avoid the problem of forcing data to fit pre-existing theories, allowing researchers to develop ideas that truly match what they see in their dataset. As noted by Hitchcock (2005), there are multiple ways to create an inductive argument including Toulmin's Model of Argumentation (Hitchcock, 2005), the Defeasible Reasoning Framework proposed by Walton and Macagno (2010), and the Argument from Cumulative Case (Inductive Generalization)) (Smaling, 2003). Toulmin's Model of Argumentation lays out a structured approach for formulating arguments, which includes claims, grounds, warrants, backing, qualifiers, and rebuttals (Hitchcock, 2005). The Defeasible Reasoning Framework emphasizes that arguments can be retracted or altered in light of new evidence (Walton & Macagno, 2010). This suggests that conclusions drawn from qualitative data should remain open to further scrutiny which reflects the logical process of reasoning where new evidence may overturn prior conclusions. The Argument from Cumulative Case (inductive generalization) serves as a foundation for showing how cumulative cases build up knowledge gradually (Smaling, 2003). In this argument process, each premise independently supports the conclusion; together, they create a strong probabilistic inference (Smaling, 2003). The Argument from Cumulative Case will be adopted here where some touches of Toulmin's Model and Defeasible Reasoning Framework will also be visible. The formal inductive argument against saturation is grounded in the same logic that underpins core qualitative methodologies. Approaches like Grounded Theory (Charmaz, 2017), Interpretative Phenomenological Analysis (Smith et al., 2017), and Reflexive Thematic Analysis (Braun & Clarke, 2021) all employ forms of inductive reasoning to build understanding from data. The argument synthesizes this methodological spirit into a philosophical critique, demonstrating why the concept of saturation is incompatible with the iterative, open-ended, and interpretive nature of qualitative knowledge production (Figure 2).

Discussion and Conclusion

The debate over sample size and data saturation in qualitative research is not just a technical discussion, it is a fundamental question about how we understand human experiences. Unlike quantitative research, where numbers and statistical significance rule, qualitative research thrives on depth, context, and the richness of individual stories (Palinkas et al., 2013). In this paper, we have argued that a preoccupation with data saturation and predetermined sample sizes is misguided because it imposes a rigid framework on a process that is inherently fluid and subjective (Frost et al., 2010). The idea that a researcher can declare "saturation" and stop collecting data assumes that human experiences can be neatly packaged and fully understood at some fixed point. But as Wittgenstein (1968) showed, language and meaning are constantly shifting. What a word or story means to one person or generation may differ for another. Similarly, a researcher's own experiences, situations and perspectives shape how they interpret data, making the notion of an objective stopping point logically impossible. The examples of Oscar Lewis's *The Children of Sánchez* and Freud's case study of *Anna O.* demonstrate that groundbreaking insights can come from small, deeply examined samples, and are not

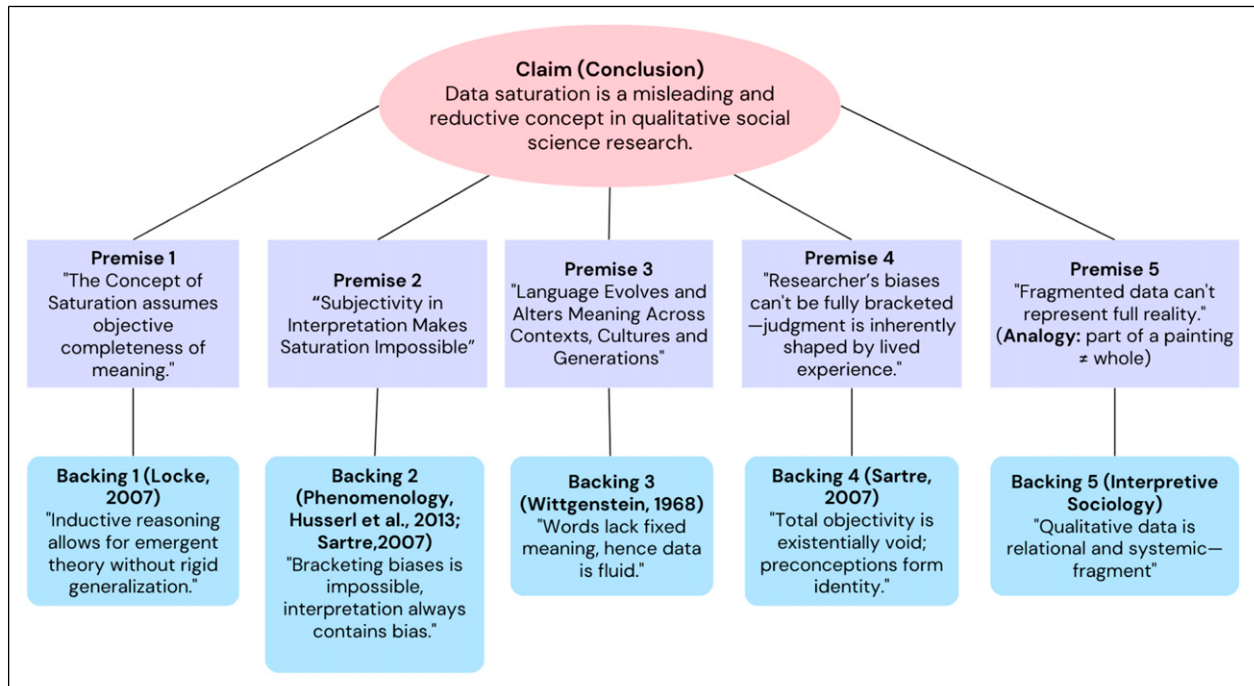


Figure 2. Formal Inductive Argument against Data Saturation

dependent on large, statistically representative ones. These works were influential not because they achieved saturation, but because the researchers delved deeply into their subjects' worlds, offering nuanced interpretations that resonated beyond their immediate context. This highlights a core strength of qualitative research: to provide a compelling, insightful narrative that offers depth and understanding, acknowledging that such narratives are constructed interpretations rather than claims of exhaustive completeness. The rain scene in Oscar-winning Korean movie *Parasite* (2019), directed by Bong Joon-ho, gives a powerful example to question the idea of data saturation in qualitative research. For the rich Park family, the rain seems peaceful and refreshing. But for the poor Kim family, the same rain brings suffering, damage, and loss. This difference is not about the rain itself, but about how each family experiences it based on their socio-economic position, memory, and emotion. Their different reactions show that meaning comes from situational experience, which is often ignored when researchers assume that no new information will appear after a certain point. Also, the researcher's interpretation is itself conditioned by their own positionality and experiential horizon. A researcher who has faced poverty might interpret findings related to Kim's pain, while another researcher from a privileged background may focus more on the Park family. Therefore, interpretation is never neutral; it is co-constructed, shaped by the researcher's ontological and epistemological stance. Just as the movie shows us different truths about the same rain, research needs to accept that there is never just one version of events. Our job is not to find some imaginary "complete" picture, but to understand that every story has multiple sides. Trying to force all these different experiences into a neat box called "saturation" doesn't make sense.

Ultimately, this paper calls for a shift toward reflexive sufficiency in qualitative research. Rather than clinging to the illusion of saturation, researchers should embrace their role as interpreters of complex human experiences. Descartes' axiom *'I think, therefore I exist'* reminds us that all inquiry presupposes an interpretive agent, the researcher (Kochashvili, 2015). While an independent reality may exist, its representation is inevitably mediated by the researcher's perspective, judgment, and engagement (Kochashvili, 2015). Foucault's work further destabilizes the illusion of objectivity, revealing how knowledge is produced within systems of power and discourse (Khan & MacEachen, 2021). Just as he exposed the ways institutions shape what counts as 'truth,' qualitative researchers must acknowledge that their narratives are not neutral reflections of reality but *constructed* accounts, shaped by their positionality, the questions they ask, and the voices they amplify or marginalize (Khan & MacEachen, 2021). The researcher's interpretive judgment, not a one-size-fits-all concept like saturation, should guide when enough data has been collected. In the end, the value of qualitative research lies in its ability to reveal the depth and diversity of human life rather than in meeting artificial benchmarks. By letting go of saturation as a gold standard, researchers can focus on what truly matters: telling stories that are meaningful, insightful, and true to the lived experiences of the people they study. This aligns with the critical perspective of Foucauldian Discourse Analysis (FDA), which examines how methodological choices themselves are shaped by institutional power and discursive norms. It reveals that the pursuit of 'saturation' can be a rhetorical performance of rigor, serving to legitimize findings within certain academic power structures rather than reflecting an authentic epistemological endpoint.

The paper acknowledges the diverse epistemological goals within qualitative research, which can range from generating transferable theory to providing rich, contextualized descriptions. The argument that the value of qualitative work lies in providing compelling, insightful narratives rather than conforming to external standards of completeness highlights the primacy of depth and meaning-making. This is a common thread across many qualitative paradigms. By adopting a framework of reflexive sufficiency and transparent interpretive judgment, qualitative researchers can uphold methodological rigor while honoring the nuanced, contingent, and profoundly human nature of the phenomena they study.

Acknowledgements

The authors would like to thank to the members of the Dr. Khan Social Impact Lab for the engagement in brainstorming and critical thinking.

ORCID iDs

Tauhid Hossain Khan  <https://orcid.org/0000-0002-7502-7377>

Md. Shawn Sheikh  <https://orcid.org/0009-0001-3316-3821>

Ellen MacEachen  <https://orcid.org/0000-0001-6477-7650>

Author Contributions

THK conceptualized the study and prepared the initial draft of the manuscript. MSS contributed to the refinement of ideas, provided further input, and assisted in finalizing the manuscript. EM, in the capacity of senior researcher, offered critical feedback and scholarly insights that enhanced the quality of the work. All authors have read and approved the final manuscript.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Declaration of Conflicting Interests

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

References

- Ayasreh, I., Hayajneh, F., Awamleh, R., ALBashtawy, M., & Alkhalwaldeh, A. (2022). Triggers of workplace violence in emergency departments: a qualitative study. *Nurse Media Journal of Nursing*, 12(3), 340–352. <https://doi.org/10.14710/nmjn.v12i3.44914>
- Baker, S. E., & Edwards, R. (2012). How many qualitative interviews is enough.
- Bartlett, H., Lewis, O., Zavattini, C., Quinn, A., Del Rio, D., & Ferrer, L. (1961). *The children of Sanchez*. <https://ci.nii.ac.jp/ncid/BA52221378>
- Boddy, C. (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>
- Borgstede, M., & Scholz, M. (2021). Quantitative and qualitative approaches to generalization and replication—A representationalist view. *Frontiers in psychology*, 12(5), 605191. <https://doi.org/10.3389/fpsyg.2021.605191>
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Brayda, W., & Boyce, T. (2014). So you really want to interview me?: navigating “sensitive” qualitative research interviewing. *International Journal of Qualitative Methods*, 13(1), 318–334. <https://doi.org/10.1177/160940691401300115>
- Breuer, J. (1893). Fräulein Anna O, Case Histories from Studies on Hysteria. The Standard Edition of the Complete Psychological Works of Sigmund Freud, Volume II (1893-1895): Studies on Hysteria, 19–47.
- Carpenter-Song, E., Jonathan, G., Brian, R., & Ben-Zeev, D. (2020). Perspectives on mobile health versus clinic-based group interventions for people with serious mental illnesses: a qualitative study. *Psychiatric Services*, 71(1), 49–56. <https://doi.org/10.1176/appi.ps.201900110>
- Charmaz, K. (2017). Constructivist grounded theory. *The Journal of Positive Psychology*, 12(3), 299–300. <https://doi.org/10.1080/17439760.2016.1262612>
- Cope, D. (2013). Methods and meanings: credibility and trustworthiness of qualitative research. *Oncology Nursing Forum*, 41(1), 89–91. <https://doi.org/10.1188/14.onf.89-91>
- Danermark, B., Ekström, M., & Karlsson, J. C. (2019). *Explaining society: Critical realism in the social sciences*. Routledge.
- Francis, J., Johnston, M., Robertson, C., Glidewell, L., Entwistle, V., Eccles, M., & Grimshaw, J. (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>

- Frost, N., Nolas, S., Brooks-Gordon, B., Esin, C., Holt, A., Mehdizadeh, L., & Shinebourne, P. (2010). Pluralism in qualitative research: the impact of different researchers and qualitative approaches on the analysis of qualitative data. *Qualitative Research*, 10(4), 441–460. <https://doi.org/10.1177/1468794110366802>
- Gadamer, H.-G. (1975). *Truth and Method* (J. Weinsheimer & D. G. Marshall, Trans.; 2nd ed.). Continuum.
- Gaskell, G. (2000). *Qualitative research in sociology: A review of the literature*. Sage.
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods*, 18(1), 59–82. <https://doi.org/10.1177/1525822x05279903>
- Guest, G., Namey, E., & Chen, M. (2020). A simple method to assess and report thematic saturation in qualitative research. *Plos One*, 15(5), Article e0232076. <https://doi.org/10.1371/journal.pone.0232076>
- Hagaman, A., & Wutich, A. (2016). 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>
- Heidegger, M. (1967). *Being and Time* (J. Macquarrie & E. Robinson, Trans.). Blackwell Publishers Ltd.
- Hennink, M., Kaiser, B., & Marconi, V. (2016). Code saturation versus meaning saturation. *Qualitative Health Research*, 27(4), 591–608. <https://doi.org/10.1177/1049732316665344>
- Hennink, M., Kaiser, B., & Weber, M. (2019). What influences saturation? estimating sample sizes in focus group research. *Qualitative Health Research*, 29(10), 1483–1496. <https://doi.org/10.1177/1049732318821692>
- Hitchcock, D. (2005). Good reasoning on the Toulmin model. *Argumentation*, 19(3), 373–391. <https://doi.org/10.1007/s10503-005-4422-y>
- Husserl, E., Moran, D., Findlay, T. J., & Dummett, M. (2013). *Logical Investigations Volume 2*. Routledge.
- Khan, T. H., & MacEachen, E. (2021). Foucauldian Discourse Analysis: Moving Beyond a Social Constructionist Analytic. *International Journal of Qualitative Methods*, 20, 16094069211018009. <https://doi.org/10.1177/16094069211018009>
- Khan, T. H., & MacEachen, E. (2022). An Alternative Method of Interviewing: Critical Reflections on Videoconference Interviews for Qualitative Data Collection. *International Journal of Qualitative Methods*, 21, 16094069221090063. <https://doi.org/10.1177/16094069221090063>
- Kochashvili, K. (2015). I Think, Therefore I Am ("I Exist as a Subject of Right"). *J. Law*, 58. https://heinonline.org/HOL/login-hol?redirect_url=https%3A%2F%2Fheinonline.org%2FHOL%2Fmojo%3F%252Fmojo%3D%26auth_token%3DaFFW6nrgDgPAUYPSzz3RFUs63KoAHP4TW8iS9gJsnuQ%253D%26casa_token%3D%26div%3D8%26g_sent%3D1%26paywall_url%3D%26referrer%3Dhttps%253A%252F%252Fheinonline.org%252FHOL%252FPage%253Fhandle%253Dhein.journals%252Ftsujrnl2015%26timestamp%3D1776348734%26url%3D%252FHOL%252FPage%253Fhandle%253Dhein.journals%252Ftsujrnl2015%2526div%253D8%2526id%253D%2526page%253D
- Kuzel, A. J. (1992). Sampling in qualitative inquiry.
- Locke, E. (2007). The case for inductive theory building. *Journal of Management*, 33(6), 867–890. <https://doi.org/10.1177/0149206307307636>
- Majid, A., Roberts, S., Cilissen, L., Emmorey, K., Nicodemus, B., O'Grady, L., Levinson, S., LeLan, B., de Sousa, H., Cansler, B. L., Shayan, S., de Vos, C., Senft, G., Enfield, N. J., Razak, R. A., Fedden, S., Tufvesson, S., Dingemanse, M., & Ozturk, O. (2018). Differential coding of perception in the world's languages. *Proceedings of the National Academy of Sciences*, 115(45), 11369–11376. <https://doi.org/10.1073/pnas.1720419115>
- Malterud, K., Siersma, V., & Guassora, A. (2016). Sample size in qualitative interview studies. *Qualitative Health Research*, 26(13), 1753–1760. <https://doi.org/10.1177/1049732315617444>
- Mays, N., & Pope, C. (2000). Qualitative research in health care: assessing quality in qualitative research. *BMJ*, 320(7226), 50–52. <https://doi.org/10.1136/bmj.320.7226.50>
- McGrath, C., Palmgren, P., & Liljedahl, M. (2018). Twelve tips for conducting qualitative research interviews. *Medical Teacher*, 41(9), 1002–1006. <https://doi.org/10.1080/0142159x.2018.1497149>
- Mistler, B., & Mandelkorn, J. (2008). Book review. *Journal of Constructivist Psychology*, 21(2), 178–181. <https://doi.org/10.1080/10720530701853818>
- Morse, J. M. (2015). Data were saturated: Then what? *Qualitative Health Research*, 25(5), 711–718. <https://doi.org/10.1177/1049732315575316>
- Nelson, J. (2017). Using conceptual depth criteria: addressing the challenge of reaching saturation in qualitative research. *Qualitative research*, 17(5), 554–570. <https://doi.org/10.1177/1468794116679873>
- Omona, J. (2013). Sampling in qualitative research: improving the quality of research outcomes in higher education. *Makerere Journal of Higher Education*, 4(2), 169–185. <https://doi.org/10.4314/majohe.v4i2.4>
- Palinkas, L., Horwitz, S., Green, C., Wisdom, J., Duan, N., & Hoagwood, K. (2013). 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(5), 533–544. <https://doi.org/10.1007/s10488-013-0528-y>

- Parasite, Bong, J.-h. (Director). (2019). *Parasite* [Film]. Barunson E&A.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Sage.
- Plato. (2000). *The Republic* (G. R. F. Ferrari & T. Griffith, Trans.). Cambridge University Press.
- Rahimi, S., & Khatooni, M. (2024). Saturation in qualitative research: an evolutionary concept analysis. *International Journal of Nursing Studies Advances*, 6, 100174. <https://doi.org/10.1016/j.ijnsa.2024.100174>
- Roulston, K. (2010). Considering quality in qualitative interviewing. *Qualitative Research*, 10(2), 199–228. <https://doi.org/10.1177/1468794109356739>
- Ruddin, L. P. (2006). You can generalize stupid! Social scientists, Bent Flyvbjerg, and case study methodology. *Qualitative inquiry*, 12(4), 797–812. <https://doi.org/10.1177/1077800406288622>
- Sale, J. (2008). How to assess rigour or not in qualitative papers. *Journal of Evaluation in Clinical Practice*, 14(5), 912–913. <https://doi.org/10.1111/j.1365-2753.2008.01093.x>
- Sartre, J. P. (2007). *Existentialism is a Humanism*. Yale University Press.
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., & Jinks, C. (2017). Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & Quantity*, 52(4), 1893–1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Smaling, A. (2003). Inductive, analogical, and communicative generalization. *International Journal of Qualitative Methods*, 2(1), 52–67. <https://doi.org/10.1177/160940690300200105>
- Smith, J. A. (2017). Interpretative phenomenological analysis: Getting at lived experience. *The Journal of Positive Psychology*, 12(3), 303–304. <https://doi.org/10.1080/17439760.2016.1262622>
- Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, 19(6), 349–357. <https://doi.org/10.1093/intqhc/mzm042>
- Tutar, H., Şahin, M., & Sarkhanov, T. (2023). Problem areas of determining the sample size in qualitative research: a model proposal. *Qualitative Research Journal*, 24(3), 315–336. <https://doi.org/10.1108/qjrj-06-2023-0099>
- Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, 18(1), 148. <https://doi.org/10.1186/s12874-018-0594-7>
- Walton, D., & Macagno, F. (2010). Defeasible classifications and inferences from definitions. *Informal Logic*, 30(1), 34–61. <https://doi.org/10.22329/il.v30i1.692>
- Williamson, K. (2009). Evidence-based practice: critical appraisal of qualitative evidence. *Journal of the American Psychiatric Nurses Association*, 15(3), 202–207. <https://doi.org/10.1177/1078390309338733>
- Wittgenstein, L. (1968). *Philosophical investigations*. Wittgenstein. Macmillan.
- Wittgenstein, L. (2013). *Tractatus Logico-Philosophicus*. Routledge eBooks. <https://doi.org/10.4324/9781315884950>
- Young, D., & Casey, E. (2018). An examination of the sufficiency of small qualitative samples. *Social Work Research*. <https://doi.org/10.1093/swr/svy026>

Author Biographies

Dr. Tauhid Hossain Khan obtained his Ph.D. in Public Health Sciences from the University of Waterloo, Canada, focusing on the work and health. His main research interests include the precarious labor market (gig, Gen-Z, vulnerable, and equity-deserving workers), occupational health and safety, boy's health, health education/promotion, critical pedagogy, and critical qualitative methods. Dr. Khan is an associate professor in Sociology at Jagannath University, Bangladesh. His research output includes approximately 25 research articles published in reputed national and international refereed journals, with 15 of them as the lead author.

Md Shawn Sheikh is a lecturer in the General Education Department at the University of Liberal Arts Bangladesh (ULAB). With an academic background in Philosophy, he brings a strong foundation in critical thinking and ethical analysis to his teaching and research. Currently, he is pursuing his Master's by research in the Department of Government and Politics at University Putra Malaysia. His broader research interests lie in political philosophy and the evolving dynamics of global political and technological developments, particularly their promises and perils.

Dr. Ellen MacEachen is a professor and the director of School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada. Her research examines the design and performance of work and health systems to identify how they can be improved and adapted to fast changing economic, social, and technological environments of our global economy. She is particularly interested in international work disability policy and the health risks and opportunities of new forms of work (such as “sharing economy” gigs).