American Journal of Qualitative Research
2025, Vol. 9 No. 1, pp. 95-124
https://doi.org/10.66815/ajqr/18485
© 2024 AJQR. http://www.ajqr.org

AMERICAN JOURNAL OF QUALITATIVE RESEARCH
ISSN: 2576-2141

Building the Thematic Structure in Qualitative Research: The Theme Construction Tool

David C. Coker[1]
West Liberty University, WV, USA

ABSTRACT
Themes and thematic analysis are very popular in qualitative research methods and analysis. The very definition of theme varies widely, from patterns that exclude the researcher as biased to highly interpretive work that prizes the researcher's instrumentality. Exploring the historical development of themes, the many definitions reveal that different levels of analysis produced varied understandings of the meanings of concepts and themes. A qualitative analysis of educational leadership dissertations in the US using thematic analysis showed that a common practice was to state that themes emerged; regardless of the claims, the resultant themes were mostly topical and descriptive. The construction of themes beyond the descriptive is a practice for which many researchers lack systematic, rigorous processes and fail to show visible methods. To develop a methodical, visible practice, a theme construction tool provides a three-step framework to use a storyboard approach to develop concepts, produce hypotheses using abductive reasoning, and then evaluate hypotheses to make an inference to the best explanation. Criteria for evaluating themes provide a reference point for researchers to apply concrete skills across a range of qualitative methods. The paper serves as a call to action for researchers to adopt structured, transparent, and methodologically flexible tools for theme construction. By doing so, researchers can move beyond simple summarization and opaque claims of "emerge" to produce more credible, nuanced, and genuinely interpretive themes.

KEYWORDS: Qualitative Methods, Themes, Thematic Analysis, Qualitative Analysis, Interpretative

Themes and thematic analysis are very popular in qualitative research, leading some to conclude that thematic analysis far outstrips use by researchers than all other qualitative methods. Whether Braun and Clarke's 2006 work or their subsequent remodeling as reflexive thematic analysis (Braun & Clarke, 2006, 2019), themes and thematic analysis have grown in popularity across all fields and types of qualitative studies. Themes have many different meanings, from patterns to core concepts that describe the participants and phenomenon (Morgan, 2018). Developing themes can be more than a summary of transcripts:

> Themes are not as easily distinguished from their surroundings as specks of gold in a pan. They "are not objects one encounters at certain points or

[1] Corresponding Author: Associate Professor/Doctoral Methodologist, College of Education and Human Performance, West Liberty University, WV, US. E-Mail: david.coker@westliberty.edu

> moments in the text" (van Manen, 1987). They do not stand out as if they were underlined. Anything in the data could be a theme. And then again, perhaps what one sees is only a mirage. Themes can be elusive; they don't have the character of definite entities whose discovery depends on nothing more than technique (Tesch, 1987, p. 235).

Many books and articles describe processes, steps, and problems in thematic analysis and theme formation (Vaismoradi & Snelgrove, 2019), but it is still largely not well documented in many research publications how one leaps from codes to themes.

The paper is not directed toward any specific qualitative methodology, but themes in qualitative research seem to be the most popular end result (Wieler, 2025). Braun and Clarke's 2006 paper is a short 25 pages long, and though they decry that it is spurious that researchers treat thematic analysis as a methodology instead of a method (Braun & Clarke, 2019, footnote), there is talk of vague coding schemes, recoding, and even citation of a popular textbook while stating grounded theory lite is akin to thematic analysis (Braun & Clarke, 2006). A great deal of the initial short paper discusses the need for conceptualization (Braun & Clarke, 2006), but precise procedures are largely absent by design. One notices that their Figure 2 (Braun & Clarke, 2006), which provides an example of themes and relationships, shows no relationships and provides topical themes that they later decry. (The article for Figure 2 is one of the authors from 2003, which shows the obvious influence of grounded theory's coding methods, such as open and axial coding; see Braun and Wilkinson, 2003). While reflexive thematic analysis has a much better developed framework and methods, the purpose of this research and conceptual paper is that many practices are often poorly described because exact procedures were intentionally left to a flexible, open design: "Further, the flexibility of the method / which allows for a wide range of analytic options / means that the potential range of things that can be said about your data is broad" (Braun & Clarke, 2006, p. 97).

Braun and Clarke vociferously proclaim that themes do not emerge but that the researcher is an integral instrument that consciously shapes themes (Morriss, 2024). Coker (2021) argued that the emerge debate is often pedantic, as researchers who use emerge can be akin to an archaeologist digging up the entire artifact for the first time. Despite this argument, Braun and Clarke (2019) rightly pointed out that emerge obscures the processes and thoughts used to construct themes. With the call to move beyond emerge, one can conduct an academic search of major qualitative journals, such as The Qualitative Report, International Journal of Qualitative Methods, Qualitative Health Research, and Journal of Mixed Methods Research, etc. (many more examples abound!), and find many examples that emerge is the process. While no comment on the quality in these journals or the papers using the ominous word is made, the issue highlights a widespread gap in the qualitative analysis literature, and by no means disparage the use of "emerge" to explain processes: Moving from coding and categories to themes is poorly understood and executed, resulting in a lack of clarity in theme formation.

"Theme identification is one of the most fundamental tasks in qualitative research. It also is one of the most mysterious" (Braun & Clarke, 2019, abstract). An extensive literature review, by decade since the 1950s, provides a comprehensive account of the evolution of qualitative research regarding themes and thematizing. The following study has two aims. First, an exploration of patterns and trends in dissertation themes at US universities to elucidate practices for developing qualitative research themes. Secondly, the analysis was used for a metamethodological review and development of a systematic thematic tool that could be used by novice and experienced researchers to improve systematicity. The results of the research and the author's experiences lead to a systematic method—a storyboard for developing, evaluating, and refining themes—with the goal of enhancing rigor, transparency, and results. An agnostic approach is at the center, with the belief that any qualitative analysis could be improved by concrete, systematic methods flexibly applied. The many definitions of theme reveal insights into the conflicting definitions, with three types of themes of interest to qualitative researchers. The following article asks the general research questions: How do doctoral students in US educational administration programs (Ed.D./Ph.D.) describe the development and formation of themes in the analysis and results section of their dissertations? Next, the discussion centers on the problem of "themes emerge," with historical roots that problematize the idea. The theme construction tool and the three-step storyboarding process provide a concrete, actionable method for developing and reporting themes.

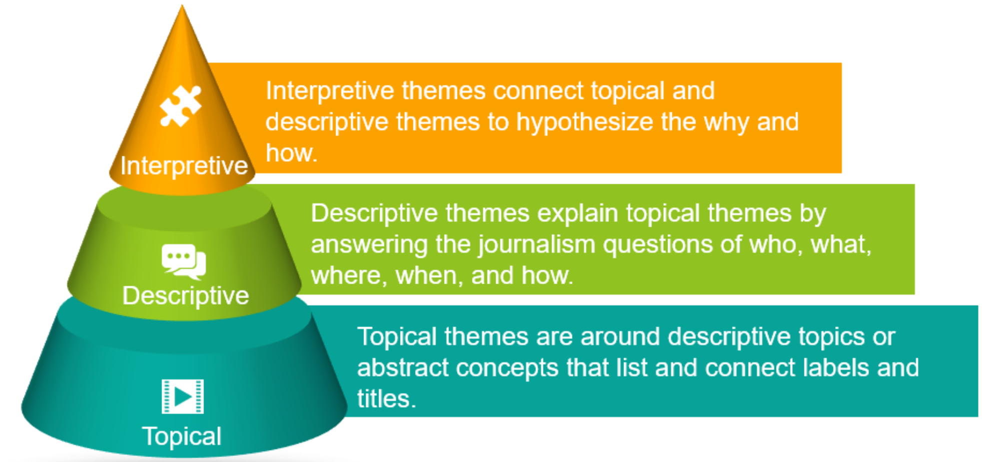
Background

Thematic analysis is perhaps the most common qualitative research method in the 21st century (Fryer, 2022; Naeem et al., 2023). Braun and Clarke's 2006 article is internationally known and widely used in most countries worldwide. The non-theoretical methodology, with its flexibility and adaptability to most any issue, is a key factor in its popularity. Besides the high citation count, ease of use has led to adoption as a first method in teaching qualitative research in many universities. Notwithstanding the belief that Braun and Clarke originated many of the ideas, the following review provides a comprehensive view of the development of themes and thematic analysis since the 1970s.

If one is to construct a theme, then one will have to know what one is looking for. Qualitative research was in a nascent period from the 1950s to 1970s, with little development in the shift from content analysis to thematic analysis, as Denzin and Lincoln termed it, through the stages of doubt, questioning, growth, and eventual acceptance (Jules & Salajan, 2024). The definitions of theme spread in many directions, from the simple to the complex, varying from one-word topics more akin to titles and summaries to interpretations and inferences (Braun & Clarke, 2019; Francis, 1989). Ryan and Bernard (2003) and DeSantis and Ugarriza (2000) provide a laundry list of possible definitions of themes over the past century: categories, conceptual connections, units, inferred rather than concrete, larger units of thought showing interrelationships, a recurrent pattern controlling, etc. More recent scholarship identifies specific components of themes that connect subjective meanings to broader social and cultural patterns to convey core concepts to an audience (Morgan, 2018; Vaismoradi & Snelgrove, 2019). Themes should tell a coherent, logical story that interprets the core concepts and essence of the data, grounded in a single assertion (Braun & Clarke, 2022; Kassarjian, 1977).

Despite many definitions, researchers generally agree that themes unify disparate analyses into a cogent whole that tells a multifaceted story, though some disagree about the simplicity, complexity, or function (Braun & Clarke, 2022; DeSantis & Ugarriza, 2000; Morgan, 2018; Tesch, 1987). Themes can range from grand narratives that touch all participants in a study to diverse, conflicted themes that apply only to different subsets of participants under particular contexts (Coker et al., 2025). We propose a taxonomy of themes, as shown in Figure 1, that provides a hierarchy necessary to develop qualitative analysis from the purely descriptive to the interpretive, if that is indeed the goal (which may not be the case for many qualitative methodologies, content to remain at the descriptive level). The hierarchy moves from the manifest to the latent and from the specific to the general.

Figure 1
Taxonomy of Themes



Rather than seeing themes as better or worse than others, the taxonomy of themes builds on each layer in a reciprocal fashion, with little researcher erudition and abstraction in lower layers beyond the spotlight focus and increasing abstraction and theorizing as researchers develop interpretive themes. Topical themes start from the empirical and can be wholly descriptive, or the development of conceptual; descriptive themes build off topical themes to construct narratives and stories; inferential themes move beyond the data and connect to broader social phenomena, using the researcher's knowledge and theory (Luborsky, 1994; Wiltshire & Ronkainen, 2021). Inferential themes rely on all manners of data analysis supporting one another: "Therefore, themes are inferred or discovered from behavior or symbolic verbal expressions that stimulate activity, direct interpersonal relationships, and unite or illustrate the common meaning of experiences that occur in multiple and varied situations and contexts" (DeSantis & Ugarriza, 2000, p. 335). Only by moving through the taxonomy—from the foundation level to the stages—can researchers be both systematic and open to the research and honor Becker and Geer's (1957) admonition to connect near and far inferences.

The three levels of theme interact: the topical and descriptive serve as the building blocks, and the inferential themes serve as the researcher's translation of the data into results (DeSantis & Ugarriza, 2000; Spiggle, 1994). Topical themes provide labels and can be rudimentary categorizations to the budding of abstract conceptualization, but descriptive and inferential themes comprise two components: the theme, or main idea, and the rheme, which provides new information and elaboration of the theme by explicitly using concepts with linkages between each other within a particular context (Belmonte & McCabe-Hidalgo, 1998; Morgan, 2018):

> By contrast, themes as interpretative stories built around uniting meaning cannot be developed in advance of analysis. They contain diversity, but they have a central idea that unifies the diversity (instead of 'good experiences of healthcare' you might have the theme "validation of my personhood") (Braun & Clarke, 2023, Telling-Meaning, para 1).

By analyzing topical and descriptive themes, analysis can identify linkages and factors that lead to the development of abstract themes from both an insider and an outsider perspective, describing broad, general social practices with concepts at the core (Miles et al., 2014; Spiggle, 1994). A significant concern is research that stops at the topical and/or descriptive levels, which is underdeveloped and provides little erudition beyond outlines and summaries (Sandelowski, 1995). If themes are poorly understood, the processes that develop them are often even more mysterious.

"Shared meaning themes should be rich and multi-faceted, encompassing multiple observations about the central concept (uniting meaning) of the theme" (Braun & Clarke, 2024a, p. 9). The problem is that most texts, such as Braun and Clarke (2006, 2019, 2022), leave a gap: The authors do not explicitly state a leap from analysis to coding and categories to themes. Explanations often default to sorting, categorizing, and interpreting (cf. Özden, 2024), but what these processes look like—and how they operate—remains hidden. Authors of popular textbooks, such as Miles et al. (2014) and Saldaña (2015), provide specificity but lack descriptions of levels, false starts, and rejections. Morgan (2018) provides informative models, but he does not offer concrete methods for transforming analyses into interpretive themes.

How to define and construct themes often remains a mysterious endeavor, with modern scholarship offering few clues. Much of the recent literature either talks in generalities or assumes a hidden link between the data. Braun and Clarke (2006) lack specific methods: "You end this phase with a collection of candidate themes, and sub-themes, and all extracts of data that have been coded in relation to them" (p. 90). Braun and Clarke (2022) state that themes are "developed from codes, are constructed at the intersection of the data, the researcher's subjectivity, theoretical and conceptual understanding, and training and experience" (Conceptualizing, para 7). Thompson (2022) does not even discuss candidate themes; his ideas generally comport with those of Terry et al. (2017), which direct researchers to categorize, promote codes, and test ideas, but all lack substantive methods for concrete cognitive processes. Many authors offer vague possibilities in developing themes: organizing and modeling (Naeem et al., 2023), problems of ending with topical themes instead of interpretive themes (Braun & Clarke, 2024b), interpretation and abstraction (Vaismoradi & Snelgrove, 2019), patterning and collating (Finlay, 2021), conceptualizing patterns and relationships (Herzog et al., 2019), and patterns and interpretation (Peel, 2020). While not all authors had a clear focus on theme construction, the complete lack of mention of candidate themes or prototypes (e.g., Christou, 2022; Dawadi, 2020; Kushnir, 2025; Majumdar, 2019) in articles and guides means researchers have few concrete tools to develop themes or even consider alternatives.

Theme construction beyond generalities and positivist methods, with patterns and collating reduced to absolute counts, fails to show the link between codes and the interpretive process. Braun and Clarke (2024b) are considered the foremost authorities on themes and thematic analysis, but specific methods and routines are left out when they describe how to construct themes: ". . . collapsing or clustering codes that seem to share some unifying feature together, so that they reflect and describe a coherent and meaningful pattern in the data (Braun & Clarke, 2021, p. 63). Popular articles, such as Christou (2022) and Dawadi (2020), summarize Braun and Clarke while omitting transparent, concrete cognitive processes for the construction of themes. Surveying current dissertation practices proved apt, as articles typically face space limitations that obscure many processes.

Theoretical Framework

Many researchers call for developing epistemological and ontological reasoning to position one's research, but this research minimizes philosophizing in favor of a pragmatist approach. Instead of philosophy, pragmatism moves beyond the simplicity of what works to recognize that inquiry is a process in which knowing cannot be separated from doing (Morgan, 2014). Within this philosophical understanding, the researcher is not a static entity with fixed worldviews that bend through philosophy but is inextricably entwined in the inquiry process of seeking knowledge and taking action. Instead of philosophizing, we propose that researchers reclaim an epistemic humility in which they seek to know their beliefs, values, and assumptions by actively seeking disconfirmation and falsification. Epistemic humility is a stance in which the researcher situates actions and results within the specific context, challenging taken-for-granted assumptions that are often uncritically accepted (Kidd, 2016).

Combined with a pragmatic approach that leaves philosophy to the philosophers, ecological systems theory is well-suited to qualitative content analysis dissertations. The different levels are the micro, meso, exo, macro, and chronosystems, from the individual to the connections of different settings to the larger experience without an active role in the whole culture (Onwuegbuzie et al., 2013). All researchers play many roles within research, within the microsystem, on the dissertation journey (which is highly individualistic), and within the university setting, which entails expectations, beliefs, and values. The second level, the mesosystem, is the interaction among the class setting, the class/department, and the social world, all within the local community. The third level, the exosystem, is where the doctoral student is not an active participant but still operates within its confines, such as the discipline, the broader university, and life outside school. Next, the macrosystem imparts cultural values and expectations, such as the larger fields of study and research traditions, as well as, within qualitative research, a person's connections to groups and movements. Finally, the chronosystem involves all the changes over time.

Methodology

A qualitative content analysis of educational dissertations examined thematic analysis practices from a metamethodological frame. The research question was exploratory: How do doctoral students in US educational leadership programs (Ed.D./Ph.D.) describe the process to develop themes? How do doctoral students in US educational leadership programs (Ed.D./Ph.D.) describe the formation of themes in the analysis and results section of their dissertations? The gap in the literature is that it is unknown what specific cognitive processes are used to construct themes. The sampling size was open to being extended beyond any set number, provided that all dissertations met the following criteria: A.) All dissertations must be in educational leadership or educational administration; B.) All dissertations must be traditional five-chapter dissertations in either an Ed.D. or Ph.D. program; C.) All dissertations were only qualitative and used some form of analysis resulting in themes; D.) No more than two dissertations from any one institution; E.) All dissertations were from 2024-2025 and must be full text; F.) Google and Google Scholar were used to search for dissertations, as Google Scholar indexes ProQuest and dissertations in repositories. Dissertations that did not meet the criteria were excluded. Once a relevant dissertation was found, relevant portions were downloaded in Excel, with each dissertation numbered, author/title noted, and the university/college of origin noted.

The results were not designed to be exhaustive, which would be better suited to quantitative research, but rather to assess the transferability of how analytic frameworks are adopted and reported. The author, a frequent contributor to dissertations worldwide as a consultant and professor, sought to generate a sample that could provide insight into patterns and trends in thematic analysis and theme development. Selecting dissertations only in one discipline reduced external variance, as different disciplines enact qualitative practices differently. Homogeneity of the sample also improves the identification of commonality and saturation (Hennink & Kaiser, 2022). The projected range for producing a detailed but not definitive account was projected to be a sample of 15 to 25 dissertations; after 17, no significant divergences or new results emerged, and analysis was stopped. The 17 themes were assigned pseudonyms from P1 to P17.

The qualitative content analysis was based on Lyhne et al. (2025), with the following deviations: 1.) Reading the text, with annotations and notes; 2.) Coding and memoing using in vivo, descriptive codes, categories, concepts, and analytic memos that asked questions, compared/contrasted, and generated hypotheses; 3.) Using categories and descriptive inferences, two stages of sorting and classifying categories were used to develop hypotheses; 4.) Analysis stayed close to descriptive themes, formed and selected by considering the idiographic nature of each dissertation, checking for negative cases and divergences, and considering the relationship between methods and results. Not all steps happened sequentially; once some dissertations were coded, other stages began with memoing and development. Excel was used to organize data analysis, with an iterative process to examine different schemas, as the process was not strictly linear. Saturation was a formal proposition, using intermittent thematic formation (ITF) to evaluate for significant deviations or promising new directions throughout the analysis (Coker, 2022). Two key themes were constructed to describe the data, and, as this was an exploratory study, several miscellaneous but noteworthy results are reported.

Results

The results can be divided into two areas: how themes were constructed and what themes represented. Examples are provided in Appendix A. The first theme is that theme construction depends on a leap from the coding and categorizing to the final results without concrete methods being described. Three major concepts connected to the themes emerge: three central mechanisms for theme development were categorized as methodological nominalism, passive synthesis, and statistical reduction. Methods commonly stated but never made transparent, such as using all data, coding and matrices to allow themes to emerge, codes being good and obvious, or categories being sorted, etc. Formats for themes were primarily descriptive, one-word, and lacked erudition. For most dissertations, results emerged in a leap of faith from the data analysis steps to the final rendition of the themes. All dissertations planned interpretive themes, but most resulted in descriptive themes.

Methodological nominalism meant doctoral students named processes without explanation or demonstration. The following processes were dropped without formal explanation: coded, highlighted, triangulated, delineated, bracketed, refined, classified, synced, and sorted. “Looking for overall impressions (P1), “initial exploration” (P10), and “similar codes were grouped” (P13) were common, where participants named methods but did not demonstrate how and why; the process was assumed without explaining what looking and sorting meant in practice. For example, triangulation was popular but used unevenly. The participants with distinct subgroups (P3, P4, P11, & P16) found convergences and divergences, but for many doctoral students, triangulation was a completely different process through data sources and often included outside sources, such as the literature review and publication documents, such as P1 comparing data to established theory, P3 using public databases, “comparing to the wealth of empirical literature” (16), and many others.

The second aspect was passive synthesis, in which themes emerged without showing the methods used to theorize and thematize. Most doctoral students defaulted to using the verb “emerge” to explain the development, seemingly distilled from “broad categories, patterns” (P2), “allowed me to find the themes” (P4), themes “directly mapped from the research questions” (P11), and the rhetorical question if the theme did not exist then it would now alter the study (P16), as if one had to uncover the hidden gems. Two processes were commonly used. Passive processes for harvesting themes from the data were common in theme construction, using the following verbs: emerged, discovered, revealed, surfaced, coalesced, collapsed, grouped, captured, and extracted. Open coding or other methods were rarely defined or operationalized, and the leap from codes to categories was explained through the process of emergence. Emerge replaced the systematic reporting of analysis, showing that, voilà, the names of methods became themes. Absent were relationships and connections of codes and categories to concepts; false starts and hypotheses were nowhere to be found. “Themes naturally emerged” (P1) summed up the definition of passive synthesis.

The third aspect was statistical reduction, where counts and frequencies dominated the promotion of codes to themes. Some participants offered exact counts of codes, and technology such as computer-aided qualitative data analysis software (CAQDAS) was the driver. Exact numbers and cutoffs for some were used as if qualitative work could be reduced to a frequency, such as “10/16” with a frequency table for the highest occurring phrases (P2), reducing Braun and Clarke’s reflexive thematic analysis to a specific keyword/phrase count (P3), percentage weighting such as 100% (P6), or demonstrating thoroughness by the number of codes and the resultant categories (P2, P17, & P16). Others offered statistical support as the primary reason for the theme's emergence, such as “69% of participants” (P11) and “total meaningful statements and averages” (P3), with both citing standard deviations as well. Some did not offer concrete numbers but conveyed statistical use, such as “popularity,” to derive themes (P16).

Many participants considered methods such as triangulation, member checking, and saturation in Chapter 3 and either never mentioned again or shifted to different activities, such as highlighting (P1), missing the audit trail (P2), and the use of constant comparison, though never disclosed in Chapter 4 (P5), with many more common examples. Saturation was often not defined and a quantitative endeavor, from the completely arbitrary three minimum of the sample reporting a code/category called “consensus” (P9) to topical redundancy of hearing the same thing twice (P9, P11, & P16). Noticeably absent from most reports on theme development was the examination of significatives and whether codes were evenly distributed (that is, the number of codes might be from a few participants, masking commonality).

The incoherent use of methods related to how themes were formed. For example, Braun and Clarke hold that researcher instrumentality is a core concept, yet many participants transformed reflexivity not as a core concept but as a process to “control influence” (P11) and “minimize preconceptions” (P9) to claim reflexivity as a protective shield against bias to ensure trustworthiness and credibility. Another common problem was the use of phenomenology, where participants developed plans rooted in popular phenomenologists (e.g., Moustakas, Smith) and used the language correctly in Chapter 3 (e.g., bracketing, intentionality, noesis/noema, etc.) but reverted to other methods in Chapter 4. Shifting to Braun and Clarke, concrete and popular methods were used by many as substitutes for phenomenological methods in Chapter 3 (P2, P4, P8, & P16), and there was also misuse or absence of terminology. For example, bracketing was operationalized in unintended ways that defeated its purpose, such as “bracketed … into meaningful themes” (P3) and “reflexive bracketing,” yet it was used with preconfigured deductive and topical themes (P14). More striking is that most participants who planned on bracketing never mentioned the process again in Chapter 4 (e.g., P9, P13, P14, etc.). These same participants did not mention noesis, noema, essence, etc., which are core concepts in phenomenology.

The second theme was what themes represented, with themes largely concrete, descriptive products represented in a linear, smooth fashion. Themes were often descriptive and more like category titles (e.g., P1 had labels for categories, P2 had statistics to derive themes, P4 had grouped by one-word emotion titles, etc.). One-word themes or short phrases predominated, such as “fear” (P4), “collaboration” (P6, P9), categories (P13), and “friendship” (P1). Most analyses produced 3-15 themes, with most in the recommended range of 3-7, with some using subthemes and subheadings. The ones with deductive frameworks lacked divergences of challenges to the a priori framework (P1, P2, & P14). Universally, most everyone then began theme explication with the format “participants described,” with few participants offering the how/why of a theme beyond the data. Hypertypicality was common, where for many researchers, everyone, everything, all the time fit into a neat, wrapped package of themes that lacked divergences, questions, and contradictions: "There were no discrepant cases reported" (P4) was the norm, not an anomaly. Themes as topics seemed to be the root cause, as mapping relationships to develop concepts proved elusive.

While many produced no relationships (P1, P4, & P15 jumped directly from categories to themes), some stated relationships by being a "primary driver" (P17), "interact" (P9), and "mediated" (P11, P12) without formal elucidation of the dimensionality and contextual variables. Providing descriptions rooted in the participants' language was the majority, with few offering an interpretive leap by expanding themes to include a "because" and "though" (P1) framework or a how/why, such as "based on" (P17), that were beyond the transcripts and involved the researcher using abductive reasoning. Macronarratization, through grand narratives, guided most results, leading to stories that lacked multivocality, outliers, and stubs that did not fit through narrative smoothing by "collapsed," "merged," and "coalesced" (P1, P2, P3, P4, P16, etc.) to create "composite summaries" (P3). Most researchers lacked specifics on contradictions, ideas on the fringe, and nuances. Divergences that did exist were minor and rare, such as P1 noting one difference between teachers, P6 stating a difference between views of timing and accessibility, and P16 constructing a "miscellaneous theme" for that which did not fit.

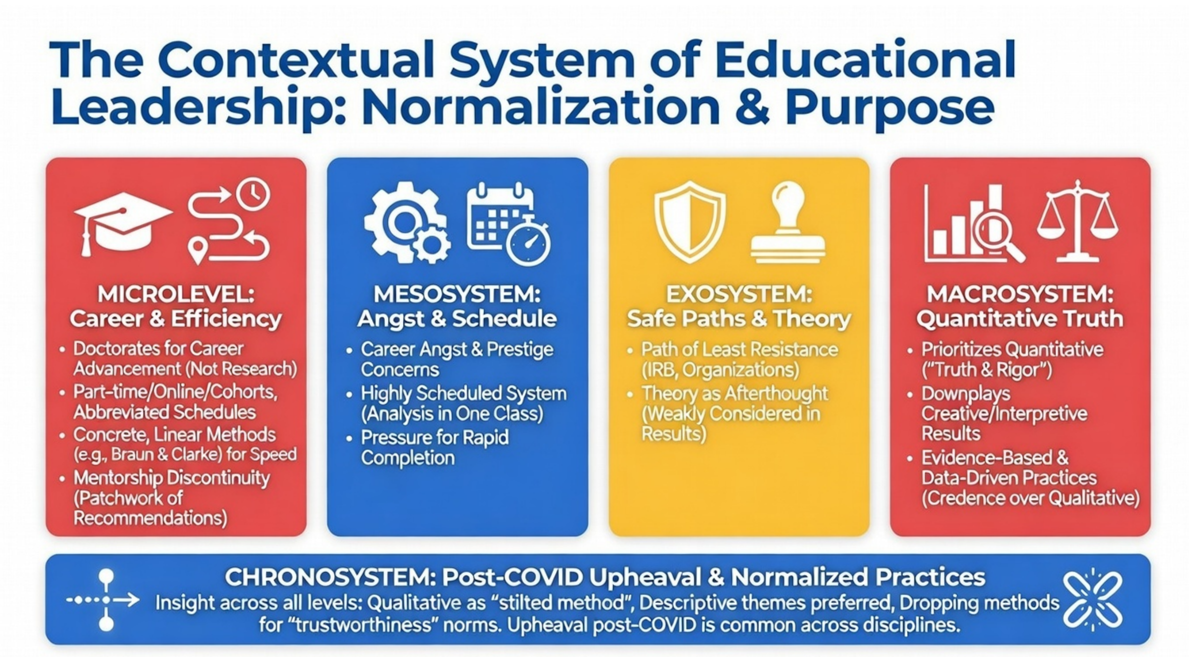
Evidence to support themes was quotes, with some Chapters 4s having almost 30% of quotes. The data shows that the most frequent shift is the simplification of logic. Researchers often promised hermeneutic or interpretive depth in Chapter 3 but delivered statistical or topical descriptions in Chapter 4. This is typically driven by software (P2, P3, & P14), which rewards counting and merging over deep theorizing. Quotes were the results and reduced to phatic statements, where the data speaks for itself, and were often substituted for analysis. Long quotes, often exceeding 180 words and presented in series, were not uncommon, and instead of explicating quotes as analytical or illustrative, researchers often introduced them as if they were the results, with phrases such as "participant stated." The majority of the sample defaulted to descriptive themes, presenting quotes as results with little explanation of why a quote was chosen or whether it was used for analytical or illustrative purposes. Common introductions to quotes used words such as described, stated, and shared, with no explanation or analysis. For example, giving the name and a verb, followed by a quote, was a standalone process used by many. The few researchers who were interpretive frequently used explanations to justify why a quote was representative and often used snippets rather than long quote strings. P17 is a prime example, aggregating different participants into a single sentence and explaining whether a quote illustrated a point or was used in analysis to develop themes. P1 was another interpretive researcher: Quotes began with explanations, and snippets were frequently used with contextualization and for comparison.

The final common issue, though not an explicit purpose of the research, was the division of what can be labeled limitations and delimitations. Limitations are problems outside the researcher's control, such as participants dropping out, poor interview participation, participants not being qualified, or technology problems, which were largely ignored (P3, P4, P11, & P16). Delimitations are bounded, planned research parameters to manage a project and were an apologist tour; examples include stating an intentionally small sample chosen was a limitation (P1, P4, P5, P7, P8, P9, P10, P13, P16, & P17) but was the defining feature of the qualitative study; complaints of selecting from a small region (P1, P2, P5, P6, P7, P8, P13, P14, P15, & P17), which was used as a facet of backyard research; complaints of generalizability (P4, P5, P7, P8, P11, P14, & P17), though the researcher intentionally did not chose a quantitative study with a representative population. Researchers often confused delimitations with limitations.

While it is easy to consider these problems failures, the context in which educational leadership operates suggests that these practices are normalized and serve a purpose, as shown in Figure 2. At the microlevel, educational leadership students often pursue doctorates, even a Ph.D., not for future research but to advance their careers as practitioners. Most doctoral students in educational leadership, like students in many other fields, work part-time, online, and are in cohorts with abbreviated schedules. Shifting from the time-consuming development of interpretive themes or the messiness of phenomenology means being concrete and linear in methods and steps, such as frequencies, the six steps of Braun and Clarke, and descriptive themes become default processes to complete within the truncated timeline. Another salient issue at the microlevel is that educational dissertations are often initiated under multiple professors and, with changes in dissertation committees, lead to discontinuity in mentorship and a patchwork of recommendations.

Figure 2

The Hidden System of Theme Development in Educational Leadership



Within the mesosystem, students are under great pressure to finish, with career trajectories and prestige concerns to keep up with a very structured system (most educational leadership students complete their analysis in one class, which might span 5 weeks to a semester). Moving on to the exosystem, researchers in educational leadership play it safe, trying to find the path of least resistance with the institutional review board and mindful of organizations that are often the research sites. Theory is often an afterthought in educational leadership as a necessity in Chapters 1-2 and then weakly considered in the results; the overwhelming number of dissertations scarcely considered theory in the results. At the macrosystem level, educational leadership shares much in common with other fields, such as public health and the so-called hard sciences, which prioritize quantitative results as a mark of truth and rigor while downplaying highly creative or interpretive outcomes. Common practices in educational leadership are evidence-based practices and data-driven instruction, which give credence to quantitative research over strict qualitative research. The chronosystem provides great insight across all levels: Many educational leadership programs approach qualitative research as a stilted method, with descriptive themes preferred as interpretive methods might be considered opinions, and dropping a list of methods to "ensure" trustworthiness and credibility becomes an unstated norm to move in lock step through the dissertation process.

Educational leadership is not alone in these practices and problems, as many disciplines have experienced upheaval post-COVID.

Validity and Reliability

Validity and reliability in qualitative research entail examining if the results are supported by the data and are trustworthy through transparent, rigorous adoption of methods (Golafshani, 2003). Lyhne et al. (2025) posited that reliability and validity can be assessed through systematic coding and analysis, intracoding by a single researcher, and the researcher's background and experience. Systematic coding began with reading and retaining the dissertation sections that detailed the methods and results for each theme. Then there were in vivo coding and descriptions, with memoing and self-questioning throughout. The project came with starts and stops over two years and, at the end, included an entire reconsideration of the methods and results by returning to the original data. Using intermittent thematic formation, the researcher cycled back and forth between the analysis and the raw data to evaluate convergences, divergences, and whether any new ideas had been overlooked previously (Coker, 2022). By being recursive, the researcher moved beyond confirmation to consider negative cases that challenge thinking and the null-hypothesis trick, in which the researcher attempts to falsify nascent arguments. By being recursive in this manner, the researcher also continuously returned to check if the results are grounded in the data.

The second aspect of validity and reliability, according to Lyhne et al. (2025), is the researcher's experience and background. The researcher works in educational leadership, teaching doctoral students, serving on dissertation committees as chair, committee member, and methodologist, and providing consultancy in the US and nationally. In the researcher's background, he has conducted extensive qualitative content analysis of dissertations across many publications; through this study and practice, the researcher has analyzed 100s of dissertations formally. Like Braun and Clarke (2023), he has found that researchers sometimes struggle to understand descriptive and interpretive themes and lack transparency about how they formed their interpretations.

Part of the examination of validity and reliability is cycling back to theory and the extant literature. Besides applying systems theory to one's own work, the current study suggests that researchers often clashed between a closed system and an open system in theory, but, in practice, operated as an open, fluid system with eclectic practices. The researcher was mindful of Kim's (2015) admonition to avoid the grand narrative by not overly smoothing the data, which resulted in descriptions that went beyond what fit the data's dimensionality and outliers, and by being open to new directions, not intended for data-driven decision-making. As mentioned about incoherent methods, many other researchers have found similar: Braun and Clarke (2023) term a mishmash; Holton (2008) discussed Glaser's rumination on the misapplication of grounded theory; Tracy (2010) postulated a global view that a study is incoherent if it does not achieve the stated goal, is not connected to the literature, and uses methods not suitable. Coker (2023), in a direct study of dissertations, found that the methods reported in dissertations often did not align with actual practices. As a reflexive researcher, there is the hypothesis that the dissertation is a long, complex process with many starts and stops, different mentors and guides, and the researcher growing and learning, which creates numerous revisions that often cannot be used or become overwhelming.

Moving beyond emerge requires an active, recursive process chronicled throughout. The lack of transparent reporting of the transformation from topical identification, descriptions, and narratives, hallmarks of deductive and inductive reasoning, to abductive reasoning is probably much more complicated than "it did not happen." First, many doctoral students probably had a limited repertoire of tools for abductive reasoning. Secondly, chronicling and analyzing abductive reasoning is a process in and of itself that must be formally planned and executed before collecting data. Finally, abductive reasoning takes an inordinate amount of time and theorizing, which is probably why students who usually lack the skills and time shy away from interpretive work. Abductive reasoning requires a conflictual view of analysis in both hypothesizing and finalizing, challenging the self, the process, the raw data, and the results, which many might find uncomfortable to engage with.

Discussion

The research, along with my positionality as a dissertation consultant and professor, suggested the results were the norm and aligned with the findings of Braun and Clarke (2019) that theme development often departs significantly from recommended practices, lamenting that thematizing was difficult and often misunderstood in the leap from the topical to the interpretive. A key departure is analyzing the nuts and bolts, or actual practices of developing themes from raw data. Qualitative research exists on a continuum, with four quadrants: exploratory-descriptive, explanatory-descriptive, exploratory-interpretive, and explanatory-interpretive. All research is interpretive by the very nature of what to include and exclude as topics and what theories informed the research, but some research is much more so than others (Finlay, 2021). Exploratory is open to the data, guiding many of the results and often is inductive, while explanatory generally has predefined purposes or aims to evaluate, often with deductive reasoning. Descriptive is near inferences, sticking much closer to the data, while far inferences rely on abductive reasoning beyond the data. Most research is a blur of the different quadrants, as few research will be completely pure.

Research is only as strong as its weakest link in the chain of evidence, and without considering, evaluating, and transparently presenting multiple hypotheses for refutation in part or whole, results lack the credibility beyond more than the researcher being attached to an idea or position by foreclosing possibilities before or at the early stages of a research project (Burnard, 2024; Sagan, 2007). Although claims of far-reaching inferences are common, the ubiquity of descriptive themes and the corresponding rise of exploratory-descriptive research can be traced to three primary reasons. The first reason is that exploratory-descriptive (often explanatory-descriptive despite protestations) is a linear, straightforward project: code the data, group by similarities, and write up summaries. In exploratory or explanatory inferential research, researchers do not know what they are seeing or where they are going, leading to many iterations, possibilities, and false starts (de Farias et al., 2021; Luborsky, 1994). As opposed to interpretive qualitative research, which is rhizomatic, ambiguous, and one loses one's sense of direction through the internal dialogue of doubt (Tesch, 1987), descriptive research is well-suited to artificial intelligence techniques, such as autocoding and counting (Xu, 2026). In many ways, the data speaks for itself, as the chain of evidence is straightforward, but issues such as metonymy and synecdoche may not be as well understood (Spiggle, 1994). Courses and textbooks often provide scant coverage in qualitative analysis and even less guidance in how to synthesize beyond codes and categories to concepts and interpretive themes (Coker, 2023; Lucas & D'Enbeau, 2013). Such research is important, but it is probably a byproduct of a skills gap paired with a rush to complete a dissertation/article, receive promotion, or earn tenure. Replacing exploratory-interpretive with explanatory-descriptive, especially with small samples and one interview per participant, is decidedly more straightforward and often accepted without the complaints that interpretive methods are merely opinions.

In in-depth, well-designed interviews, a refracted space is created between the participants and the interviewer, diving into an unknown sea of possibilities (Coker & Akande, 2025; Tavory, 2020). Highly descriptive qualitative research operates in a much different fashion, as many research studies suffer from a data desert. Unlike quantitative studies, which must demonstrate the reliability, validity, and applicability of the instrumentation, interview protocols in qualitative research are often poorly developed and lack evaluation. In articles, due to space limitations, there is little mention of how protocols were prepared, developed, and refined (Coker & Akande, 2025). Reviewing protocols, there are often too many questions in too little time, making it difficult to construct interpretive themes. Sometimes researchers even map each interview question as if each question produces a theme, foreclosing interpretive schemes because the results are known a priori. Questionable or poor instrumentation makes it difficult to dig deeper than the surface level. In highly descriptive qualitative research, short interviews are often more like interrogations than conversations.

Finally, descriptive qualitative research as a proxy for claims of interpretive research is probably a product of compartmentalization, in which researchers place themselves in silos and believe that qualitative methods must be exact and pure, thereby limiting growth, skills, and possibilities (Guest et al., 2011). Some researchers claim that they are limited because so-and-so researchers did not mention something, though most researchers reject proceduralism and rigid methods (Braun & Clarke, 2019). This scenario would be like a researcher conducting a simple linear regression and claiming they cannot perform imputations or handle extreme nonlinearity. While not advocating anything goes, there is no single scientific method, so researchers should practice flexibility and adopt hybrid methods where appropriate, as long as their aims remain in focus (Sandelowski, 1995; Swain, 2018). The key is that researchers need to be systematic and explicit, as many overly descriptive results seem to stem from a lack of a method for constructing highly interpretive work. Philosophical debates are best left to philosophy; practical solutions that can be implemented rigorously to meet the research's purposes and aims should be the guiding principle. Part of the shift in thinking is that researchers need to replace "following" a method with "based on." Glaser once famously declared that all data is data, and the recommendation is that researchers should get lost and find oneself as the goal of the knowledgeable researcher.

The results, along with a plethora of searches across major journals in most disciplines, unequivocally show that themes do emerge. The claim is so prevalent that one sees many results that wave the wand, conduct the analysis, and produce a ready-made theme that just emerges out of thin air. Braun and Clarke (2019) complain vociferously about this nefarious practice:

> Themes do not passively emerge from either data or coding; they are not 'in' the data, waiting to be identified and retrieved by the researcher. Themes are creative and interpretive stories about the data, produced at the intersection of the researcher's theoretical assumptions, their analytic resources and skill, and the data themselves. (Braun & Clarke, 2019, Conceptualizing TA, para 6)

No matter what is said, many researchers cannot escape that for all the methods, analysis, and sophistication, the entire process devolves into an esoteric, invisible claimed process of emerged. Coker (2021) pointed out that the level of analysis and connection to data make claims of emerge plausible when researchers are stuck in unstructured processes that lack a clearly defined model.

Modern researchers might be surprised to learn that Braun and Clarke's complaint of emerge has a history that long predated either, or that some research traditions embrace the concept of emerge. Glaser (2008) stated repeatedly over a long career that emerge is a right and proper method. Others, such as Sandelowski (1995), complained that emerge is false, as the researcher is the one who constructs a meaning inexorably constructed out of one's head: ". . . no themes, categories, theories, or anything else of analytic importance will 'emerge' without the researcher who must (in the pithy words of Captain Picard on Star Trek) 'make it so'" (p. 317). Citing Pierce, inferences and analyses were often beyond participants' awareness, rendering it challenging to claim systematicity (Fischer, 2001). The late Renate Tesch (1987) pointed out that researchers were often so enmeshed in their phenomenon that their intuition and deep immersion in the data made anything more than emerge difficult to bring to consciousness.

There is a suggestion that researchers should evaluate the emerge debate as a larger structural deficiency: Researchers might fail to transparently report methods because they lack skills with concrete, reproducible methods to construct concepts, hypothesize themes, and evaluate for final results. Themes might be less about results and absolute truths and more about constructing themes from data digested through the researcher's psyche (Braun & Clarke, 2019), but such a transformation would require methodical, visible methods. One such method is abduction, in which forms of logic such as induction and deduction are used to formulate possible interpretations of the phenomenon under study (Karlsen et al., 2021). The following suggestions give researchers a concrete method to move beyond emerge by providing interlocking stages for creating, hypothesizing, and evaluating themes in a rigorous manner. The methods are eclectic, making use of strategies and processes from a diverse set of qualitative analysis traditions.

Thematic analysis. Grounded theory. Basic qualitative. Phenomenology. Qualitative research can be wholly descriptive and attempt to exclude the researcher's opinions and views, but most qualitative methodologies believe that researchers can perceive relationships by applying external sources of knowledge beyond observations and directly acquired data to develop interpretations that offer possible causal conclusions (Maxwell, 2021; Russell, 1935). One recommended, but underused practice is creating narratives and story frames (Vaismoradi & Snelgrove, 2019). One can go on and on about different methodologies and methods, and most methods have different proponents, often with radically different aims and purposes, but there is the proposal that an eclectic view that any method that works and is congruent with one's purpose can and should be used: ". . . it does not make sense to exclude a particular technique because of personal discomfort with it, or misconceptions about or prejudices regarding how and why it might be used. . . ." (Guest et al., 2011, Intro, para 2).

Future Recommendations

As Coker et al. (2025) stated, "... themes need to be reimagined from the ground up, starting with the union of themes and rhemes. By systematically designing, executing, and reporting each stage, themes can be improved, grounded in the data, and account for researcher instrumentality" (Themes, para. 1). For most researchers, providing a structure to the analysis ends in creatively developing themes that are tied to broader social phenomena instead of particular contexts, with a structure, coherence, and order that brings a semblance of meaning to the data (Luborsky, 1994; Mishra & Dey, 2022; Spiggle, 1994). Common problems extend in two directions: underdeveloped themes that are essentially summaries and at the surface level, and overelaborated, prescriptive themes with many conditional statements (Vaismoradi & Snelgrove, 2019). The theme construction tool provides a three-stage storyboard process to develop topical, descriptive, and interpretive themes in a methodical, transparent, and visible way. The application is not defined by any particular qualitative methodology or methods, and researchers should not think the process is rigid or linear; recursive, iterative, and adaptive should apply to every exploratory qualitative research process, which Braun and Clarke (2006, 2019) remind researchers of numerous times. The results suggest researchers need a systematic approach to developing themes in thematic analysis, with a three-step approach as a possible solution. The following methods, developed after reviewing the history of theme development since the 1950s, the analysis results of dissertations, and the author's experience, provide concrete approaches for conceptualizing themes across three levels.

The following procedures, refined through the study as both a literature review and a formal analysis of dissertations, offer methods that could be adopted by researchers across all disciplines. A particular use could also be to teach would-be researchers systematic methods to conduct qualitative analysis. Building on a pragmatist approach, the recommendations move beyond philosophizing and hold that what works and one's reasoning are shaped by the social context in which research is conducted, a context that researchers must recognize (Morgan, 2014). There is a logical structure: No researcher (should) starts with the conclusions or works back (though some theorize researchers unwittingly do just that). All researchers, whether inductive or deductive, descriptive or interpretive, exploratory or explanatory, commence with familiarizing and describing the data, and through multiple, recursive stages, researchers analyze the data to develop descriptive and then interpretive themes (cf. Miles et al., 2014; Saldaña, 2015; as well as any seminal author on grounded theory, thematic analysis, phenomenology, etc.). The steps are envisioned as tools-at-hand, where researchers can map systematic practices and be eclectic in adopting methods as the purpose and needs of the research dictate.

Storyboard Stage 1: Conceptual Clarity

In the first stage of the theme construction tool, the gateway to interpretive themes involves developing concepts from the categories and topics constructed through coding or other analysis. A holistic picture should be kept in mind: topical themes around concepts and descriptive themes must be formed first before one can start hypothesizing and evaluating potential themes. Lacking clear models for conceptual development, researchers tasked with creating concepts often resort to the default process of sorting and summarizing topics (Ahmed et al., 2025). Concepts are the beginning of a linking process, where the final themes provide researchers with their own interpretations to explain the links (relationships and associations) of the concepts to the phenomenon, which in turn reciprocate because the concepts are grounded in the data and move beyond a list or pile of categories (Miles et al., 2014). The proposal is that researchers use concept boxing to move beyond crude classifications to the construction of concepts that explain the linkages within and between categories.

"Concepts are socially constructed abstract linguistic terms that cluster related categories of knowledge or experience into generalized, collective meaning" (Wynn, 2024). The first step in concept boxing is for researchers to develop categories linked to proposed concepts, which can proceed either deductively, inductively, or, as is most likely, a combination of both (Spiggle, 1994). Categories can start with, but must be much more than simple titles, list sorts or counting codes; being methodical should involve the use of systematic methods and models, such as matrices, network arrays, counts, patterns, clusters, literary devices, etc. (Miles et al., 2014; Sandelowski, 1995; Tesch, 1987). Two primary methods exist for results patterns. Weak patterning is strictly counting codes, sorting under a classification or category title, and then moving to developing summaries. Strong patterning using the techniques in weak patterning but applies systematic, documented constant comparative methods at the individual, subset, and group level both horizontally and vertically, considers contextual matrices using journalism questions, the dramaturgical of before/during/after, conditional connections, dimensionality, and variation appreciation (Spiggle, 1994).

A conundrum exists here that is the axis of most problems that turn highly inferential, interpretive themes into descriptive, topical themes, even with patterning: Few researchers know or seem to fail to report how to free themselves of the positivist trap of being quantitative researchers in sheep's clothing by reducing everything to a tidy label (a code) and then a discrete count. Several methods used concurrently can move beyond counts, which often are the foundation of all categories and concepts.1 One such method is significatives-evaluatives-evocatives (SEE): 1.) Significantives: Indications of importance; 2.) Evaluatives: Indications of judgment; 3.) Evocatives: Indications of emotion (cf. Buetow, 2010; Luborsky, 1994). Using SEE at the individual, subset, and group level can move beyond a strict count level, where researchers can code and memo the whole and part by what participants claim they think is/is not important, judgments made about topics in the data collection, and strong emotions or feelings by analyzing multiple vignettes with consideration of the 3-D environment of time, space, and people (Coker et al., 2025).

As shown in Table 1, the antecedents-phenomenon-consequences (APC) concept boxing provides researchers with a concrete iterative method to move beyond categorization and classification schemes while still producing near-inferences to the data. Concepts, at the most basic level, are abstract constructs, more general than categories/classifications, possess relationships or linkages to attributes, characteristics, and/or variables, operate on a continuum (dimensionality), might be connected to theory, and demonstrate mechanisms around three focuses: conditions, actions-interactions, and consequences-outcomes (Spiggle, 1994; Vollstedt & Rezat, 2019). Mechanisms, not processes, are the purpose of concepts: Processes connect different concepts to describe the larger pattern and movement (which will later be used to construct themes), while a mechanism is a discrete construct abstracted for specific contexts and situations (Van Meegdenburg, 2023). A method for considering possibilities is to use Glaser's 6 C's: causes, context, contingencies, consequences, covariances, and conditions (Vollstedt & Rezat, 2019).

Table 1
A Framework for Concept Boxing

Component	Description and Guiding Questions
Core Structure: A-P-C	The iterative framework is built around the Antecedent-Phenomenon-Consequence model. Start by hypothesizing links between categories, using "???" for unknown elements.
Antecedent	What happened before the phenomenon?
Phenomenon	What happened during the phenomenon?
Consequence	What were the results or outcomes of the phenomenon?
Analytical Lenses	Methods to apply across each part of the A-P-C structure: • **Phenomenon Level:** Conditions, actions/interactions, consequences/outcomes. • **Factors to Consider:** Causes, context, contingencies, consequences, covariances. • **Quadfecta:** The interplay of thinking, feeling, doing, and valuing. • **Constant Comparative:** Compare findings within single participants and across the entire dataset.
Variable Linkages	Identify the types of relationships between variables (categories). These can be linear, cyclical, complex, recursive, etc. Examples include: • **Causal:** A → B • **Sequential:** A happens before B • **Bidirectional/Simultaneous:** A ↔ B • **Mediating/Moderating:** A → B → C • **Covariance:** A and B both increase/decrease; or as one increases, the other decreases.
Contextualization	Situate the A-P-C model within its broader context:

Table 1
Contunied

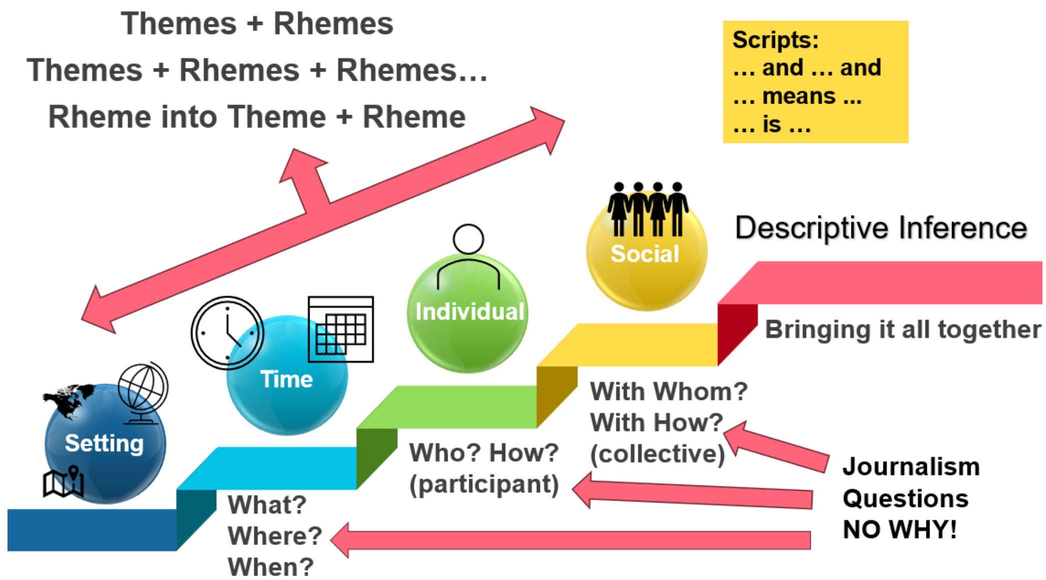
	• **Temporal:** When did events happen (before, during, after)? • **Setting:** What are the circumstances (where, when, what objects were involved)? • **Participants:** Consider the participant as an individual and within their social role.
Dimensionality	Analyze concepts as a continuum rather than as binary states: • **Intensity:** What is the strength of a concept/component? • **Frequency:** How often does a concept/component occur? • **Duration:** How long does a concept/component last when it occurs?
Concept Test	A litmus test to distinguish a true concept from a simple category: • **Concreteness:** Is the concept abstract, or can you physically "touch" it? (If concrete, it is not a concept.) • **Linkages:** Does it describe a relationship between variables, or is it just a title for a list? • **Origin:** Is the concept implicit and inferred by the researcher, or was it explicitly stated by participants? • **Contradiction:** Are there contradictory or alternative concepts that could also explain the phenomenon? • **Narrative Richness:** Can you write a rich, descriptive narrative around the concept?

Note. A helpful tip is to turn the variables into simple sentences to clarify the proposed linkages between them.

In concept boxing, multiple variables will be considered under categories to construct mechanisms. Three important points must be considered. First, linkages should drive the connection of variables to develop categories; linkages can have relationships, but associations with unknown relationships will be common. Rarely will a concept not involve multiple linkages, but care must be taken to address issues of vagueness and ambiguity (Miles et al., 2014; Wynn, 2024). Though Braun and Clarke have evolved in their thinking, their recommendation to use constant comparative in interpretive qualitative design still holds promise: "Although identified here as a key feature of GT [grounded theory], we recommend a recursive approach like this [constant comparative] as essential for rigorous qualitative analysis in general" (2013, p. 214). Secondly, variation at the concept level is to be expected with descriptors such as most, some, and few that change over time and space; unlike final themes, which must be internally and externally coherent, concepts can rival each other and even be contradictory, such as one subset of participants acting in a certain manner contrary to another subset. Third, a general rule of thumb is 30-50 codes, which will be placed into categories, and then concepts. The number of concepts will probably be 10-20, but it should generally be fewer than the number of categories.

Once researchers have a list of tentative concepts and topical themes, the next step is to map each topical theme (concept) onto descriptive themes, as depicted in Figure 3. Descriptive themes are low-level inferences in salience that provide summaries and syntheses about each concept through thick descriptions; researchers must make conscious decisions about what to include, what to exclude, the order, and importance (Luborsky, 1994; Saldaña, 2015; Villamin et al., 2025). There will be a significant amount of data shedding from coding and categories to develop summaries. Matrices around the journalism questions of who, what, where, when, and how—noticeably excluding the why—create systematic, multiple descriptions of narratives, with qualifiers of intensity, duration, and frequency, using words such as most, some, and all (Miles et al., 2014). Each concept can have contradictory descriptive themes, as not all participants act uniformly (Liamputtong, 2009). Themes need rhemes, which are elaborations and clarifications of the order of events, sense-making, and recognition, with linear types, continuations of themes with further elaborations, and the elaborations themselves as themes with further elaborations (Belmonte & McCabe-Hidalgo, 1998; Tesch, 1987). Themes might have multiple rhemes, and rhemes can be promoted to themes in follow-ups if researchers deem them important. Going back to the original data sources, codes, and categories adds extra information and can provide analytic and illustrative codes for further analysis. Again, there should be variation and contradictions at this level.

Figure 3
Constructing Descriptive Themes

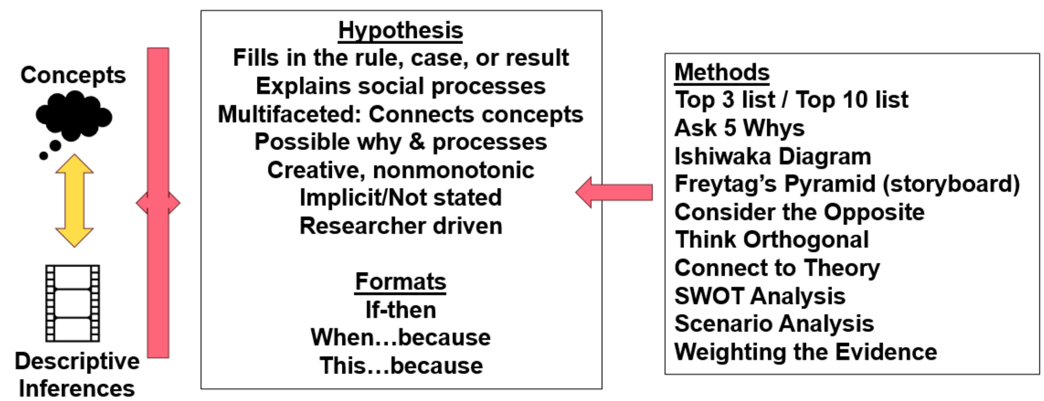


Topical (categories and concepts) and descriptive themes are the gateway to interpretive themes that rely on greater inferential power that develops "far inferences." Instead of standing in opposition to interpretive themes, topical and descriptive themes tell the story and provide a framework for developing hypotheses and later evaluations, serving as more than summaries but as syntheses (de Farias et al., 2021). Topical and descriptive themes are not set in stone; they are not a one-and-done activity. Rather, ending with descriptive themes lays the foundation for the following two stages of the storyboard by producing a narrative, showing important actors, and any change (Polkinghorne, 1995). The entire process is recursive and iterative, with constant attention to detail and to challenging one's tentative results. Closing analysis too early and staying at the superficial level are always concerns, but researchers are also wed to the quality and quantity of their data. Using the topical and descriptive themes provides a springboard to the next stage, and a common shortcoming is the absence of concepts and descriptive themes, which can cause a short circuit when necessary intermediary steps are missing to construct interpretive themes.

Storyboard Stage 2: Hypothesizing Possibilities

In the second stage of the thematic construction tool, researchers generate possibilities to explain processes and the why. Braun and Clarke (2006, 2024b) recommended candidate themes that tell the story, or narrate phenomena, which are hypotheses that theorize explanations and processes to interpret broader social patterns of behavior and thought. Narratives have a beginning, middle, and end, with characters, goals, and objectives, moving beyond a collation of events to movements and change that link one event to another (Kim, 2015). The goal of interpretive themes is to produce narratives that transform descriptions of events into coherent plots that merge the participants' data with the researcher's interpretations (Polkinghorne, 1995; Spiggle, 1994): "But compelling insights and aha moments do not come from an inventory of responses. They arise from deeper analysis and interpretation requiring extensive engagement with participants' emergent discourses, existing theory, and the contexts in which they all are embedded" (Lucas & D'Enbeau, 2013, p. 214). One way to accomplish hypothesizing is using abductive reasoning with concrete methods and formats, as shown in Figure 4, which has been variously called hypothesizing, inference to the best explanation, and retroduction (Fischer, 2001; Paul, 1993).

Figure 4
Abductive I Processes to Develop Hypotheses



The term Abductive I means hypothesizing a posteriori, using topical and descriptive themes to project possible themes that move beyond the concrete and descriptive to include either one or both components: hypothesizing the why or implicit processes by moving backward from the results, forming a possible explanation/process (rule) to explain the case by testing different frames of reference (Fischer, 2001; Khurshid et al., 2024). Formats should ask if-then, if-but, if-however, and/both, or because statements to develop conditional hypotheses (Saldaña, 2015), and one should intentionally form rival and dueling hypotheses. Abductive reasoning involves reading beyond the data to form explanations and processes to describe why and how persons and groups operate; instead of reading between the lines, abduction is reading beyond the lines to create new data that does not exist in the data directly collected (Miles et al., 2014).

Abductive hypotheses work backward for sense making by the researcher, with the case and results (the phenomenon) generally known, but researchers must be creative to go beyond the data to formulate rules—the explanations that answer the why or the process that answers the how (Fischer, 2001; Karlsen et al., 2021; Mukumbang, 2023; Sandelowski, 1995; Tesch, 1987). Good themes involve linking multiple concepts to form hypotheses, so researchers need to be aware of overstretching and overelaboration when forming themes (Calzavarini & Cevolani, 2022). Rich, thick descriptions presented in a creative format (Vaismoradi & Snelgrove, 2019) connect the topical and descriptive to the hypothesized themes. Morgan (2018) stated that a collection of concepts constitutes a low-level theory, whereas better theory explains the inclusion of concepts and their interconnections. While Morgan envisioned strong themes similar to theory, the pragmatics is that concepts generally should not be mapped 1:1 to themes. Nonmonotonicity means that concepts do not all have to move in unison or uniformly over time (that is, change can be multidirectional and might apply to a subset but not the whole), and the more parsimonious and plausible, the better (Paul, 1993).

Interpretations operate beyond observable, direct data by theorizing linkages between what was observed and what was hypothesized, if all other contexts and conditions remain stable and unprovable absent a follow-up study (Miles et al., 2014; Thagard & Shelley, 1997; Van Meegdenburg, 2023). Two interpretative traditions, developed by Ricoeur, are the hermeneutics of suspicion and the hermeneutics of empathy (or faith). Both manners provide interpretive insights that postulate conditions not directly stated in the data, as suspicion involves reading beyond the text to construct hidden meanings (a common practice in all coding), while empathy seeks to discuss meanings directly with participants (Willig, 2012). Both approaches provide different insights, and when addressed with disruptive reflexivity—where researchers actively consider their biases from the positionality of self—hypotheses and evaluations can go beyond preconceived notions while speaking from below (the empathic) and from above (the suspicious) (Reichertz, 2010). Reciprocal determinism, from Bandura's theories, provides further nuance by suggesting that causes and effects often occur in tandem, as people can act vicariously.

Storyboard Stage 3: Interpretive Decision Making

The final stage of the theme construction tool, Abductive II, is the systematic evaluation of hypotheses—those candidate themes in Stage 2—for inference to the best explanation (IBE) (Calzavarini & Cevolani, 2022). Evaluating hypotheses is more than mere description or further elaboration; researchers must consider their thinking in relation to the raw data, previous analyses, and concepts (Wiltshire & Ronkainen, 2021). Nothing is sacrosanct at this point, as all themes are tentative in selecting the IBE. Evaluative methods can only work if researchers construct a range of hypotheses in Stage 2 that rival each other, such as opposites (on a gradient) or orthogonal hypotheses. To conduct evaluations in Abductive II, researchers must wage a spirited debate between themselves, the data, and their proposals to foster disagreements and questions and arrive at an informal range from possible to probable that the end result seems most likely (Calzavarini & Cevolani, 2022; Carter & Carter, 2014; Khemlani & Johnson-Laird, 2013). Disconfirmation, by tearing down hypotheses and building new ones, provides great insight and analytical power to demonstrate thinking and decision-making processes (Spiggle, 1994). Rejection of hypotheses, in part or whole, will show readers the contrastive reasons why a final theme was chosen. Tesch (1987) suggested that researchers embrace their anxiety and ambiguities to plunge into the great unknown.

Evaluative criteria, shown in Table 2, provide concrete strategies by starting with refuting hypotheses and then reworking them to select the best possible explanation. Miles et al. (2014) term the evaluative stage "taking a skeptical, demanding" (p. 258) position on the hypotheses generated in Stage 2, using explicit tactics. Part of the evaluation is that existing hypotheses will be continuously scrutinized, reconceptualized into new hypotheses, and constantly compared/contrasted with all data collected or generated (Tesch, 1987). Some common problems are fitting one's theory, positionality, personal motivation, perspective (going native), or research questions to the data rather than letting the data drive the results (Braun & Clarke, 2019; Kruglanski & Ajzen, 1983; Mahoney & Goertz, 2006; Miles et al., 2014). Many tactics are families, such as equifinality, multifinality, and counterfactuals, all of which are methods for negative case analysis.

Table 2
Evaluative Criteria for Thematic Hypotheses in Qualitative Research

Criterion	Guiding Question	Application
Foundational Analysis		
Relationships	How do variables link to each other?	Apply established typologies (e.g., Spradley's) or use analogies to map connections.
Context, Causes, Consequences	What is the holistic view?	Develop a situationally connected schema that outlines the A-P-C (Antecedent-Phenomenon-Consequence) model.
Dimensionality	What is the shape and strength of the concepts?	Describe the frequency (how often), intensity (how strong), and duration (how long) of variables.
Direction	What is the increase or decrease of concepts over time?	Describe the trajectory of change across time, space, and social interactions.
Temporal	What is the time frame and sequence of events?	Describe the timing of variables and their importance before, during, and after the phenomenon.
Narrative	What is the micro and macro story being told?	Contextualize variables within a coherent plot, including setting, characters, and events.
Causal & Relational Complexity		
Moderators	Does a third factor change the relationship between two variables?	Identify factors that act as catalysts, strengthening or weakening the link between other variables.
Mediators	What intervening factor is necessary for a cause to create an effect?	Identify the mechanism through which one variable influences another.
Proxies	What factors are absent but might be noteworthy?	Describe concepts that may be influential but were not directly observed or measured.
Reverse Causation	What if the cause and effect were reversed or are reciprocal?	Challenge the assumed direction of causality to see if alternative explanations are plausible.
Case Logic & Negative Case Analysis		
Multifinality	How do similar starting conditions lead to different results?	Examine outcomes to find differences that challenge a single hypothesized process.

Table 2
Contunied

Equifinality	How do different paths lead to the same result?	Consider how different variables or processes can create the same outcome.
Counterfactuals	What negative cases do not support the hypothesis?	Actively search for data that contradicts or offers an alternative to the primary conclusion.
Alternatives	What if my own views or the existing literature are wrong?	Deliberately consider opposite or orthogonal hypotheses to challenge the primary interpretation.
Researcher Positionality & Bias		
Transference	How do my beliefs, values, and attitudes translate into the results?	Articulate personal positionality and compare it to the final themes to check for bias.
Researcher	What are my aims and purposes with this research?	Re-integrate the findings from the specific perspective of the original research questions.
Isolation	Would these results make sense if disconnected from the original context?	Test variables out of context as a thought experiment to check their robustness and plausibility.
Data Sufficiency & Nuance		
Sample Individuation	What are the variations among individuals?	Describe the unique nuances of specific participants that exist outside the main group trends.
Multimeanings	Are there different meanings beyond the aggregate theme?	Describe multivocality by highlighting individual or subgroup differences in interpretation.
Distances	How well do the findings connect with both the core data and the outliers?	Describe the fit (and misfit) of the final themes with both commonalities and oddities in the data.
Cognitive-Behavioral-Emotional	What are the thoughts, behaviors, and emotions connected to a variable?	Connect the event to the whole person or group by analyzing the interplay of these three domains.
Ambiguities	What remains unclear, weak, or underdeveloped in the analysis?	Honestly describe the shortcomings, remaining questions, and areas needing further research.

Three important global strategies should be incorporated into evaluative abduction. First, researchers should test hypotheses using the necessary and sufficient criteria: Necessary means that some condition must be present for the results, outcome, or phenomenon, while sufficient means that a combination of conditions can also create the same or similar scenario (Mahoney & Goertz, 2006). Constraints are contexts, variables, or concepts that, when present, change processes and outcomes by establishing contradictions and partial fits that should be examined for constraint mechanisms and processes; constraints can be granular and specific or broad social systems (Ross, 2024). What makes searching for necessary, sufficient, and constraints necessary is that concepts and variables rarely vary monotonically across contexts, which is a sophisticated way of saying that one variable might increase in one context and decrease in the next (Mahoney & Goertz, 2006). Second, researchers need to examine hypotheses at the idiographic, subgroup, and group level by checking fit, relevance, and plausibility (Luborsky, 1994). This strategy means recursively going back and forth between the original data and all the analytical steps. Third, evaluations pit hypotheses against the data, concepts, and descriptive inferences, but researchers must place the hypotheses in the context of broader conditions and variations (Spiggle, 1994). That is, themes will not be universally applicable, will not make sense out of context, and will contain some contradictions at the idiographic or subgroup level.

The final product should then be evaluated. The following criteria could be useful, but there is no exhaustive, one-size-fits-all scheme: A.) Parsimony, or no more complex than necessary; B.) Singularity, or each theme has boundaries that separate it clearly from other themes; C.) Conditional, as each theme is wed to a particular context under specific conditions; D.) Variation, as themes work much like averages, with range as not all participants fitting exactly into any scheme and multiple meanings across subsets to be expected; E.) Rules of thumb are helpful but hold mainly for homogeneous samples, as the more heterogeneous, the more variation to be expected; F.) Contradictions should be common, as subsets and individuals can and will operate in contradiction, either in part or as a totality; G.) Missing data should be common, as some categories, concepts, and themes might be underdeveloped, especially in exploratory research with complex problems or where problems and possibilities were not forecasted but were developed in the research; H.) Researchers, if truly reflexive, should clearly document and discuss how their internal dialogue developed and might differ from that of participants (Elliott & Timulak, 2005; Mahoney & Goertz, 2006; Morgan, 2018; Ryan & Bernard, 2003; Sandelowski, 1995; Spiggle, 1994).

Conclusion

Researchers need to define the meaning and level of theme as well as the procedures and evaluations to construct, as there is no universal understanding (Mishra & Dey, 2022). Calls for rigor and explicit methods for generating themes, rather than searching for themes as if they were waiting to be found (cf. Morriss, 2024), collide with the concomitant lack of specific, concrete methods beyond counting and sorting. In a study of 4-year-olds, researchers found that children struggled to identify themes but were competent at describing their elements (Lehr, 1990); researchers often seem to be similarly lost. Themes are not just objects and actions out there, waiting to be harvested and collected by would-be researchers. Processes are not a leap of faith; fuzzy methods often mean that researchers lack the mental models needed to adopt and adapt in their own thinking.

Mysticism often appears in the methodology and methods sections of dissertations and articles, with unclear processes that move from codes to themes, with little or no specificity (Coker, 2021). Though reflexive thematic analysis has gone in a much different direction from the original prescription, we think all research should adhere to their earlier advice: "As thematic analysis is a flexible method, you also need to be clear and explicit about what you are doing, and what you say you are doing needs to match up with what you actually do" (Braun & Clarke, 2006, p. 96). Researchers struggle to develop interpretive themes (Braun & Clarke, 2022), but many prescriptions offer little new direction—the point of this article. Improving qualitative research starts with reading and studying a wide range of qualitative methods and then applying a systematic approach to all stages of qualitative research. There is no standardized method.

Despite claims to improve qualitative researchers, many still succumb to positivists views such as "a focused reduction of individual bias and enhancement of interrater reliability through the authors' use of multiple analysts in the coding process to ensure that emerging themes were robust and sufficiently reflective of the data," with the practices the antithesis of researcher instrumentality as ascribed by Braun and Clarke (Watson & Jackson, 2026, 2.8, para 3). Without robust memoing and documentation, researchers will struggle to develop an analytical position and lack transparency. Auditing should be procedures designed before a study commences and adapted as problems and new directions arise, as described in Appendix B. While sounding like a grounded theorist, the admonition seems applicable: A theme should have good fit and grab. Hypothesizing in Abductive I serves as a powerful means to describe the decisions made, what was rejected, and which competing claims remain, as well as to provide insight into a researcher's personal logic in Abductive II. As many researchers have pointed out, analysis could continue ad infinitum, so intentional choices must be made to close research when derived conclusions seem to move beyond adequate (a minimalist perspective) to superb. There is no magic stopping point, so researchers will have to document why they chose what they did.

Call to Action

- **Adopt systematic methods** to make the leap from codes to themes a transparent and rigorous process, moving beyond the vague claim that "themes emerge."
- **Recognize the taxonomy of themes** by consciously working to move analysis beyond simple topical and descriptive summaries toward more powerful, multi-faceted inferential interpretations.
- **Implement the three-stage storyboard process**—Conceptual Clarity, Hypothesizing Possibilities, and Interpretive Decision Making—to provide a structured yet flexible framework for theme construction.
- **Move beyond simple categorization** by using structured tools like concept boxing to build abstract concepts that explain the relationships, context, and mechanisms within the data before forming themes.
- **Employ abductive reasoning** to generate multiple, competing hypotheses (candidate themes) that explore the potential "why" and "how" behind the data, reading beyond the lines to create new interpretations.
- **Engage in rigorous self-critique** by systematically evaluating candidate themes against specific criteria, actively seeking disconfirmation, and considering negative cases to select the most plausible explanation.
- **Embrace methodological flexibility** by adapting and combining tools from different qualitative traditions, focusing on the research purpose rather than rigidly adhering to a single named method.
- **Prioritize transparency** by maintaining a detailed audit trail and using robust memoing to document the entire analytical journey, including false starts, rejected hypotheses, and the final decision-making process.

Footnote

1.The elimination of counting codes is not advocated; however, researchers must be aware of both absolute counts (the total number of codes) and relative counts (the number of codes per participant). Published results often lack details if a code counting is absolute, which could cause the following problem, as demonstrated by this fictitious example: If a code exists 25 times for a sample of 15 (an absolute count), what was the distribution to the participants; what if the code derived from three participants for 24 times and 1 code was from another participant. Notice that code counting does not specify the context or the dimensionality. Great care must be taken to qualify counting beyond a number and a percentage. See Christou (2022) for a more extensive discussion.

Conflicts of Interest

There are no known conflicts of interest.

Funding

There was no funding for the study.

Acknowledgment

Gratitude to the graduate students who use the methods described herein. The reviewers' thoughtful feedback greatly strengthened the manuscript.

Reference

Ahmed, S. K., Mohammed, R. A., Nashwan, A. J., Ibrahim, R. H., Abdalla, A. Q., Ameen, B. M. M., & Khdhir, R. M. (2025). Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health, 6*, 100198. https://doi.org/10.1016/j.glmedi.2025.100198

Becker, H., & Geer, B. (1957). Participant observation and interviewing: A comparison. *Human Organization*, *16*(3), 28-32. https://doi.org/10.17730/humo.16.3.k687822132323013

Belmonte, I. A., & McCabe-Hidalgo, A. (1998). Theme-rheme patterns in L2 writing. *Didáctica, 10*(1), 13-31.

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

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage.

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage.

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology, 9*(1), 3–26. https://doi.org/10.1037/qup0000196

Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be (com)ing a knowing researcher. *International Journal of Transgender Health, 24*(1), 1-6. https://doi.org/10.1080/26895269.2022.2129597

Braun, V., & Clarke, V. (2024a). A critical review of the reporting of reflexive thematic analysis in Health Promotion International. *Health Promotion International, 39*(3), daae049. https://doi.org/10.1093/heapro/daae049

Braun, V., & Clarke, V. (2024b). Supporting best practice in reflexive thematic analysis reporting in Palliative Medicine: A review of published research and introduction to the Reflexive Thematic Analysis Reporting Guidelines (RTARG). *Palliative Medicine, 38*(6), 608-616. https://doi.org/10.1177/02692163241234800

Braun, V., & Wilkinson, S. (2003). Liability or asset? Women talk about the vagina. *Psychology of Women Section Review, 5*(2), 28-42. https://doi.org/10.53841/bpspow.2003.5.2.28

Buetow, S. (2010). Thematic analysis and its reconceptualization as 'saliency analysis'. *Journal of Health Services Research & Policy, 15*(2), 123-125. https://doi.org/10.1258/jhsrp.2009.009081

Burnard, K. J. (2024). Developing a robust case study protocol. *Management Research Review*, *47*(2), 204-225. https://doi.org/10.1108/MRR-11-2021-0821

Calzavarini, F., & Cevolani, G. (2022). Abductive reasoning in cognitive neuroscience: Weak and strong reverse inference. *Synthese, 200*(2), 70. https://doi.org/10.1007/s11229-022-03585-2

Carter, M. J., & Carter, S. (2014). How themes in literature can inform sociological theory, research, and teaching. *The American Sociologist, 45*, 388-411. https://doi.org/10.1007/s12108-014-9221-1

Christou, P. A. (2022). How to use thematic analysis in qualitative research. *Journal of Qualitative Research in Tourism, 3*(2), 79–95. https://doi.org/10.4337/jqrt.2023.0006

Coker, D. C., & Akande, A. (2025). Preparation, development, and refinement of the interview protocol in qualitative research. *The Qualitative Report, 30*(8), 4205-4235. https://doi.org/10.46743/2160-3715/2025.7964

Coker, D. C. (2021). Making thematic analysis systematic: The seven deadly sins. *Journal of Studies in Education, 11*(3), 126-146. https://doi.org/10.5296/jse.v11i3.18882

Coker, D. C. (2022). A thematic analysis of the structure of delimitations in the dissertation. *International Journal of Doctoral Studies, 17*, 141-159. https://doi.org/10.28945/4939

Coker, D. C. (2023). An integrative qualitative framework: Improving research through strategic mapping. *International Research in Education, 11*(1), 66-109. https://doi.org/10.5296/ire.v11i1.20921

Coker, D. C., Kaaki, W., & Miao, X. D. (2025). Qualitative research in leadership and organizational studies. In A. Akande (Ed.), *Organizational behavior: Current science, models, and applications*. Springer.

Dawadi, S. (2020). Thematic analysis approach: A step by step guide for ELT research practitioners. *Journal of NELTA, 25*(1-2), 62-71. https://doi.org/10.3126/nelta.v25i1-2.49731

de Farias, B. G., Dutra-Thomé, L., Koller, S. H., & de Castro, T. G. (2021). Formulation of themes in qualitative research: Logical procedures and analytical paths. *Trends in Psychology, 29*(1), 155-166. https://doi.org/10.1007/s43076-020-00052-0

DeSantis, L., & Ugarriza, D. N. (2000). The concept of theme as used in qualitative nursing research. *Western Journal of Nursing Research, 22*(3), 351-372. https://doi.org/10.1177/019394590002200308

Elliott, R., & Timulak, L. (2005). Descriptive and interpretive approaches to qualitative research. In J. Miles & P. Gilbert (Eds.), *A handbook of research methods for clinical and health psychology* (pp. 147-159). Wiley.

Finlay, L. (2021). Thematic analysis: The 'good', the 'bad', and the 'ugly'. *European Journal for Qualitative Research in Psychotherapy, 11*, 103–116. https://doi.org/10.24377/EJQRP.article3062

Fischer, H. R. (2001). Abductive reasoning as a way of worldmaking. *Foundations of Science, 6*, 361-383. https://doi.org/10.1023/A:1011671106610

Francis, G. (1989). Thematic selection and distribution in written discourse. *Word, 40*(1-2), 201-221. https://doi.org/10.1080/00437956.1989.11435804

Fryer, T. (2022). A critical realist approach to thematic analysis: producing causal explanations. *Journal of Critical Realism, 21*(4), 365-384. https://doi.org/10.1080/14767430.2022.2076776

Glaser, B. G. (2008). The constant comparative method of qualitative analysis. *Grounded Theory Review, 7*(3), 3-12. https://groundedtheoryreview.org/index.php/gtr/article/view/467

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report*, *8*(4), 597-607. https://doi.org/10.46743/2160-3715/2003.1870

Guest, G., MacQueen, K. M., & Namey, E. E. (2011). *Applied thematic analysis*. Sage Publications.

Hennink, 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

Herzog, C., Handke, C., & Hitters, E. (2019). Analyzing talk and text II: Thematic analysis. In H. Van den Bulck, M. Puppis, K. Donders, & L. Van Audenhove (Eds.), *The Palgrave handbook of methods for media policy research* (pp. 395–410). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-16065-4_22

Holton, J. A. (2008). Grounded theory as a general research methodology. *Grounded Theory Review*, *7*(02), 27-45. https://groundedtheoryreview.org/index.php/gtr/article/view/429

Jules, T. D., & Salajan, F. D. (2024). Toward the (post) qualitative turn in comparative and international education. *Comparative Education Review*, *68*(2), 199-211. https://www.journals.uchicago.edu/doi/full/10.1086/730517

Karlsen, B., Hillestad, T. M., & Dysvik, E. (2021). Abductive reasoning in nursing: Challenges and possibilities. *Nursing Inquiry, 28*(1), e12374. https://doi.org/10.1111/nin.12374

Kassarjian, H. H. (1977). Content analysis in consumer research. *Journal of Consumer Research, 4*(1), 8-18. https://doi.org/10.1086/208674

Khemlani, S., & Johnson-Laird, P. N. (2013). The processes of inference. *Argument & Computation, 4*(1), 4-20. https://doi.org/10.1080/19462166.2012.674060

Khurshid, F., van der Veen, M., Thompson, J., & Hegazi, I. (2024). Navigating thematic analysis: Practical strategies grounded in abductive reasoning. *Teaching and Learning in Medicine*. Advance online publication. https://doi.org/10.1080/10401334.2024.2349887

Kidd, I. J. (2016). Inevitability, contingency, and epistemic humility. *Studies in History and Philosophy of Science Part A*, *55*, 12-19. https://doi.org/10.1016/j.shpsa.2015.08.006

Kim, J. H. (2015). *Understanding narrative inquiry: The crafting and analysis of stories as research*. Sage Publications.

Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology, 13*(1), 1-44. https://doi.org/10.1002/ejsp.2420130102

Kushnir, I. (2025). Thematic analysis in the area of education: A practical guide. *Cogent Education, 12*(1), 2471645. https://doi.org/10.1080/2331186X.2025.2471645

Lehr, S. (1990). Literature and the construction of meaning: The preschool child's developing sense of theme. *Journal of Research in Childhood Education, 5*(1), 37-46. https://doi.org/10.1080/02568549009594801

Liamputtong, P. (2009). Qualitative data analysis: Conceptual and practical considerations. *Health Promotion Journal of Australia, 20*(2), 133-139. https://doi.org/10.1071/HE09133

Luborsky, M. R. (1994). The identification and analysis of themes and patterns. In J. F. Gubrium & A. Sankar (Eds.), *Qualitative methods in aging research* (pp. 189–210). Sage Publications, Inc.

Lucas, K., & D'Enbeau, S. (2013). Moving beyond themes: Reimagining the qualitative analysis curriculum. *Qualitative Communication Research, 2*(2), 213-227. https://doi.org/10.1525/qcr.2013.2.2.213

Lyhne, C. N., Thisted, J., & Bjerrum, M. (2025). Qualitative content analysis–framing the analytical process of inductive content analysis to develop a sound study design. *Quality & Quantity*, 1-21. https://doi.org/10.1007/s11135-025-02220-9

Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis, 14*(3), 227-249. https://doi.org/10.1093/pan/mpj017

Majumdar, A. (2019). Thematic analysis in qualitative research. In *Qualitative techniques for workplace data analysis* (pp. 197-220). IGI Global Scientific Publishing.

Maxwell, J. A. (2021). The importance of qualitative research for investigating causation. *Qualitative Psychology, 8*(3), 378–388. https://doi.org/10.1037/qup0000219

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A sourcebook of new methods*. Sage.

Mishra, S., & Dey, A. K. (2022). Understanding and identifying 'themes' in qualitative case study research. *South Asian Journal of Business and Management Cases, 11*(3), 187-192. https://doi.org/10.1177/22779779221134659

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry*, *20*(8), 1045-1053. https://doi.org/10.1177/1077800413513733

Morgan, D. L. (2018). Themes, theories, and models. *Qualitative Health Research, 28*(3), 339-345. https://doi.org/10.1177/10497323177501

Morriss, L. (2024). Themes do not emerge. An editor's reflections on the use of Braun and Clarke's thematic analysis. *Qualitative Social Work, 23*(5), 745-749. https://doi.org/10.1177/14733250241277355

Mukumbang, F. C. (2023). Retroductive theorizing: A contribution of critical realism to mixed methods research. *Journal of Mixed Methods Research, 17*(1), 93-114. https://doi.org/10.1177/15586898211049847

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A step-by-step process of thematic analysis to develop a conceptual model in qualitative research. *International Journal of Qualitative Methods, 22*, 16094069231205789. https://doi.org/10.1177/16094069231205789

Onwuegbuzie, A. J., Collins, K. M., & Frels, R. K. (2013). Foreword: Using Bronfenbrenner's ecological systems theory to frame quantitative, qualitative, and mixed research. *International Journal of Multiple Research Approaches*, *7*(1), 2-8. https://doi.org/10.5172/mra.2013.7.1.2

Özden, M. (2024). Content and thematic analysis techniques in qualitative research: Purpose, process and features. *Qualitative Inquiry in Education: Theory & Practice, 2*(1), 64-81. https://doi.org/10.59455/qietp.20

Paul, G. (1993). Approaches to abductive reasoning: An overview. *Artificial Intelligence Review, 7*(2), 109-152. https://doi.org/10.1007/BF00849080

Peel, K. L. (2020). A beginner's guide to applied educational research using thematic analysis. *Practical Assessment Research and Evaluation, 25*(1). https://doi.org/10.7275/ryr5-k983

Polkinghorne, D. E. (1995). Narrative configuration in qualitative analysis. *International Journal of Qualitative Studies in Education, 8*(1), 5-23. https://doi.org/10.1080/0951839950080103

Reichertz, J. (2010). Abduction: The logic of discovery of grounded theory. *Forum: Qualitative Social Research, 11*(1). https://doi.org/10.17169/fqs-11.1.1412

Ross, L. N. (2024). Causal constraints in the life and social sciences. *Philosophy of Science, 91*(5), 1068-1077. https://doi.org/10.1017/psa.2023.165

Russell, B. (1935, January). The limits of empiricism. In *Proceedings of the Aristotelian Society* (Vol. 36, pp. 131-150). Aristotelian Society, Wiley.

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. *Field Methods, 15*(1), 85-109. https://doi.org/10.1177/1525822X02239569

Sagan, C. (2007). The fine art of baloney detection. In B. Farha (Ed.), *Paranormal claims: A critical analysis* (pp. 1-12). Bloomsbury Publishing.

Saldaña, J. (2015). *The coding manual for qualitative researchers*. Sage.

Sandelowski, M. (1995). Qualitative analysis: What it is and how to begin. *Research in Nursing & Health, 18*(4), 371-375. https://doi.org/10.1002/nur.4770180411

Spiggle, S. (1994). Analysis and interpretation of qualitative data in consumer research. *Journal of Consumer Research, 21*(3), 491-503. https://doi.org/10.1086/209413

Swain, J. (2018). *A hybrid approach to thematic analysis in qualitative research*. SAGE Publications, Ltd. https://doi.org/10.4135/9781526435477

Tavory, I. (2020). Interviews and inference: Making sense of interview data in qualitative research. *Qualitative Sociology, 43*(4), 449-465. https://doi.org/10.1007/s11133-020-09464-x

Terry, G., Hayfield, N., Clarke, V., & Braun, V. (2017). Thematic analysis. In *The SAGE handbook of qualitative research in psychology* (2nd ed., pp. 17-37). Sage.

Tesch, R. (1987). Emerging themes: The researcher’s experience. *Phenomenology + Pedagogy*, 230-241. https://doi.org/10.29173/pandp15058

Thagard, P., & Shelley, C. (1997). Abductive reasoning: Logic, visual thinking, and coherence. In *Logic and scientific methods: Volume one of the Tenth International Congress of Logic, Methodology and Philosophy of Science, Florence, August 1995* (pp. 413-427). Springer.

Thompson, J. (2022). A guide to abductive thematic analysis. *The Qualitative Report*, *27*(5), 1410-1421. https://doi.org/10.46743/2160-3715/2022.5340

Tracy, S. J. (2010). Qualitative quality: Eight “big-tent” criteria for excellent qualitative research. *Qualitative Inquiry*, *16*(10), 837-851. https://doi.org/10.1177/1077800410383121

Vaismoradi, M., & Snelgrove, S. (2019). Theme in qualitative content analysis and thematic analysis. *Forum: Qualitative Social Research, 20*(3). https://doi.org/10.17169/fqs-20.3.3376

Van Meegdenburg, H. (2023). Process tracing: An analyticist approach. In P. A. Mello & F. Ostermann (Eds.), *Routledge handbook of foreign policy analysis methods* (pp. 405-420). Taylor & Francis.

Villamin, P., Lopez, V., Thapa, D. K., & Cleary, M. (2025). A worked example of qualitative descriptive design: A step-by-step guide for novice and early career researchers. *Journal of Advanced Nursing, 81*(8), 5181-5195. https://doi.org/10.1111/jan.16481

Vollstedt, M., & Rezat, S. (2019). An introduction to grounded theory with a special focus on axial coding and the coding paradigm. In G. Kaiser & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education* (pp. 81-100). Springer.

Watson, A., & Jackson, D. (2026). The art of analysis: Crafting credible and authentic qualitative results. *Journal of Advanced Nursing*, *82*(2), 1716-1724. https://doi.org/10.1111/jan.17032

Wieler, L. A. (2025). Qualitative data analysis framework for early career researchers: ARISE. *American Journal of Qualitative Research*, *9*(3), 101-107. https://doi.org/10.29333/ajqr/16529

Willig, C. (2012). *Qualitative interpretation and analysis in psychology*. McGraw-Hill Education (UK).

Wiltshire, G., & Ronkainen, N. (2021). A realist approach to thematic analysis: Making sense of qualitative data through experiential, inferential and dispositional themes. *Journal of Critical Realism, 20*(2), 159-180. https://doi.org/10.1080/14767430.2021.1894909

Wynn, G. (2024). Reformative concept analysis for applied psychology qualitative research. *Qualitative Research in Psychology, 21*(3), 251-278. https://doi.org/10.1080/14780887.2024.2346851

Xu, W. (2026). Doing thematic analysis in the age of generative AI: Practices, ethics and reflexivity. *International Journal of Qualitative Methods*, *25*, 16094069261425173. https://doi.org/10.1177/16094069261425173

Notes on Contributor

David C. Coker, Ed.D., is an Associate Professor and Doctoral Methodologist at West Liberty University. He holds a doctorate in Educational Leadership and is a specialist in qualitative, quantitative, and mixed-methods research design. Dr. Coker researches, publishes, and presents frequently, with over 60 publications. His work is multidisciplinary, focusing on education, leadership/organizational behavior, research methods, and the integration of artificial intelligence in scholarly research. He bridges academic theory with practical application, emphasizing methodological integrity and the development of audit tools to foster transparency and empower practitioners.

ORCID

David C. Coker https://orcid.org/0000-0002-9138-2125