



Virginia Braun, Victoria Clarke, Nikki Hayfield, and Gareth Terry

Contents

1	Introduction	844
2	Thematic Analysis: A Brief History	844
3	Mapping the Terrain of Thematic Analysis: What Is a Theme?	845
4	Mapping the Terrain of Thematic Analysis: Different Schools of TA	847
5	Some Design Considerations for (Reflexive) Thematic Analysis	849
6	Six Phases of Reflexive Thematic Analysis	852
7	Conclusion and Future Directions	857
	References	858

Abstract

This chapter maps the terrain of thematic analysis (TA), a method for capturing patterns (“themes”) across qualitative datasets. We identify key concepts and different orientations and practices, illustrating why TA is often better understood as an umbrella term, used for sometimes quite different approaches, than a single qualitative analytic approach. Under the umbrella, three broad approaches can be identified: a “coding reliability” approach, a “codebook” approach, and a “reflexive” approach. These are often characterized by distinctive – sometimes radically

V. Braun (✉)

School of Psychology, The University of Auckland, Auckland, New Zealand
e-mail: v.braun@auckland.ac.nz

V. Clarke · N. Hayfield

Department of Health and Social Sciences, Faculty of Health and Applied Sciences, University of the West of England (UWE), Bristol, UK
e-mail: victoria.clarke@uwe.ac.uk; nikki2.hayfield@uwe.ac.uk

G. Terry

Centre for Person Centred Research, School of Clinical Sciences, Auckland University of Technology, Auckland, New Zealand
e-mail: gareth.terry@aut.ac.nz

different – conceptualizations of what a theme is, as well as methods for theme identification and development, and indeed coding. We then provide *practical* guidance on completing TA within our popular (reflexive) approach to TA, discussing each phase of the six-phase approach we have developed in relation to a project on men, rehabilitation, and embodiment. We conclude with a discussion of key concerns related to ensuring the TA you do – within whatever approach – is of the highest quality.

Keywords

Code · Codebook · Coding reliability · Epistemology · Latent · Reflexive thematic analysis · Semantic · Thematic map · Theme

1 Introduction

Thematic analysis (TA) is often misconceptualized as a single qualitative analytic approach. It is better understood as an umbrella term, designating sometimes quite different approaches aimed at identifying patterns (“themes”) across qualitative datasets. In this chapter, we first define key concepts and map the terrain of TA; we identify three distinct “schools” of TA, highlighting differences between these schools, particularly in relation to underlying philosophy and approach to data analysis. We then provide practical guidance on completing TA, focused on one of the most popular approaches – developed by two of the authors of this chapter (Braun and Clarke 2006, 2012, 2013).

2 Thematic Analysis: A Brief History

Philosopher of science Gerard Houlton has been credited with inventing TA in the 1970s, in his work on “themata” in scientific thought (Holton 1973; see also Joffe 2011). But the term was in use well before then: musicologists in the 1930s described the analysis of musical scores as TA (e.g., Kinsky and Strunk 1933); sociologists in the 1940s used the term to describe a method for analyzing mass propaganda (e.g., Lazarsfeld and Merton 1944); psychoanalysts in the 1950s used it to refer to techniques for analyzing the results of projective tests (e.g., Winder and Hersko 1958). The conceptualization of TA as an approach for analyzing patterns of meaning *may* reflect a methodological evolution from (quantitative) content analysis. The terms “TA” and “content analysis” have often been used interchangeably, and the hybrid term “thematic content analysis” is also common (e.g., Brewster et al. 2014). Regardless of its developmental origins, TA clearly has shared history with content analysis (see ► Chap. 47, “Content Analysis: Using Critical Realism to Extend Its Utility”).

In the 1980s and 1990s – around the time there was a general explosion of interest in qualitative research – TA started to appear as a particular approach for analyzing *qualitative data* in the health and social sciences (e.g., Dapkus 1985; Aronson 1994). But it was the countless published papers that described some version of “themes emerging” from data, without reference to an established methodological approach to TA, that led us to describe TA as “a poorly demarcated and rarely acknowledged, yet widely used qualitative analytic method” (Braun and Clarke 2006, p. 77). That was 2006; just over a decade later, how things have changed! TA is now increasingly recognized as an approach to analysis in its own right (there is still some debate around this; see (Willig 2013)), and there are *many* different approaches to TA. The shared name “TA” obscures divergence, both in terms of procedures, and, more importantly, in underlying philosophy and the conceptualization of key elements of the method (e.g., a theme, a code, or coding). It is not uncommon to see researchers cite sources on, and sometimes follow procedures for, approaches to TA that do not align, conceptually or in practice. Not grasping these distinctions can result in published papers where the approach to TA used is unclear, procedures and assumptions are misattributed or mixed up, and underlying conceptual clashes between different approaches are not recognized. This does a disservice to TA. Avoiding such errors requires understanding of the conceptual and procedural differences within the terrain of TA. To aid clarity, we will define some key concepts in TA and consider the distinctive features of three “schools” of TA – which we refer to as “coding reliability,” “codebook,” and “reflexive TA.”

3 Mapping the Terrain of Thematic Analysis: What Is a Theme?

First up, it is vital to understand how “a theme” is conceptualized, as there are two competing ideas in TA research: domain summaries versus shared meaning-based patterns. We – and many others – view themes as reflecting a *pattern* of shared meaning, organized around a core concept or idea, a central organizing concept (see Braun et al. 2014). In this conceptualization, themes capture the essence and spread of meaning; they unite data that might otherwise appear disparate, or meaning that occurs in multiple and varied contexts; they (often) explain large portions of a dataset; they are often abstract entities or ideas, capturing implicit ideas “beneath the surface” of the data, but can also capture more explicit and concrete meaning; and they are built from smaller meaning units (codes) (DeSantis and Ugarriza 2000). An example of this type of theme comes from our research on meaning around male body hair (Terry and Braun 2016). A theme “men’s hair as natural” captured the way body hair was often described as natural for men and “a dominant expression of masculine embodiment” (p. 17). As well as reporting participants’ overt statements about the naturalness of male body hair (a “surface” level of meaning) – the theme

explored more nuanced manifestations of this idea – gendered assumptions that men *should* be hairy and women hairless and that men’s embodiment is biological (natural) while women’s is socially produced (worked upon), constructions which were naturalized and essentialized in the dataset.

In contrast to our conceptualization, a theme in a “domain summary” conceptualization *summarizes* what participants said in relation to a topic or issue, typically at the *semantic* or surface level of meaning, and usually reports multiple or even contradictory meaning-content. The (“theme”) issues are often based around data collection tools, such as responses to a particular interview question. Take the first theme in Roditis and Felsher’s (2015) research on adolescents’ perceptions of the risks and benefits of conventional cigarettes, e-cigarettes, and marijuana. The title – “perceived risks and benefits of conventional cigarettes compared to marijuana” – indicates a theme-as-domain-summary conceptualization, not least because it combines risks *and* benefits. And indeed, that is what is reported. Overview-type statements – “Youth either stated there was nothing good about using conventional cigarettes or stated that using cigarettes could help someone relax. Students easily recited a long list of negative consequences related to conventional cigarette use such as . . .” (p. 182) – highlight that the theme is a summary of youth perceptions in relation to a particular topic area.

Although some see domain summaries as a meaningful and useful conceptualization of a theme, others (e.g., Sandelowski and Leeman 2012; Connelly and Peltzer 2016) characterize them as at best underdeveloped or not fully realized themes and, at worst, misconceptualized. Some TA reports *do* read as if the analysis is only partly developed. For example, in Weatherhead and Daiches’s (2010) paper on Muslim views on mental health and psychotherapy, the seven themes – “causes,” “problem management,” “relevance of services,” “barriers,” “service delivery,” “therapy content,” and “therapist characteristics” – were effectively domain summaries. Discussing this paper in class, one of our students evocatively dubbed them “bucket themes”: you collect all the information gathered about X in one place, without considering shared meaning or difference. The theme “causes,” for example, described participants’ attributions for their mental health problems – explanations as diverse as reactions to life events and secular or religious notions that “life is a test.” Yet Weatherhead and Daiches’ discussion explored the “continued interweaving of religious and secular influences in participants’ account of mental distress and well-being” (p. 85), hinting at the potential for themes as shared-meanings, where the analysis is developed further and deeper. A domain-summary approach risks conceptualizing TA as simply a data reduction activity, where the purpose of analysis is to succinctly summarize the diversity of responses across the scope of a project. However, this *can* sometimes simply be an issue of ensuring shared-meaning themes are well-named (see Braun and Clarke 2013); we recommend avoiding one-word theme names to avoid this.

Approaches to TA also vary on whether themes are conceptualized as analytic *inputs* – patterns identified and developed at the *start* of the analytic process (usually following some data familiarization) which guide the data coding process – or as analytic *outputs*, patterns identified and developed *later* in the analytic process,

building on, and representing the *outcome* of, coding. To some extent, these conceptualizations align with the two different ideas about what a theme is.

4 Mapping the Terrain of Thematic Analysis: Different Schools of TA

We refer to *schools* of TA when we describe three broad “types” of TA, because there is not just *one* approach associated with each type. The names we use for the schools – coding reliability, codebook, and reflexive – emphasize the key distinctive element of each approach.

Coding reliability approaches – associated with authors like Boyatzis (1998), Guest et al. (2012), and Joffe (2011) – represent what we characterize as a *partially* qualitative approach to TA. Qualitative data are collected and analyzed using qualitative techniques of coding and theme development; the data are reported qualitatively as themes, typically illustrated by extracts of data. However, the underlying logic of these processes is firmly (post-)positivist, and some characterize these coding reliability approaches as “bridging the divide” between qualitative and quantitative methods. According to Boyatzis (1998, p. vii), (coding reliability) TA “is a translator of those speaking the language of qualitative analysis and those speaking the language of quantitative analysis.” Coding reliability TA researchers share values with quantitative researchers’ – for example, the importance of the reliability and replicability of observation – values at odds with (fully) qualitative paradigms.

In coding reliability TA, themes are often conceptualized as domain summaries (often derived from data collection questions), and as analytic inputs, as well as outputs – they *drive* the coding process and are the *output* of the coding process. The coding *process* is designed to prioritize “reliable” data coding, by which they mean identification of “accurate” codes/themes within data, usually based on agreement among multiple coders. Coding is guided by a codebook/coding frame, which typically contains a list of codes/themes – each has a label/name, a definition, information on how to identify the code/theme, a description of any exclusions or qualifications to identifying the code/theme, and data examples (Boyatzis 1998). This is designed to allow the researcher to categorize the data into (predetermined) themes. Despite a sometimes interchangeable use of the terms code and theme, *coding* is essentially conceptualized as a *process* (for identification of theme-relevant data and thus themes). Ideally, the codebook is applied to the data by more than one coder, each working independently; for some, the ideal coder has no prior experience with or knowledge of the topic of concern and comes to the coding process “cold.” After coding, the level of “agreement” between the coders is calculated (using Cohen’s kappa). A score of 0.80 and above is generally thought to signal accurate or reliable coding; lower scores are problematic, and lack of agreement needs to be resolved.

This coding approach can be understood as *consensus* coding – because it builds toward a singular, shared, and “correct” analysis of the data. The process has strong

echoes of “the scientific method” – the researcher develops a hypothesis (themes), tests these (searches for evidence of the themes using the codebook), is concerned about, and seeks to control for, “researcher bias” or “influence”, and, if the right procedures are followed, claims reliable and potentially replicable results. It reflects what Kidder and Fine (1987) described as “small q” qualitative research – qualitative research conceptualized as tools and techniques, not as a paradigm or underlying philosophy for research. We consequently characterize this school of TA as “partially” qualitative. To us, the idea that (such) TA can “bridge a qualitative-quantitative divide” is problematic, because it requires discarding what we see as central to good qualitative research practice – depth of engagement (“commitment and rigor” in (Yardley’s 2015) open-ended and flexible principles for qualitative research quality), an open and exploratory design and analytic process, and a prioritization of researcher subjectivity and reflexivity (Finlay and Gough 2003; Gough and Madill 2012). The *reflexive* school of TA emphasizes these elements.

Reflexive TA approaches include our (e.g., Braun and Clarke 2006) popular version of TA, as well as others (e.g., Langdridge 2004). In these, TA is conceptualized as a *fully* qualitative approach – with data collection and analysis techniques underpinned by a qualitative philosophy or paradigm – a “Big Q” approach (Kidder and Fine 1987). Although there is no widely agreed definition of a qualitative paradigm or, indeed, whether there is just one qualitative paradigm (Madill 2015), a qualitative orientation usually emphasizes meaning as contextual or situated, reality or realities as multiple, and researcher subjectivity as not just valid but a *resource* (Braun and Clarke 2013). We characterize this school as *reflexive* TA to emphasize the active role of the researcher in the knowledge production process.

In reflexive TA, themes are conceptualized as meaning-based patterns, evident in explicit (semantic) or conceptual (latent) ways, and as the *output* of coding – themes result *from* considerable analytic work on the part of the researcher to explore and develop an understanding of patterned meaning across the dataset. Coding is an organic and open iterative *process*; it is not “fixed” at the start of the process (e.g., through the use of a codebook or coding frame). Codes – the product of coding – can evolve throughout the coding process. An initial code might be “split” into two or more different codes, renamed, or combined with other codes. The aim of such changes during coding is to better capture the researcher’s developing conceptualization of the data. It is relatively easy to determine themes as domain summaries at the start of the analytic process; it is difficult to determine themes as conceptually founded patterns at the start, because it requires depth of (close and critical) engagement to move beyond the surface or obvious content of the data and to identify implicitly or unexpected unifying patterns of meaning. The aim of coding and theme development in reflexive TA is not to “accurately” summarize the data, nor to minimize the influence of researcher subjectivity on the analytic process, because neither is seen as possible nor indeed desirable. The aim is to provide a coherent and compelling *interpretation* of the data, grounded in the data. The researcher is a *storyteller*, actively engaged in interpreting data through the lens of their own cultural membership and social positionings, their theoretical assumptions

and ideological commitments, as well as their scholarly knowledge. This subjective, even political, take on research is very different to a positivist-empiricist model of the researcher. Many reflexive TA researchers do indeed have some kind of social justice motivation – be it “giving voice” to a socially marginalized group, or a group rarely allowed to speak or be heard in a particular context, or a more radical agenda of social critique or change.

We use the term *codebook TA* to describe a third “school” of TA – although many of these, which include framework (e.g., Ritchie and Spencer 1994; Ritchie and Lewis 2003; Smith and Firth 2011), template (King 2014; Brooks et al. 2015) and matrix analysis (Miles and Huberman 1994; Nadin and Cassell 2014), among others, do not use the actual term TA. This school of TA sits somewhere *between* “coding reliability” and “reflexive” TA, sharing the structured approach to coding with coding reliability TA (though often *without* the use of coding reliability measures) with the broadly qualitative underlying philosophy of reflexive TA. In codebook TA, some if not all themes are determined in advance of full analysis, and themes are typically conceptualized as domain summaries.

Some TA researchers, including template analysis proponents Brooks et al. (2015), have argued that researchers should not be precious about their way of working with TA. Although some friends and colleagues have (jokingly) suggested Virginia and Victoria are the “TA police,” issuing edicts about how TA should be done and rigorously punishing crimes against TA that do not follow our guidelines, we actually *somewhat* agree with Brooks et al.’s sentiment. Overall, what is important is that researchers use the approach to TA that is most appropriate for their research, they use it in a “knowing” way, they aim to be thoughtful in their data collection and analytic processes and practices, and they produce an overall coherent piece of work. Yet, we do *advocate* certain practices. From our perspective, the use of a structured codebook, determining themes in advance of analysis or following only data familiarization (using themes as analytic inputs) and conceptualizing themes as domain summaries, delimits the depth of engagement and flexibility central to *qualitative* research practice. There are, however, clear pragmatic advantages to codebook approaches – the coding framework allows teams of researchers to more easily work *together* on data analysis, facilitates a relatively quick analytic process, and provides some structure for qualitative novices. Taken out of a “consensus” and reliability framework, this has potential to produce rich nuanced analysis. But pragmatic factors should not (always) be the sole determinant of method.

5 Some Design Considerations for (Reflexive) Thematic Analysis

TA offers researchers great flexibility, meaning it can be used to do lots of the things that qualitative researchers are interested in. This flexibility stems from TA’s status as an analytic method, rather than a methodology, the latter referring to a theoretically-informed *framework* for research. Although the school of TA chosen

delimits a broad paradigm ([post]positivist or qualitative), beyond that there is scope to design and locate the method – and indeed, a requirement to do so, for reflexive TA.

Locating your overall theoretical and interpretative frameworks is important. Some treat TA as particularly compatible with phenomenological approaches (e.g., Joffe 2011; Guest et al. 2012), and it is indeed often used to describe or summarize participants' experiences, rather than to do more interpretative or conceptual work. But *why* TA should be limited to such an interpretative framework is unclear, and we think treating TA as a descriptive approach focused on experience underappreciates its flexibility and full potential. Indeed, it works well with many different interpretative frameworks, ranging from phenomenological ones to critical constructionist interrogations of meaning. And it, therefore, has the potential to answer different research questions. TA *can* address questions about, and be used to describe, the “lived experiences” of particular social groups (e.g., sex workers (Mellor and Lovell 2012), people with Parkinson's disease (Redmond et al. 2012), Asian migrants (Terry et al. 2011), adults with gay, lesbian, or bisexual parents (Titlestad and Pooley 2014)) or about particular aspects of their lives (e.g., the experience of freezing for people with Parkinson's disease (Redmond et al. 2012) or the health needs of street-based sex workers (Mellor and Lovell 2012)). It can also examine the “factors” that influence, underpin, or contextualize particular processes or phenomena (such as the factors that shape nurses' values in relation to compassionate care (McSherry et al. 2017)), identify views about particular phenomena (such as contested views about who is best placed to provide expertise in legal proceedings related to children's care (Hill et al. 2017)), or interrogate dominant patterns of meaning surrounding particular phenomena (such as the discourses underpinning the normalization of female genital cosmetic surgery on a cosmetic surgery website (Moran and Lee 2013)). Research questions for TA need to be aligned with the theoretical orientation of your TA.

TA also offers flexibility around data collection: interviews are common; focus groups are popular; diaries, visual methods, participatory methods, surveys, a wide range of secondary sources – such as online forums, blogs, websites, magazines, newspaper articles, and police reports – and many other methods have been used in TA research. As TA is generally a method for *across* dataset analysis (although it has been also used in case study research; e.g., Cedervall and Åberg 2010; Manago 2013), what is an adequate sample size? How many interviews should be conducted? How many participants should be recruited? How many hours of data should be recorded? These questions are perhaps some of the thorniest for qualitative research, although some concerns around sample size justification perhaps hark back to broader positivist-empiricist concerns with representation and generalizability, now connected to power analyses in statistical research.

Perhaps the most commonly used criterion for determining sample size in TA is “saturation” – such as claims that participants “were recruited until saturation was reached” (Gershgoren et al. 2016, p. 130; see also ► Chaps. 55, “Reporting of Qualitative Health Research,” and ► 58, “Appraisal of Qualitative Studies”). Here,

“saturation” typically refers to information redundancy, or collecting data until no new information is generated (there are other definitions), and some TA researchers have suggested saturation can be achieved in as few as 6–12 interviews (Guest et al. 2006; Ando et al. 2014) or 5 focus groups (Namey et al. 2016). Such papers are often cited to provide justification for relatively small sample sizes in TA research, but bold claims about saturation warrant interrogation. One problem with an information redundancy conceptualization is that it relies on an understanding of meaning as transparent and obvious *prior to* analysis. As TA (often) involves identifying new patterns of meaning, and this usually happens *after* data collection, *analysis* is necessary to judge whether the information generated by participants offers something new or not. Researchers who claim saturation, then, seem to rely on potentially superficial impressions made of data during data collection. This approach is more compatible with coding-reliability versions of TA. Where saturation has (attempted to be) operationalized, it is often within an implicit coding-reliability approach to TA (e.g., Guest et al. 2006). With such often surface-meaning-based and early-conceptualized analysis, it is easy to see how the appearance of “saturation” might be achieved in relatively few interviews. A recent paper suggested that if coding in TA moves beyond the surface level, larger samples are needed to achieve saturation (Hennink et al. 2016), and reported more conceptual codes achieved “saturation” with 16–24 interviews or not at all.

More problematic for using saturation as the rationale for sample size relates to the underlying philosophy of the research – in qualitative approaches that emphasize the partial, multiple, and contextual nature of meaning, and view knowledge as the actively *created* product of the interpretive efforts of a particular researcher (or researchers), combined with the dataset, the concept of saturation stops making sense (Malterud et al. 2016). We remain skeptical of the usefulness of the saturation concept, particularly when conceptualized as information redundancy, for determining sample size in TA research, and do not think it is useful for much “Big Q” TA. Researchers who use saturation need to do so from a position of theoretical “knowingness,” understanding the assumptions embedded in (their particular iteration of) this concept, and whether those are compatible with the underlying philosophy of their research.

What does this mean for sample sizes in TA research? Unfortunately, there are *no magic formulas* for determining sample size in TA research! We urge readers to be skeptical of anyone proffering simple formulas (e.g., Fugard and Potts 2015), as they always contain inbuilt assumptions (see Braun and Clarke 2016). Sample size is most often informed by various contextual and pragmatic considerations, some of which cannot be (wholly) determined in advance of data collection. Imagine a PhD student conducting TA research – their sample size could be informed by local “norms” around the appropriate scope of doctoral research, what is considered an acceptable sample size in journals the student hopes to publish their research in, and other such pragmatic “rules of thumb,” as well as more contextual considerations such as the breadth of their research question, the diversity within the population of study, and the amount and richness of data collected from each participant/case. Our pragmatic “rule of

thumb” is *at least* five or six interviews for a (*very*) small project, assuming the data are rich, the sample relatively homogenous, the research question focused, and the output an unpublished dissertation (for more “rules of thumb” advice on sample size, see Braun and Clarke 2013). It is also important to reflect on the sorts of claims made about themes developed, in light of the sample size.

6 Six Phases of Reflexive Thematic Analysis

Having discussed some conceptual and design issues, we now provide research-illustrated (see Box 1) discussion around the phases of *doing* (reflexive) TA, aligned to the “six-phase” approach we have developed, noting this as a reflexive and recursive, rather than strictly linear, process. For more practical step-by-step guidelines, see Braun and Clarke (2006, 2012, 2013) and Terry et al. (2017).

Box 1 The Men’s Embodiment in Rehabilitation Study

Despite three decades of calls for more research into men’s health, it continues to be underresearched, in general and in rehabilitation studies in Aotearoa/New Zealand. This project, theoretically located at the intersection of critical health psychology, critical rehabilitation studies, and critical disability studies, was designed to explore men’s experiences, practices, and sensemaking regarding male bodies undergoing rehabilitation for illness or impairment. Data were generated through one-to-one qualitative interviews with 20 men in various states of health and fitness, who had experienced recent (and extensive) rehabilitative treatment of some kind. Gareth Terry was the primary investigator and David Anstiss a postdoc researcher on the project. Rehabilitation was a new area of research focus for both.

Familiarization, which requires the researcher to shift focus from data generation (including transcription) to analysis, is fundamentally about appreciating the data *as* data. The process involves becoming “immersed” in the data and connecting with them in different ways: engaged, but also relaxed; making *casual* notes, but being thoughtful and curious about what you are reading. It is not about attaching formal labels – that comes later – but about looking for what is interesting about the data and what you notice about possibilities, connections (between participants, data, and existing literature), and quirks, which may add depth and nuance to your later coding. It can be one of the most enjoyable phases of the analytic process, and by providing a solid foundation of interrogating and thus “knowing” your data, it certainly makes the rest of the analysis *much* more enjoyable.

Practically, familiarization includes listening to audio data, watching video data, and/or reading and rereading textual data, “noticing” interesting features, and making notes about individual data items, as well as the whole dataset. These notes should be shaped by your research question(s), as well as broader questions about

what is going on in the data. For instance, in the men and embodiment project (Box 1), some key questions that Gareth and David asked of the data related to how “typical” understandings of men’s health might intersect with newly-experienced impairments and disabling environments – both social and physical. As researchers are new to the topic, familiarization was a crucial “entry point” into the data, providing them with an opportunity to closely read and thoroughly engage with the data, and giving room for reflexivity – asking questions of themselves and how they responded to the data. Consequently, much of the familiarization that occurred in the project involved making sense of ideas in the data that were new to them, which correspondingly made them aware of their own *abled* experiences and assumptions. Gareth and David engaged in the familiarization process concurrently, meeting several times to discuss their “noticings” and notes in detail. This process was not intended to produce any “consensus,” but rather to gain greater initial insight through sharing each other’s perspective on the data.

Generating codes moves to more detailed and systematic engagement with the data. We sometimes suggest familiarization could be done with a glass of wine, but coding needs coffee (or a good cup of tea). The coding phase in (our) TA is about focused attention, to systematically and rigorously make sense of data. If the familiarization phase could be considered a somewhat “loose” route into engaging with the data, the coding phase is about succinctly and systematically identifying meaning throughout the dataset. Data are organized around similar meanings and the content reduced into collated chunks of text. As a process, coding involves attaching pithy, clear labels (codes) to “chunks” of data, to help you organize the data around meaning-patterns (developed in later phases).

There are two broad orientations to coding: an *inductive* orientation, where the researcher starts the analytic process *from* the data, working “bottom-up” to identify meaning without importing ideas, and a *deductive* orientation, where the researcher approaches the data with various ideas, concepts, and theories, or even potential codes based on such, which are then explored and tagged within the dataset. In practice, any researcher *will* approach the data with preconceived ideas based on their existing knowledge and viewpoints. Coding inductively does not mean that we assume the researcher is a “blank state,” but, instead, that the starting point of the analysis is with the data, rather than existing concepts or theories (Terry et al. 2017).

Another consideration is the level at which “meaning” is identified and coded for – something partly informed by the epistemological approach of a project. *Semantic* codes stay at the “surface” of the data, capturing explicit meaning, close to participant language. *Latent* codes focus on a deeper, more *implicit* or conceptual level of meaning, sometimes quite abstracted from the explicit content of the data. The boundaries between these types of codes in practice are not always distinct; these codes represent ends of a continuum of ways of looking at data, rather than a binary. Initial coding for most TA projects is often semantic, and it can be hard to move beyond this level, to start to see the meaning *beyond* the obvious. As researchers become more experienced, or an analysis develops, latent-level meaning can be easier to “see” – but whether latent meaning is included can depend on the aims of the project.

Table 1 Example of coding, P1 (“Derek”) from the men’s embodiment and rehabilitation study

Data	Code
GT: So has that (pause) has that (pause) situation resolved itself a little bit	
P1: Nah I didn’t see my father at all (inbreath). I was really hopeless, lots of things happened that (pause) he was not supportive at all	“Harden up” mentality Relational breakdown with father Lack of affective recognition = unsupportive
GT: Yep	A long time to begin accepting
P1: He was a real cuppa concrete guy you know which is just not how it works (pause, inbreath); you know I was a lot worse then coz that was sort of two years in	Statute of limitations on acceptance Importance of supportive partner Importance of ACC support Financial means to get best treatment
GT: Yep	Recovery as relational
P1: Sort of accept it and move on you know. I was (pause) I’m not giving up (inbreath); there have been a few people that (pause) I mean I was fortunate my wife was very into research and stuff and she found (inbreath) my [private rehab clinic], and ACC paid for it; I think about 500 hours of rehab	

Transcription notation: *underline*, participant emphasis; (pause), pause in speech; (laughs), laughter from speaker

Gareth and David coded both semantically and latently, first working independently in the early part of this phase, then more collaboratively. Earlier familiarization discussions informed coding, but their emphasis was on generating a wide variety of codes to discuss and refine (see Table 1 for a brief example). Understanding that researchers look at data through their own lenses, and make interpretative choices throughout the analytic process (see Braun and Clarke 2016), they aimed to develop a diverse range of codes to build themes from, rather than trying to reach a consensus. Their practice demonstrates how more than one coder can work effectively with reflexive approaches to TA.

Constructing themes continues the active process of the previous phases. Themes are built, molded, and given meaning at the intersection of data, researcher experience and subjectivity, and research question(s). Because themes *do not emerge* fully-formed from the data, the process of constructing them is akin to processes of engineering or design. Prototypes (or *candidate themes*) are developed from the analytic work of the earlier phases, and “tested out” in relation to the research question/dataset overall. Knowing that not all candidate themes will necessarily survive this early development process is vital to not getting too attached. Good themes are those that tell a coherent, insightful story about the data in relation to the research question.

There are two key ways to develop codes into candidate themes. The first involves using codes as building blocks – similar codes are collated, together with their associated data, into coherent clusters of meaning that tell a story about a particular aspect of the dataset. This approach is most commonly how researchers move from codes to constructing (candidate) themes. However, sometimes a code may be “substantial” enough to be “promoted” to a theme – if it contains a central organizing idea that captures a *meaningful* pattern across the dataset, as well as different manifestations of that pattern. A common pitfall in (reflexive TA) theme development is identifying a feature of the data, rather than meaning-based patterns – features are somewhat akin to the idea of themes as domain summaries. For instance, men in the embodiment and rehabilitation project would often use humor to deflect questions about emotion and rehabilitation. This is potentially important information, but “humor” in and of itself is a *feature* of the dataset, not a meaning-based pattern. If the researchers could identify a conceptual meaning related to the use of humor, it might work as a theme, but alone, it does not.

Gareth and David again worked independently and collaboratively in the early stages of theme construction, meeting regularly to discuss candidate themes. Their meetings took the form of a kind of “theme off”: each presented their candidate themes, including preliminary theme names and definitions (discussed soon); they then “tussled” with each theme, and the collection of themes, to identify the most meaningful potential themes, the ones that collectively told the best story of the data. Thematic mapping – a process of visually exploring potential themes and subthemes, and connections between them (Braun and Clarke 2006) – was useful. Figure 1 maps the six candidate themes produced through this process, with all being relevant to the research question. Three of the initial themes Gareth and David constructed independently were similar enough to collapse into the single “bodies about more than roles and functions” theme – Gareth had identified a theme called “multiple embodiments”; David had two called “demanding embodiment” and “knowledge about bodies.” Their process demonstrates a way of working together, analytically, outside a consensus-building model.

The phases of *revising* and *defining themes* are particularly important, precisely because candidate themes *are* effectively prototypes. Sometimes they do not work! It can be difficult to “let go” of our early ideas, but holding too tightly to a candidate theme can potentially result in analytic “thinness” or conceptual overlap. The story being told about the data risks being diminished in richness, or conceptually confused, by inclusion of weak or overlapping themes. Having clear definitions of each theme – a paragraph delineating the theme’s boundaries and central organizing concept (for examples, see Terry et al. 2017) – helps clarify the essence and scope of each theme. Indeed, it was such descriptions of Gareth and David’s three candidate themes that highlighted their similarity and led to combination as a single theme (note that typically, not every facet survives this review process!).

Key to reviewing and defining is compiling all coded data for each of the candidate themes and reviewing them to ensure that the data relate to a central organizing concept; another stage of review involves checking the themes against the whole dataset. It is also important to develop a clear sense of how each theme

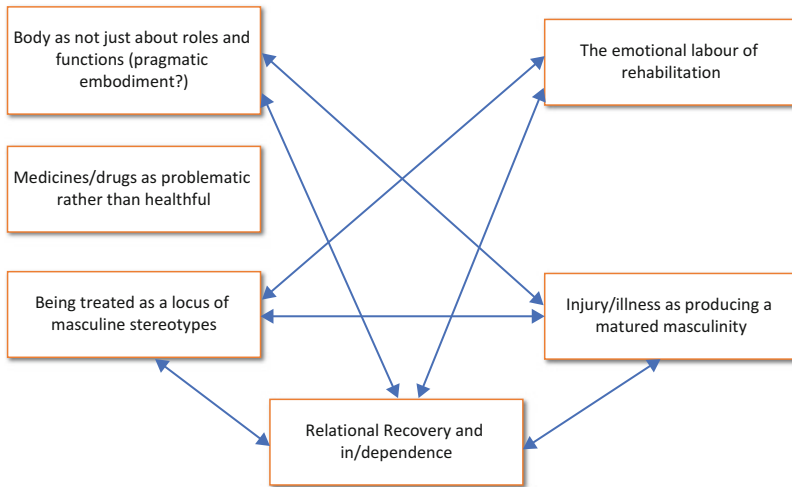


Fig. 1 Map of candidate themes from the men’s embodiment and rehabilitation study

relates to the others. Thematic maps can be useful to visualize how the themes fit together and tell the *overall story* of your data – and to check that themes do not overlap. We often move from “early” maps for candidate themes, through to “final” maps when the revising phase is complete. For the men’s embodiment project, it became clear when developing the (early) thematic map (Fig. 1), and comparing the definitions of each candidate theme, that some of the relationships between themes were stronger than others. The “resistance to medication” theme appeared quite independent (no connecting arrows); the other five themes all spoke to “relational outworking” of rehabilitation - how the men related to their bodies, and to others, and how they were related to *by* others. Thinking more deeply about the central organizing concept for each theme led Gareth and David to conclude that a notion of *recovery as relational* acted as an overarching theme (see Braun and Clarke 2013), an “umbrella” that contained three, related, themes: “bodies about more than roles and functions;” “being treated as a locus of masculine stereotypes;” and “the emotional labor of rehabilitation.” As analysis is a task of telling a compelling story about (aspects of) the data, they set aside one strong theme, “injury/illness producing a matured masculinity,” for future analysis oriented to *identity*. Through this process of revising themes, you aim for an in-depth and nuanced understanding of the central organizing concept and boundaries of each theme, including any subthemes (and overarching themes), and the overall theme story. Tables and similar tools can also facilitate in clearly identifying boundaries and structures of themes, in these phases.

Defining themes often leads to tighter/clearer theme names, which signal the scope and “core” of each theme. When you arrive at this point in the process, theme names will likely be somewhat makeshift – perhaps lengthy, or alternatively just one word – and only *provisionally* capture the content and scope of each theme. Final

theme names should succinctly cue the reader in to what they can expect to read about in the theme, *and* draw them into wanting to read the analysis!

The revising and defining phases seek to ensure that themes, *and* theme names, clearly, comprehensively *and* concisely capture what is meaningful about the data, related to the research question, getting you close to a “completed” analysis. The final phase, *producing the report*, is not, however, purely a writing-up exercise. Producing the report often serves as a final test of how well the themes work, individually in relation to the dataset, and overall. Revisiting the research question, your notes from the earlier phases of familiarization and coding, your lists of codes, and theme definitions can be useful to ensure that the final themes remain close to the data *and* answer the research question well (your research question can be “tweaked” for better fit at this point). The scholarly process of making connections to existing research and literature on the topic of interest, and weaving this in to the written results and discussion, may offer final moments of inspiration and a deeper insight into the analysis. Therefore, we urge researchers to view this phase as the final stage of analysis, and be open to making further revisions to the themes’ content, structure, and names. It may be that when you start to write your analytic narrative around the data extracts, you decide that some participant quotations do not as clearly demonstrate the point as other quotations do – your analytic claims may shift to reflect this. Or, as you write up your themes, you might notice that a change to the order in which they are presented would help with the flow of the story of your data, and so on. It can be useful to draw on our 15-point checklist (see Braun and Clarke 2006; Terry et al. 2017) to check the strength of your analysis and consistency across the report. We emphasize, again, that the epistemological position you have claimed will inform the terminology you use and the way you treat the data.

7 Conclusion and Future Directions

Despite TA providing an accessible method for (novice) qualitative researchers, there *are* plenty of potential traps and ways you can go wrong. Having read this chapter, you will be well-equipped to avoid many of these! In our view, good quality TA requires a degree of “theoretical knowingness” – an understanding of the philosophical basis of enquiry. This means, for instance, understanding the assumptions underpinning coding reliability or consensus coding practices, and understanding why these are *not* compatible with Big Q qualitative enquiry. Viewing theory as something that we *do*, rather than an abstract consideration divorced from the practical processes of conducting research, helps realize how essential this is. Imagine a supervisor telling a student “stop overthinking things and just get on with it.” Such advice suggests theory is *separate from* “getting on with it,” but theory is meshed into everything. We *do* theory all the time, in how we relate to participants, in our interviewing “style,” in how exactly we transcribe our interviews. . . All of these, and *many* other practical elements of research, reflect theoretical assumptions (implicitly made *or* actively chosen) about the nature of enquiry and what counts as meaningful knowledge.

Understanding theoretical bases also means you can avoid inconsistencies which mar too much published TA work. Too often we read research that contains statements like “following the procedures outlined by Braun and Clarke. . .” and then a description of analytic procedures that have little or no relation to those we outline. Sometimes coding reliability and reflexive approaches to TA are both cited, or the procedures combined, without any acknowledgment or justification of merging two philosophically divergent approaches. Sometimes researchers attribute analytic processes and concepts associated with grounded theory to TA (such as constant comparison, line-by-line coding, open coding, categories and subcategories, saturation; see Braun and Clarke 2013). Sometimes researchers claim domain summaries as conceptually founded patterns. . . Sometimes we really are left wondering whether the authors have actually *read* Braun and Clarke (2006)!

Is this us being picky, or does this really matter? Are we succumbing to methodolatry – the prioritizing of procedure above all else – something that “method-obsessed” psychologists, such as ourselves, are thought to be particularly vulnerable to? Does it matter if the end product (“the results” section) is good? We think it does matter! To us, these method(ological) “choices” – for example, combining consensus coding with Braun and Clarke (2006) – seem rarely to be made knowingly or reflexively. They seem instead to reflect a lack of understanding of, or caring about, the philosophical underpinnings of (Big Q) qualitative research or, perhaps, an (knowing or unknowing) acquiescence to the notion that (post)positivism is the only valid philosophy for research. Vitaly, in a context in which there is not only much confusion about qualitative research, and its philosophical underpinnings, but much critique, these practices serve to create further confusion and even give the critiques some validity!

Furthermore, when published work is internally incoherent, or does not follow best- or even good-practice guidelines for any “school” of TA, this can be confusing, particularly for qualitatively inexperienced or student readers, who may assume publication is a guarantee of quality! We have often led a critical discussion of a published paper that ends with a student asking some variety of the question: “how does this stuff get published?” For the future of TA, all of us – scholars and researchers doing TA, reviewers, and editors – need to work hard to ensure consistency and quality in published TA.

References

- Ando H, Cousins R, Young C. Achieving saturation in thematic analysis: development and refinement of a codebook. *Compr Psychol*. 2014;3:4.
- Aronson J. A pragmatic view of thematic analysis. *Qual Rep*. 1994;2(1):1–3. <http://www.nova.edu/ssss/QR/BackIssues/QR2-1/aronson.html>.
- Boyatzis RE. *Transforming qualitative information: thematic analysis and code development*. Thousand Oaks: Sage; 1998.
- Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101.
- Braun V, Clarke V. Thematic analysis. In: Cooper H, editor. *APA handbook of research methods in psychology, Research designs*, vol. 2. Washington, DC: APA books; 2012. p. 57–71.

- Braun V, Clarke V. *Successful qualitative research: a practical guide for beginners*. London: Sage; 2013.
- Braun V, Clarke V. (Mis)conceptualising themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis. *Int J Soc Res Methodol*. 2016;19(6):739–43.
- Braun V, Clarke V, Rance N. How to use thematic analysis with interview data. In: Vossler A, Moller N, editors. *The counselling and psychotherapy research handbook*. London: Sage; 2014. p. 183–97.
- Brewster ME, Velez BL, Mennicke A, Tebbe E. Voices from beyond: a thematic content analysis of transgender employees' workplace experiences. *Psychol Sex Orient Gend Divers*. 2014;1(2): 159–69.
- Brooks J, McCluskey S, Turley E, King N. The utility of template analysis in qualitative psychological research. *Qual Res Psychol*. 2015;12(2):202–22.
- Cedervall Y, Åberg AC. Physical activity and implications on well-being in mild Alzheimer's disease: a qualitative case study on two men with dementia and their spouses. *Physiother Theory Pract*. 2010;26(4):226–39.
- Connelly LM, Peltzer JN. Underdeveloped themes in qualitative research: relationships with interviews and analysis. *Clin Nurse Spec*. 2016;30:51–7.
- Dapkus MA. A thematic analysis of the experience of time. *Personal Process Individ Differ*. 1985;49(2):408–19.
- DeSantis L, Ugarriza DN. The concept of theme as used in qualitative nursing research. *West J Nurs Res*. 2000;22(3):351–72.
- Finlay L, Gough B, editors. *Reflexivity: a practical guide for researchers in health and social sciences*. Oxford: Blackwell Science; 2003.
- Fugard AJB, Potts HWW. Supporting thinking on sample sizes for thematic analysis: a quantitative tool. *Int J Soc Res Methodol*. 2015;18(6):669–84.
- Gershgoren L, Basevitch I, Filho E, Gershgoren A, Brill YS, Shink RJ, Tenebaum G. Expertise in soccer teams: a thematic enquiry into the role of shared mental models within team chemistry. *Psychol Sport Exerc*. 2016;24:128–39.
- Gough B, Madill A. Subjectivity in psychological research: from problem to prospect. *Psychol Methods*. 2012;17(3):374–84.
- Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods*. 2006;18(1):59–82.
- Guest G, MacQueen KM, Namey EE. *Applied thematic analysis*. Thousand Oaks: Sage; 2012.
- Hennink MM, Kaiser BN, Marconi VC. Code saturation versus meaning saturation: how many interviews are enough? *Qual Health Res*. 2016;27(4):591–608.
- Hill M, Welch V, Gadda A. Contested views of expertise in children's care and permanence proceedings. *J Soc Welf Fam Law*. 2017;39(1):42–66.
- Holton GJ. *Thematic origins of scientific thought: Kepler to Einstein*. Cambridge, MA: Harvard University Press; 1973.
- Joffe H. Thematic analysis. In: Harper D, Thompson AR, editors. *Qualitative methods in mental health and psychotherapy: a guide for students and practitioners*. Chichester: Wiley; 2011. p. 209–23.
- Kidder LH, Fine M. Qualitative and quantitative methods: when stories converge. In: Mark MM, Shotland L, editors. *New directions in program evaluation*. San Francisco: Jossey-Bass; 1987. p. 57–75.
- King N. Using templates in the thematic analysis of text. In: Cassell C, Symon G, editors. *Essential guide to qualitative methods in organisational research*. London: Sage; 2014. p. 256–70.
- Kinsky G, Strunk WO. Was Mendelssohn indebted to weber? An attempted solution of an old controversy. *Music Q*. 1933;19(2):178–86.
- Langdrige D. *Introduction to research methods and data analysis in psychology*. Harlow: Pearson; 2004.

- Lazarsfeld PF, Merton RK. The psychological analysis of propaganda. In: *Proceedings, Writers' Congress: Berkeley and Los Angeles: University of California Press; 1944.*
- Madill AL. Qualitative research is not a paradigm. *Qual Psychol.* 2015;2(2):214–20.
- Malterud K, Siersma VK, Guassora AD. Sample size in qualitative interview studies: guided by information power. *Qual Health Res.* 2016;26(13):1753–60.
- Manago AM. Negotiating a sexy masculinity on social networking sites. *Fem Psychol.* 2013;23(4):478–97.
- McSherry W, Bloomfield S, Thompson R, Nixon VA, Birch C, Griffiths N, Fisher S, Boughey AJ. A cross-sectional analysis of the factors that shape adult nursing students' values, attitudes and perceptions of compassionate care. *J Res Nurs.* 2017;22(1–2):25–39.
- Mellor R, Lovell A. The lived experience of UK street-based sex workers and the health consequences: an exploratory study. *Health Promot Int.* 2012;27(3):311–22.
- Miles MB, Huberman AM. *Qualitative data analysis: an expanded sourcebook.* 2nd ed. Thousand Oaks: Sage; 1994.
- Moran C, Lee C. Selling genital cosmetic surgery to healthy women: a multimodal discourse analysis of Australian surgical websites. *Crit Discourse Stud.* 2013;10(4):373–91.
- Nadin S, Cassell C. Using data matrices. In: Cassell C, Symon G, editors. *Essential guide to qualitative methods in organisational research.* London: Sage; 2014. p. 271–87.
- Namey E, Guest G, McKenna K, Chen M. Evaluating bang for the buck: a cost-effectiveness comparison between individual interviews and focus groups based on thematic saturation levels. *Am J Eval.* 2016;37(3):425–40.
- Redmond L, Suddick K, Earhart GM. The lived experience of freezing in people with Parkinson's: an interpretive phenomenological approach. *Int J Ther Rehabil.* 2012;19(3):169–77.
- Ritchie J, Lewis J. *Qualitative research practice: a guide for social science students and researchers.* London: Sage; 2003.
- Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. *Analysing qualitative data.* London: Taylor & Francis; 1994. p. 173–94.
- Roditis ML, Halpern-Felsher B. Adolescents' perceptions of risks and benefits of conventional cigarettes, e-cigarettes, and marijuana: a qualitative analysis. *J Adolesc Health.* 2015;57:179–85.
- Sandelowski M, Leeman J. Writing usable qualitative health research findings. *Qual Health Res.* 2012;22(10):1404–13.
- Smith J, Firth J. Qualitative data approaches: the framework approach. *Nurs Res.* 2011;18(2):52–62.
- Terry D, Ali M, Lê Q. Asian migrants' lived experience and acculturation to Western health care in rural Tasmania. *Aust J Rural Health.* 2011;19(6):318–23.
- Terry G, Braun V. "I think gorilla-like back effusions of hair are rather a turn-off": 'excessive hair' and male body hair (removal) discourse. *Body Image.* 2016;17:14–24.
- Terry G, Hayfield N, Clarke V, Braun V. Thematic analysis. In: Willig C, Stainton Rogers W, editors. *The sage handbook of qualitative research in psychology.* 2nd ed. London: Sage; 2017. p. 17–37.
- Titlestad A, Pooley JA. Resilience in same-sex-parented families: the lived experience of adults with gay, lesbian, or bisexual parents. *J GLBT Fam Stud.* 2014;10(4):329–53.
- Weatherhead S, Daiches A. Muslim views on mental health and psychotherapy. *Psychology and Psychotherapy: Theory, Research and Practice.* 2010;83(1):75–89.
- Willig C. *Introducing qualitative research in psychology: Adventures in theory and method.* 3rd ed. Maidenhead, Berkshire: McGraw-Hill Education (UK); 2013.
- Winder AE, Hersko M. A thematic analysis of an outpatient psychotherapy group. *Int J Group Psychother.* 1958;8(3):293–300.
- Yardley L. Demonstrating validity in qualitative psychology. In: Smith JA, editor. *Qualitative psychology: a practical guide to research methods.* 3rd ed. London: Sage; 2015. p. 257–72.