

# Conceptual and design thinking for thematic analysis

Virginia Braun and Victoria Clarke

## *Qualitative Psychology*

### Abstract

Thematic analysis (TA) is widely used in qualitative psychology. In using TA, researchers must choose between a diverse range of approaches that can differ considerably in their underlying (but often implicit) conceptualizations of qualitative research, meaningful knowledge production and key constructs such as themes, as well as analytic procedures. This diversity within the method of TA is typically poorly understood and rarely acknowledged, resulting in the frequent publication of research lacking in design coherence. Furthermore, because TA offers researchers something closer to a method (a trans-theoretical tool or technique) rather than a methodology (a theoretically-informed framework for research), one with considerable theoretical and design flexibility, researchers need to engage in careful *conceptual and design thinking* to produce TA research with methodological integrity. In this paper, we support researchers in their conceptual and design thinking for TA, and particularly for the reflexive approach we have developed, by guiding them through the conceptual underpinnings of different approaches to TA, and key design considerations. We outline our typology of three main “schools” of TA – *coding reliability*, *codebook* and *reflexive* – and consider how these differ in their conceptual underpinnings, with a particular focus on the distinct characteristics of our reflexive approach. We discuss key areas of design – research questions, data collection, participant/data item selection strategy and criteria, ethics, and quality standards and practices – and end with guidance on reporting standards for reflexive TA.

**Keywords:** Design coherence, methodological integrity, reflexivity, participants, saturation

## Conceptual and design thinking for thematic analysis

Thematic analysis (TA) is widely practiced in qualitative psychology. What distinguishes TA from most other qualitative analytic approaches – such as grounded theory and narrative analysis – is that it is more akin to a method (a trans-theoretical tool or technique) than a methodology (a theoretically-informed framework for research). Approaches like grounded theory and narrative analysis have been dubbed “off-the-shelf” methodologies (Chamberlain, 2012), in the sense that they encompass both analytic techniques *and* philosophical assumptions, a theoretical framework, and steer toward particular types of research question, participant/data item selection practices, methods of data collection, and quality procedures. This has led some qualitative methodologists to argue that the use of TA demands more conceptual and design thinking from researchers compared to the use of off-the-shelf methodologies (McLeod, 2015; Willig, 2013). For Willig (2013), TA is not the “easy option” (p. 66) it is often perceived as, because the researcher “needs to do a lot of conceptual work before they can embark upon the research itself” (p. 65). McLeod (2015) described TA as “a good choice for researchers who feel confident that they know what they are trying to achieve” (p. 147).

The status of TA as a *method* is sometimes framed as an *obstacle* to good practice, particularly for qualitative newcomers. Some have argued that the combination of the reputation of TA as an accessible method and its lack of inbuilt theory can lead researchers to make the mistake of conducting TA without explicitly locating it theoretically (King & Brooks, 2017; Willig, 2013). Brown and Locke (2017) suggested that TA is popular in applied psychological research because it is perceived to allow researchers to “analyse their qualitative data for topic content without considering any methodological horrors” (p. 425).

That perception can result in the reporting of themes that have no explicit conceptual underpinning. This is a poor practice – as captured by *Qualitative Psychology’s* submission guidelines, which note explicitly that empirical research will be evaluated as to whether there is “adequate conceptualization (as opposed to simple description or reporting of themes)” (American Psychological Association, 2020). Our perspective on TA, as a *method*, is more optimistic: as the use of TA *requires* deliberation from researchers, the importance of a thoughtful, reflective research practice – a practice emphasized as *crucial* in many quality standards and guidelines (e.g., Elliott et al., 1999; Levitt et al., 2017; Yardley, 2015) – is highlighted. We are not alone in this position: others have criticized the “predetermined” nature of off-the-shelf *methodologies* for allowing for “thought-less” qualitative research (e.g., Chamberlain, 2012). At its best, TA helps us not only make visible the various elements that need to come together for successful qualitative analysis characterized by integrity, but also to consider how they connect and build on each other. As TA may be taught or learned early as a qualitative approach,<sup>1</sup> we have an important opportunity – or obligation – to make new researchers recognize and think (deeply) about the many layers of conceptual thinking behind all (good) research practice.

Although we value the flexibility of TA, we appreciate that conceptualizing and designing a TA study can be daunting, especially for qualitative newcomers, because the

---

<sup>1</sup> Some argue that it *should* be taught early precisely *because* it makes the “mechanics” – the conceptual and design thinking – of qualitative research visible to new researchers. We thank an anonymous reviewer for raising this point.

qualitative methodological literature is vast and complex, and provides numerous contradictory and contested accounts. In this paper, we aim to support such researchers by providing guidance on *conceptual and design thinking* for TA, and particularly for the reflexive approach we have developed (e.g., Braun & Clarke, 2006, 2012, 2019a). To do so, we draw on both the TA and wider qualitative methodological literature, in psychology and related disciplines.

Conceptual and design thinking involves all the elements of a research project, assessing whether different elements will *work together*, and producing an explanation for choices made (see Box 1 for an overview). A general principle for qualitative research design is coherence or “fit” (Braun & Clarke, 2013; Willig, 2013), where the research aims and purpose, philosophical, theoretical and methodological assumptions, and methods cohere together (Chamberlain et al., 2011; Tracy, 2010). Levitt et al. (2017) proposed a similar concept of *methodological integrity* to capture when:

*research designs and procedures (e.g., autoethnography, discursive analysis) support the research goals (i.e., the research problems/questions); respect the researcher’s approaches to inquiry (i.e., research traditions sometimes described as world views, paradigms, or philosophical/epistemological assumptions); and are tailored for fundamental characteristics of the subject matter and the investigators. (pp. 9-10)*

The status of TA as a flexible method, rather than a delimited methodology, provides one challenge for qualitative newcomers in achieving methodological integrity. The multiplicity *within* TA provides another. TA is best thought of as a *family* of methods with some elements in common – alongside some substantial divergences in philosophical assumptions, conceptualizations of key constructs, and analytic procedures. However, this

diversity is often poorly understood, and rarely acknowledged, with considerable evidence of seemingly “thoughtless” muddling together of conceptually incoherent practices in published research (Braun & Clarke, 2020).

This paper is divided into two sections. The first focuses on *conceptual thinking* for TA. Understanding the conceptual underpinnings of reflexive TA, and how these differ from those associated with other types of TA, is crucial for methodological integrity. Our discussion centers on a typology of three main “schools” of TA, with a particular focus on our reflexive approach. This typology captures some of the key areas of divergence within TA as a family of methods. This section aims to assist researchers in making deliberative decisions about their approach to TA, and using that approach knowingly, “owning” the embedded research values (Elliott et al., 1999). The second section centers on *design thinking* in reflexive TA. It covers matters of research questions, data collection methods and sources, participant group/dataset constitution and size, ethics, and quality standards and practices. This section will help researchers to design a reflexive TA study with methodological integrity. We end with a discussion of reporting standards for reflexive TA research.

[Insert Box 1 about here]

### **Conceptual Thinking for TA Research**

TA methods, as a family, share the following characteristics: theoretical flexibility (albeit constrained to a greater or lesser degree by assumptions about meaningful knowledge production and how qualitative research is conceptualized); procedures of coding and theme development; the possibility of inductive and deductive orientations to analysis (although there can be marked differences in how these orientations are

conceptualized); and the possibility of coding for both manifest (semantic or descriptive) meanings – the meanings directly observable on the surface of the data – and latent (implicit or conceptual) meanings – the meanings that underlie the data surface (Boyatzis, 1998; Braun & Clarke, 2006; Joffe, 2012). At the same time, there are some notable differences between various TA approaches, underpinned in some cases by markedly different conceptualizations and values.

The challenge for the qualitative researcher, and the starting point for conceptually coherent design in TA research, is to understand the particulars of their chosen approach to TA, and where it sits on what we conceive of as a TA spectrum. This is a *challenge* because specific procedures rely on and encode sets of underlying research values, but these are not always explicitly stated (see Carter & Little, 2007). Indeed, in some contexts, particularly those dominated by quantitative positivism, research values themselves might be taught as singular and universal, or indeed just be assumed. We view clarity on research values as fundamental to quality (TA) research. This is especially important in qualitative research because of the diversity of research values and associated ontological (theories of reality and being) and epistemological (theories of meaningful knowledge and knowledge production) assumptions. Understanding that TA is not one method, but a *cluster* of methods underpinned by different conceptual models and research values, facilitates the practices of owning one's (theoretical and methodological) perspective (Elliott et al. 1999), and demonstrating sensitivity to (theoretical) context (Yardley, 2015), highlighted in quality standards and principles.

### **A Typology of Thematic Analysis: Coding Reliability, Codebook and Reflexive**

We distinguish between three main schools of TA which we call *coding reliability*, *codebook*, and *reflexive*.<sup>2</sup> We find Kidder and Fine's (1987) distinction between "small q" and "Big Q" qualitative research a useful one for understanding the differences across these types. Small q involves qualitative data, but is informed by quantitative/(post)positivist<sup>3</sup> research values and practices; Big Q involves both qualitative data, and values and practices embedded in a qualitative paradigm. Coding reliability TA exemplifies small q qualitative, reflexive TA Big Q, and codebook sits somewhere between small q and Big Q. These types can be conceptualized as located on a *spectrum* of TA, with coding reliability approaches at the small q/(post)positivist end of the spectrum and reflexive approaches at the other –

---

<sup>2</sup> These names reflect the key characteristics of coding in each type and thus only capture *one* element of differences across the approaches, and indeed the practice of doing TA.

<sup>3</sup> We use the term *(post)positivist* to signal the contested terrain of positivism within psychology. Some predominantly associate positivism and postpositivism with quantitative research and argue that the default paradigm for quantitative research is now postpositivism, rather than (naïve) positivism, following the critiques of Popper (1959) and others (e.g., Ponterotto, 2005). Others view postpositivism as spanning both quantitative and qualitative research and associate it with a critical realist ontology and qualitative methodologies such as consensual qualitative research (CQR) and objectivist grounded theory (e.g., Morrow, 2007). *Our* use of (post)positivism reflects the distinction between *small q* ((post)positivist) and *Big Q* qualitative (Kidder & Fine, 1987), connected to paradigmatic values. We view coding reliability TA, as well as CQR and objectivist grounded theory, as examples of small q qualitative.



Big Q/non-positivist, constructionist – end (see Terry & Hayfield, 2020).<sup>4</sup> It is important to note that this is our *positioned* mapping of the “landscape” of TA. We are not necessarily describing the different approaches in ways that the authors who developed these would recognize, as we have sought to unravel unstated assumptions and tease out divergences in the way key concepts such as codes and themes are understood.

### ***Coding Reliability Thematic Analysis***

Coding reliability approaches are so named because the analytic procedures are oriented to establishing the “accuracy” or “reliability” of data coding, underpinned by a (post)positivist paradigm or research values (see Ponterotto, 2005). Stemming from that, there is a concern for controlling researcher subjectivity or “bias”, reliability and replicability of measurement, and generating as-objective-as-possible knowledge. Some coding reliability authors frame their approach as “bridging the divide” between positivist (quantitative) and interpretive (qualitative) paradigms through combining the use of qualitative techniques and data with (post)positivist values around meaningful knowledge production (e.g., Boyatzis, 1998; Guest et al., 2012). In describing this small q form of TA, we

---

<sup>4</sup> This typology captures much of the diversity within the TA family but there are other (often idiosyncratic) approaches that defy easy categorisation and combine elements of the different types (e.g., Buetow, 2010; Malterud, 2013); furthermore, the use of grounded theory coding techniques and other analytic practices (such as constant comparative analysis and memo writing) to develop themes from qualitative data – both demarcated as a distinct approach to TA known as *thematic coding* (e.g., Gibbs, 2007; Flick, 2014) and used more idiosyncratically – remains relatively common.

focus on the conceptualization of coding and themes, researcher subjectivity, meaningful knowledge production, and quality standards and practices.

Coding reliability TA typically involves some or all of the following. Themes developed early in the analytic process prior to or following some data familiarization, and often reflecting data collection questions. Themes as effectively *inputs* into the coding process rather than *outputs* from it. Themes conceptualized (implicitly) as “fossil[s] hidden in a rock” (King & Brooks, 2017, p. 220) or “diamonds scattered in the sand” (Braun & Clarke, 2016, p. 740), lurking in the data awaiting “discovery” by the researcher. Themes also tend to be understood (again implicitly) as topic rather than meaning based, as topic summaries – summaries or overviews of things said by participants in relation to a particular topic. Topics often map closely onto data collection questions. This means a topic summary “theme” is effectively a summary of responses to a data collection question. For example, an interview question might focus on barriers to African heritage women accessing professional support for postnatal depression and the topic of this question – the barriers to accessing support – becomes the focus of the “theme”. The “theme” is effectively a summary of all the main barriers discussed by participants. What unites the observations reported is the topic – the barriers – rather than a pattern of shared meaning evident across responses. This theme conceptualization facilitates *early* theme development/identification (i.e., before any substantial analysis has taken place), as it is relatively straightforward to identify topics without a detailed unpacking of *how* those topics were spoken about. There is often little that *unites* the meanings within a topic summary other than the topic. Such “themes” can be developed both inductively, following some data familiarization, and deductively, from prior research or theory.

In both inductive and deductive orientations to coding reliability TA, a codebook or coding frame is constructed to guide the allocation of data to the pre-determined themes.<sup>5</sup> Codebooks typically consist of a definitive list of codes/themes, a coding label and definition for each code/theme, instructions on how to identify each code/theme, including any exclusions, and examples of each code/theme. We use “code/theme” because there is not always a clear distinction between codes and themes in coding reliability approaches. Coding as a *process* is often foregrounded over the “code,” an analytic entity distinct from, but contributing to, a “theme.” The codebook is applied to all or a portion of the data, ideally by multiple coders working independently. The level of intercoder agreement is then calculated to provide a measure of coding reliability (O’Connor & Joffe, 2020) – the assumption being that “reliable” coding is possible and two or more researchers choosing to assign the same piece of data to the same code/theme is a meaningful measure of this. Some coding reliability researchers advocate for the use of coders who are “blind” to the research question or have no prior knowledge of the research area to minimize the “contamination” of the coding process with this knowledge, and to maximize objectivity (e.g., Bond et al., 2008). Final data coding is typically determined by agreement or consensus. One of the challenges for qualitative researchers in disciplines like psychology, where *qualitative* values remain marginal, is that the (post)positivist quality standards

---

<sup>5</sup> A deductive approach within coding reliability TA is often conceptualized as providing a tool for testing – refuting or confirming – a hypothesis. This model aligns with the deductive orientation and assumptions of *the scientific method*; it sits at odds with the use of deduction in many other versions of TA and indeed qualitative approaches that take a Big Q approach.

prioritized in coding reliability TA are often equated with quality practice in all forms of TA, including reflexive TA (Braun & Clarke, 2020). Instead, they reflect a particular set of theoretically embedded research values.

### ***Codebook Thematic Analysis***

Codebook approaches sit somewhere between the coding reliability and reflexive ends of the TA spectrum. Such approaches combine a more structured approach to coding, through the use of a codebook or coding frame, (some) early theme development, a (typical) conceptualization of themes as topic summaries, all associated with small q coding reliability approaches, with the Big Q values of reflexive TA (e.g., conceptualizing researcher subjectivity as a resource for research, and coding and interpretation of data as an inherently and inescapably subjective practice, which we discuss further below). Our label *codebook* encompasses approaches like matrix (e.g., Miles & Huberman, 1994; Nadin & Cassell, 2004), framework (e.g., Ritchie & Spencer, 1994), network (e.g., Attride-Stirling, 2001) and template (e.g., King, 2012) analysis, often developed for, and popular within, applied research. In codebook TA, codebooks are not typically used to facilitate the measurement of intercoder agreement but are rather oriented to pragmatic considerations such as meeting predetermined information needs (common in some areas of applied research), a team of data analysts working together, each coding different portions of the data (with the team potentially including qualitative novices or users/stakeholders with no research training), and/or a swift and “efficient” analysis (because of working to a tight deadline – of the funder/service). The codebook is used to record and or chart the developing analysis as well as to guide data coding. Some codebook authors argue that codebook approaches reflect a pragmatic compromise of some qualitative research values.

The open, exploratory, and (sometimes) inductive elements of qualitative research pose a challenge when practical constraints such as those detailed are present (Smith & Firth, 2011).

### ***Reflexive Thematic Analysis***

Reflexive approaches prioritize the values of Big Q qualitative paradigms and emphasize the inevitable subjectivity of data coding and analysis, and the researcher's active role in coding and theme generation (e.g., Gleeson, 2011; Hayes, 2000). As our main focus in this paper is on conceptual and design thinking *for reflexive TA*, we outline the key conceptual foundations of *our* reflexive approach in some detail in the next section.<sup>6</sup>

### **Conceptualizing *Reflexive Thematic Analysis***

Our Big Q approach to reflexive TA developed from a critique and rejection of the values underlying (post)positivist TA (Braun & Clarke, 2019a). TA, and qualitative research more broadly (e.g., Morrow, 2007), is often equated with the study of subjectivity and lived experience (e.g., Flick, 2014), and phenomenology (e.g., Guest et al., 2012, Joffe, 2012). Furthermore, the mapping of the conceptual foundations of qualitative research in psychology often frame different qualitative paradigms as reflecting different orientations

---

<sup>6</sup> We do *not* overview the six phases of the analytic process of reflexive TA – familiarization with the data, coding the data, generating initial themes from the codes and coded data, reviewing and developing themes, defining, naming and refining themes, writing up the report – as these are more *practice* oriented. They are also discussed extensively elsewhere (e.g., see Braun & Clarke, 2006, 2012).

to the study of *experience* (e.g., a distinction is commonly made between interpretivist-constructivist and ideological-critical qualitative paradigms, but both are conceptualized as oriented to the study of experience and subjectivity, see Morrow, 2007). However, as researchers schooled in social constructionism (see Gergen, 2015), we (and others) understand TA, and qualitative research, as extending *beyond* a concern for experiential phenomena to social processes and the social construction of meaning.<sup>7</sup> We are relatively unique among TA authors in making a distinction between experiential and constructionist orientations to TA (see also King, 2012; King & Brooks, 2017).

Broadly speaking, *experiential TA* (including reflexive TA when used in experiential orientations) is concerned with exploring the truth or truths of participants' contextually-situated experiences, perspectives and behaviors. It is typically underpinned by some form of realist (naïve and critical) ontology (see Maxwell, 2012) and a range of intersecting and overlapping epistemologies including interpretivism-constructivism, ideological-critical (see Morrow, 2007), contextualism (see Madill et al., 2000) and phenomenology (see Willig, 2013). The conceptualization of language is key to the experiential/constructionist distinction (Reicher, 2000). In experiential TA, language is conceptualized as reflecting the true nature of things or participants' contextually situated unique realities or truths (Braun & Clarke, 2013). *Constructionist* orientations to language are concerned with interrogating the rhetorical implications and effects of particular patterns of meaning and linguistic

---

<sup>7</sup> An orientation similarly captured in constructivist versions of grounded theory (e.g., Charmaz, 2006) and constructionist versions of narrative analysis (e.g., Sparkes & Smith, 2008).

practices (Braun & Clarke 2013). Language is conceptualized as active and symbolic, as creating rather than simply reflecting meaning. In constructionist TA, language is not treated as a simple conduit to access information. Constructionist TA research takes different forms; researchers can make claim to both relativist and critical realist ontologies and postmodern and poststructuralist epistemologies and methodologies (Clarke & Braun, 2014).

The theoretical flexibility of reflexive TA is often mistaken for theoretical *neutrality*. Like all forms of TA, reflexive TA reflects various theoretically-based assumptions about how knowledge is (best) produced (Mauthner & Doucet, 2003), and these are associated with *qualitative* paradigms. This Big Q position makes “pure” induction impossible; the researcher always brings philosophical meta-theoretical assumptions and *themselves* to the analysis, meaning an inductive orientation is better understood as “grounded” in data. A deductive orientation in reflexive TA involves using pre-existing theory as *a lens through which to interpret* the data; deductive reflexive TA is *not* about “testing” a pre-existing theoretical framework or hypothesis.

The core assumptions of reflexive TA can be summarized across ten points:

- 1) Researcher subjectivity is the primary “tool” for reflexive TA; subjectivity is not a problem to be managed or controlled, it is a resource for research (Gough & Madill, 2012). The notion of “researcher bias,” which implies the possibility of unbiased or objective knowledge generation, is incompatible with reflexive TA, as knowledge generation is inherently subjective and situated.
- 2) Following on from this, analysis and interpretation of data cannot be accurate or objective, but can be weaker (e.g., under-developed, unconvincing, thin,

- superficial, shallow) or stronger (e.g., compelling, insightful, thoughtful, rich, complex, deep, nuanced).
- 3) Good quality coding and themes result from dual processes of immersion or depth of engagement, and distancing, allowing time and space for reflection and for insight and inspiration to develop.
  - 4) Coding quality is not dependent on multiple coders; a single coder/analyst is typical in reflexive TA. Good coding (and theme development) can be achieved singly, *or* through collaboration, if it seeks to enhance reflexivity and interpretative depth, rather than consensus between coders.
  - 5) Themes are analytic outputs not inputs and are developed after coding and from codes (which are also analytic outputs); as Saldaña (2013) noted, a theme is “an *outcome* of coding... not something that is, in itself, coded” (p. 14).
  - 6) Themes are patterns of meaning anchored by a shared idea or concept (central organizing concept), not summaries of meaning related to a topic.
  - 7) Themes are not waiting in the data to “emerge” when the researcher “discovers” them; they are conceptualized as *produced* by the researcher through their systematic analytic engagement with the dataset, and all they bring to the data in terms of personal positioning and meta-theoretical perspectives.
  - 8) Data analysis is *always* underpinned by theoretical assumptions, and these assumptions need to be acknowledged and reflected on.



9) Reflexivity, the researchers' insight into, and articulation of, their generative role in research, is key to good quality analysis. Researchers *must* strive to “own their perspectives” (Elliott et al., 1999).

10) Data analysis is conceptualized an art not a science;<sup>8</sup> creativity is central to the process, within a framework of rigor.

This list places researcher subjectivity front and center in reflexive TA. We view researcher subjectivity, and the aligned practice of reflexivity, as the *key* to successful reflexive TA – hence the label *reflexive*. We refer here to a “deep” process of reflexive interrogation of researcher assumptions and practice, rather than a simple listing of identity or experience categories when reporting research (for an example, see Trainor & Bundon, 2020).

Coding, for example, is a process not of simple identification, but of *interpretation* – and researcher subjectivity fuels this process. Good coding (coding that is more complex and nuanced) is often the result of a deep and prolonged engagement with the data; codes can and should evolve in an organic way over the coding process, as insight shifts and changes. Individual codes can expand and contract in scope, be collapsed together with other codes, split into two or more codes, and coding labels can be refined. The point of this organic coding process is precisely to capture the researcher's developing and deepening interpretation of their data. Even at the endpoint of coding, things are still *provisional*. This

---

<sup>8</sup> Although definitions of *art* and *science* are variable and contested, here we evoke a (naïve) realist positivist empiricism with our demarcation of science.

organic process makes the use of a codebook to direct data coding *incompatible* with reflexive TA. A codebook does not allow for this type of data engagement, as it can delimit coding at the start of the analytic process (particularly the more fixed codebooks preferred by coding reliability practitioners). There is also little sense in developing a codebook *after* coding stops (and then re-coding the data using this codebook), because there is no fixed endpoint for coding, any further engagement with the data could lead to new insights (Trainor & Bundon, 2020, illustrate this point nicely).<sup>9</sup>

Themes, like codes, are understood as the *output* of the analysis; the “identification” of themes very early on risks underdeveloped themes and analytic foreclosure, where analysis stops at the level of superficial findings (Connelly & Peltzer, 2016). Themes, developed *from* codes, are constructed at the intersection of the data, the researcher’s subjectivity, theoretical and conceptual understanding, and training and experience. A dataset does not “hold” a single TA analysis within it. Multiple analyses are possible, but the researcher needs to decide on and develop the *particular* themes that work best for their project – recognizing that the aims and purpose of the analysis, and its theoretical and philosophical underpinnings, will delimit these possibilities to some extent. Existing theories, concepts and knowledge are part of the reflexive TA researcher’s set of resources for analyzing the data. How much each contributes during the analysis process depends on

---

<sup>9</sup> This organic and developing coding process still requires systematic tracking and record keeping. For practical advice on ways of tracking your coding process in reflexive TA, see Braun and Clarke (2012), and Trainor and Bundon (2020) for a richly illustrated example.

where on the inductive-deductive spectrum of reflexive TA an analysis sits. Even in a *deductive* or theory driven orientation, these serve to *guide* data coding and the exploration and determination of final themes for the analysis, rather than provide a predetermined structure to code the data within or test the data against.

### **Design Thinking for Reflexive Thematic Analysis**

Design thinking is needed not just for coherent research, but at many points of research assessment, such as ethics review, research proposals, or funding applications. In these contexts, the researcher lays out what they intend to do, with justification of their design decisions. There is no single starting point for, and route through, research design for reflexive TA. Sometimes, meta-theoretical philosophical assumptions and political commitments comprise one point of departure – these are frequent starting points in feminist and other politically-oriented research (see Braun & Clarke, 2013). Research *questions* also constitute a common starting point. Using research question as a starting point for design means that the question should guide the choice of methods of data collection and analysis, and participant/dataset selection strategies, and the location of the research in relation to specific philosophical meta-, methodological and explanatory theories. In practice, more pragmatic and indeed emotional considerations often come to the fore in research design – such as a student researcher choosing an analytic orientation because it is one they have used before, so feel somewhat confident using it, or it is the preferred approach of their academic advisor. Such pragmatic starting points do not necessarily lead to poor design, as long as design is thoughtfully considered, and the overall design is conceptually *coherent*.

This key principle of *design coherence* (Braun & Clarke, 2013; Willig, 2013), or methodological integrity (Levitt et al., 2017), is very important in TA research, because there are few *inherent* limits or prescriptions in research design for TA. In general, as well as having the theoretical flexibility to be used within a wide range of philosophical meta-theoretical, methodological, explanatory and political/ideological frameworks, TA can be used to address a wide range of research questions, analyze almost any type of data, and analyze smaller and larger datasets, collected from participant groups/datasets that are more homogenous or heterogeneous (King & Brooks, 2017). We start our discussion of design thinking for *reflexive* TA with research questions, then move to data collection methods and sources, participant group/dataset constitution and size, quality standards and practices, an end with an overview of reporting standards for reflexive TA.

### **Research Questions**

Research is guided by a question that captures *what* it is the researcher is trying to understand through their data analysis. Qualitative researchers are interested in understanding a diverse range of phenomena; these can be clustered into different “types” of questions (see Braun & Clarke, 2013). Reflexive TA can address most of these types of research question – see Table 1. If the “essence” of what it is a researcher seeks to understand fits within one of these research question types, then reflexive TA is likely a method that will suit their research, as long as it is used within a conceptually coherent design. The types of questions reflexive TA cannot address are those that require technical understanding of language *practice* and/or narrative structure – these are associated with some types of discursive psychological (e.g., Wiggins, 2017), conversation analytic (e.g.,

Schegloff, 2007) and narrative (e.g., Reissman, 2007) approaches; those approaches are best suited for addressing such questions.

[Insert Table 1 about here]

Questions centered on exploring participants' *experiences* and sense-making (variously described as understandings, perceptions, motivations, needs and views), seem to be of most interest to psychologists. Other questions focused on experiential phenomena include those concerned with understanding people's behaviors or practices (the things they *do*), and their sense-making around these, the factors and processes that shape and influence particular phenomena, and the rules and norms that regulate and govern human behavior or practices. *Constructionist* research questions typically interrogate meaning making in the social (and psychosocial) world. They often center on the social construction of reality, and the meaning-frameworks or discourses that surround and constitute the phenomena of interest, and the implications of these (Gergen, 2015).

One important thing to note about research questions in TA research: in some cases, they might be (more or less) fixed from the outset, and strictly adhered to – this is particularly the case in some forms of applied research and in more (post)positivist TA. In contrast, research questions in reflexive TA more commonly *evolve* throughout the course of the research. The initial research question(s) can be quite open and constitute a “starting point” that might become more focused or expand or even shift in focus, as data collection and analysis progresses. They are a starting point for, but are not necessarily the endpoint of, the analysis. Reflexive TA involves a “dialogue” between the interpretation of patterned meaning and the research question. Honing and refining your research question are not

indicators of poor-quality practice, of poor design, but of a process through which deeper insight has been generated.

### **Methods for Data Collection**

There are few in-built restrictions around data collection methods or sources in reflexive TA research. A wide range of data sources have been used in published TA research, including everything from more conventional and extensively used methods such as interviews (e.g., Robinson-Wood et al., 2020) and focus groups (e.g., Tebbe et al., 2018), to other self-report techniques such as open-ended/qualitative surveys responses (e.g., Blackie et al., 2020) and solicited diaries (e.g., Schnur et al., 2009). From innovative and creative methods such as story completion (e.g., Jennings et al., 2019) and visual methods (e.g., Devine-Wright & Devine-Wright, 2009), with forms of reflexive TA specifically developed for the analysis of imagery (e.g., Gleeson, 2011), to “naturalistic” and pre-existing data sources such as psychotherapy sessions (e.g., Willcox et al., 2019), online forum posts (e.g., Fletcher & StGeorge, 2011), and political speeches (e.g., Pilecki, 2017). Analysis can be conducted across more than one *different* data type – such as interview and survey data – (with a clear rationale). The theoretical flexibility of reflexive TA means that it can be incorporated into ethnographic designs (e.g., Devaney et al., 2018) and participatory methodologies such as memory work (e.g., Delgado-Infante & Ofreneo, 2014). There is a wide range of research designs that sit within a “community-located” model, from those “community based and located,” which involve the community, but are researcher driven and directed, to those which are more fully “participatory,” and involve participants as co-researchers and a “power-sharing” model between researchers and community participants (see Coughlin et al., 2017). TA can be used within both of these broader models. Its relative

accessibility (both in terms of procedures and outputs), and the previously noted potential to side-step “methodological horrors” (Brown & Locke, 2017, p. 452) – here knowingly, and for pragmatic and political purposes (e.g., facilitating community members contributing to the analysis) – means it is particularly well suited to power-sharing participatory methodologies (e.g., Rowley et al., 2020).

Data *quality* is another important design consideration for reflexive TA, as good quality analysis depends on having good *quality* data (Connelly & Peltzer, 2016), even more than having a sufficient quantity of data. Data should *ideally* be rich, nuanced, complex and detailed. Connelly and Peltzer highlighted “at-surface interviewing” (p. 53) as one reason for poor quality data, by which they meant little to no attention given to prompts and probes and the relationship between researcher and participant. Data quality is an important consideration *before* analysis begins. It is important to consider the fit between data collection methods and the research question, theoretical frameworks, analytic orientations and, for participant generated data, the characteristics and needs of participant group. It is also important to consider *how* methods will be used. For example, coding reliability researchers prioritize a relatively structured and consistent approach to interviews – asking the same questions in the same order to facilitate the determination of “data saturation” and a more structured coding approach, and consonant with the (post)positivist conceptual underpinnings of coding reliability TA (e.g., Guest et al., 2006). By contrast, reflexive TA, in keeping with its qualitative sensibility, prioritizes a more flexible and fluid approach to

interviewing that more closely resembles the “messier” flow of real-world conversation:<sup>10</sup> questions and topics are carefully considered but the interview centers the interaction and co-construction of meaning between researcher and participant; there is considerable scope for the researcher to be spontaneously responsive to the participant’s unfolding account. The goal is to be “on target while hanging loose” (Rubin & Rubin, 1995, p. 42), gaining an in-depth exploration of each participant’s story, not a uniformly-structured account.

If using interactive methods of data collection, such as interviews and focus groups, reviewing transcripts of the initial interviews or focus groups is vital to check they are generating rich “on target” data. Inexperienced researchers should ask mentors or advisors for feedback on this – interviewing and focus group moderation are skilled activities that do not come “naturally” to most (see Braun & Clarke, 2013, for some suggestions; Connelly & Peltzer, 2016, usefully provide examples of transcripts of interviews with and without sufficient probing). If using non-interactive data collection tools like qualitative surveys, solicited diaries, vignettes and story completions (see Braun et al., 2017), piloting is crucial to assess data “fit” and “quality.” With such methods, richness is often assessed *across* a dataset, as well as within each data items. It is important to *design in* such reviewing or piloting into the research design and timetable.

---

<sup>10</sup> Conversation analysts (e.g., Schegloff, 2007) would point to the patterning and structure of real-world interactions; our emphasis on looseness or messiness here does not negate this aspect.



## Participant Group/Dataset Selection and Constitution

Another important design consideration is the selection (or generation or construction, depending on the conceptualization of the research) of the dataset, whether through the recruitment of participants into a project, the selection of social media posts on a topic of interest, or one of a myriad of other ways. In quantitative and (post)positivist research terms, this constitutes a “sample” – a framing that remains pervasive in qualitative research (including some of our own). Conceptualizing data as a sample reflects the idea that relevant information has been selected from the total *possible* sources (the “population”), and this *sample* is used to address the research question. Here, we try to avoid this simple representational inference, as we discuss participant group/dataset selection strategy, size of the participant group/dataset, and (size) justifications (where authors use sample, we retain their language). Before we do, we note that the emphasis in TA is on *themes*, patterns of meaning *across* cases, rather than on meaning *within* individual cases. Therefore, the participant group/dataset needs to be large *enough* to justify the claims regarding patterned meaning. This contrasts with more idiographic approaches such as narrative analysis, where the specific characteristics of a small number of cases, and even one case (e.g., Josselson, 2009), are analyzed in depth and detail.<sup>11</sup>

---

<sup>11</sup> Reflexive TA *has* been used in case study research, where the focus is on a small number of cases, or even one case (e.g., see Cedervall & Åberg, 2010; Manago, 2013). Furthermore, some researchers have combined reflexive TA with narrative methodologies and procedures to produce distinct

With the caveat that TA is typically concerned with meaning across data items, there are no *particular* participant group/dataset selection requirements for reflexive TA research, neither regarding how many data items, nor how the participant group/dataset is *selected* – what is often known as the “sampling” method or strategy. Robinson’s (2014) four-step “pan-paradigmatic” guide to sampling provides a useful starting point for thinking about the different aspects of selecting and generating participant groups/datasets in reflexive TA:

- Define a “sampling universe” using *inclusion* criteria (attributes that participants or data items *must* possess) and *exclusion* criteria (attributes that *disqualify*). The more specific the criteria, the more homogenous (in certain ways) the sampling universe likely becomes. For example, moving from “people who do not have children” to “people who are childfree by choice” to “men who are childfree by choice” narrows the pool in particular ways.
- Determine a sample size (or size *range*) by reflecting on what is *ideal* (consonant with purpose of research, analytic orientation, theoretical underpinnings) and what is *practical* (e.g., time, resources, or norms or expectations of the local – institutional, research field – context).
- Develop a sampling *strategy* for selecting items or participants for inclusion.
- Source the sample by recruiting participants or selecting items from the sampling universe.

---

“hybrid” methods that are concerned with both narrative structure and “across case” patterning of meaning (e.g., Palomäki et al., 2013; Ronkainen et al., 2016).

### ***Strategy for Selecting Participants/Data Items***

In qualitative research conducted within qualitative paradigms, the aim of research, and thus participant/data item selection, is generally to capture some of the range and diversity of meaning within the “population,” rather than providing some “quantified representation” of it (Gaskell, 2000). And, to allow for an in-depth exploration of the research question(s), which maximizes the opportunity for “transferability” of results (Spencer et al., 2003). What are known as *convenience* and *purposive* sampling (Patton, 2015; Sandelowski, 1995) are seemingly the most common participant/data item selection strategies in TA research. *Convenience sampling* involves selecting “cases” (participants or data items) that are *easily* accessed by the researcher. In practice, this often means advertising a project, and the participant group constitutes whoever happens to respond. Psychologists have commonly – and problematically – recruited psychology undergraduates through research participation schemes, another form of a convenience sample (Arnett, 2008). Convenience sampling is often considered the least rigorous and justifiable participant selection method (Sandelowski, 1995) – especially when the wider group of interest is not *specifically* psychology students. However, the critique has not dented its popularity as a participant/data item selection strategy and there are ways to facilitate diversity within a convenience strategy (e.g., where and how the research is advertised). *Purposive sampling* can involve deliberately selecting “information rich” cases (Patton, 2015) that have the potential to *maximize understanding* of the phenomena under investigation. Purposively selected participant groups/datasets can be homogenous or heterogeneous in constitution; deliberately seeking diversity is referred to as maximum variation (Sandelowski, 1995) or heterogeneity (Fassinger, 2005) sampling (see Onwuegbuzie & Leech, 2007). Participant group/dataset selection strategies can be

combined and can blur; real-world practice is often not particularly like textbook descriptions. There is no ideal “sampling” strategy for reflexive TA: what matters most, from a reflexive TA design coherence standpoint, is that researchers understand what their strategy is, and why they have chosen it, its strengths and limitations, and they can articulate how and why it provides a dataset to meaningfully address their research questions. This connects to participant group/dataset size.

### ***Participant Group/Dataset Size***

How large should a participant group/dataset be? This is a tricky question not just for TA, but qualitative research more broadly (e.g., see Malterud et al., 2016; Morse, 2000; Sim et al., 2018). Determining a participant group/dataset size for TA is not as simple as identifying the “correct number” of participants or data items – for a start there is data *type* to consider, and the related consideration of the “volume” and *richness* of each data item, as well as considerations of homogeneity and heterogeneity.

Larger participant groups/datasets can be useful when the scope of the study is relatively broad, the topic is potentially “difficult to grab” (Morse, 2000, p. 4) and/or sensitive for participants, and there is considerable diversity within the wider group of interest. When working with smaller participant groups/datasets (e.g., 10 interviews or fewer), homogeneity (which could, for example, be based on demographics, experience, location, and many other things; Robinson 2014) may help to facilitate theme development. But even inclusion and exclusion criteria expressly designed to produce a homogenous participant group/dataset cannot guarantee homogeneity of sense-making. Then there is *also* the purpose and the context of the study (see Morse, 2000), as well as pragmatic considerations such as norms around publishability (Dworkin, 2012). Context is part of what

determines what is manageable – the analysis should not just be completable but done *well*. In the context of a study with a concrete deadline, there is a need to balance a dataset with enough breadth and depth to give validity to TA, with time to analyze the data meaningfully before the deadline. Researchers should not constantly feel overwhelmed and like they are “drowning in data” or fail to do it justice (though this *can* happen, Braun & Clarke, 2020). TA research is rarely conducted under “perfect” conditions, with all the resources, time and skills needed to execute a study in a textbook manner, so participant group/dataset size decisions invariable involve compromises born of negotiating competing priorities – pragmatic/practical and methodological/theoretical.

General guidelines around participant group/dataset selection in qualitative research are useful for reflexive TA research (e.g., Malterud et al., 2016; Morse, 2000), but, reflecting the fuzzy nature of the qualitative research terrain (Madill & Gough, 2008), there are no widely agreed on and precise criteria for determining participant group/dataset *size*. Furthermore, “what constitutes an adequate sample size to meet a study’s aims is one that is necessarily a process of ongoing interpretation by the researcher. It is an iterative, context-dependent decision” (Sim et al., 2018, p. 630). Two formulae that appear to offer more precise criteria for determining participant group/dataset size, even in advance of analysis, are saturation and statistical models. As these have been widely discussed in the TA methodological literature, we now explore the relevance of both of these for reflexive TA.

### ***Saturation***

“The data were saturated” is one of the most ubiquitous participant group/dataset size rationales in TA research, and is often taken to be so self-explanatory it is not even

defined (Braun & Clarke, 2019b). The concept of “data saturation” or “information redundancy” seems to have evolved from “theoretical saturation” (Glaser & Strauss, 1967) – tightly defined and connected to the methodology of grounded theory (O’Reilly & Parker, 2012). Theoretical saturation is inextricably linked to the practice of theoretical sampling and concurrent data collection and analysis in grounded theory (Morse, 2015; O’Reilly & Parker, 2012). Data saturation (Fusch & Ness, 2015), and its variants of code- (Hennink et al., 2016), theme- (Guest et al., 2006), and meaning- (Hennink et al., 2016) saturation, have now become embedded as a “gold standard” (Guest et al., 2006, p. 60) for dataset generation in TA research. However, its use as a generic measure of dataset adequacy is problematic because it is not philosophically and methodologically consistent with reflexive TA.<sup>12</sup> For Morse (1995), data adequacy was operationalized as “collecting data until no new information is obtained” (p. 147); this notion of “no new” is common across different varieties of saturation (data, code or theme). However, not only is the definition and meaning of claimed saturation often fuzzy, but precisely *how* saturation might have been achieved is commonly not discussed (Bowen, 2008). A claim of saturation – whether defined and/or explained in any way or not – is often provided as the rationale for stopping data collection in TA studies (e.g., Grabe et al., 2015; Staneva & Wittkowski, 2013). It is often positioned as separate from, as preceding, data analysis (Saunders et al., 2017).

---

<sup>12</sup> The concept or use of saturation *is* referenced in many general qualitative quality guidelines and criteria (e.g., Levitt et al., 2018; Morrow, 2005); the best make clear that criteria such as saturation are *not* universally applicable or theoretically neutral (e.g., Levitt et al., 2018).

Claims of data saturation often work (whether knowingly or unintentionally) as a rhetorically robust rationale for *sample size* in TA research. Most uses of this term seem to evoke the previously noted idea that the researcher gets to a point in data collection where no new insights will be generated by gathering additional data. This is problematic from a reflexive TA perspective, because it effectively positions the researcher's task as discovery – as recognizing themes that are waiting to be discovered – which is not how themes or the research process are conceptualized in reflexive TA (Braun & Clarke, 2019b).<sup>13</sup> Within a conceptualization of qualitative research as a reflexive process of knowledge *generation* or *construction*, rather than discovery, there is always the potential for new understandings (Mason, 2010), developed through ongoing data engagement, or through reading the data from different perspectives (nicely illustrated by Ho et al., 2017). The notion that it is possible to “saturate” stops making sense (Malterud et al., 2016) if we envisage analysis as a process of analytic insight developed *through* engagement with the data and in line with our positionings (at the time of analysis).

In the last decade or so, concern about a lack of concrete guidance for determining saturation has led several authors to “operationalize” saturation, and try to determine how many interviews (or focus groups) are enough to reach saturation, or a certain level of saturation, in TA research (e.g., Guest et al., 2006; Hancock et al., 2016; Hagaman & Wutich, 2017; Hennink et al., 2016). Such guidelines have been taken up and used as rationales for

---

<sup>13</sup> See also Charmaz (2006), for a more constructivist reworking of saturation within grounded theory.

TA participant group/dataset sizes, but they are *not* suitable as guidance for reflexive TA, because they contain some (sometimes unacknowledged) assumptions that are at odds with the conceptual bases, values and practice of reflexive TA (Braun & Clarke, 2019b). These are evident in various ways.

Most use a coding reliability or codebook version of TA – although rarely acknowledged as a *particular* iteration of a broader approach – and evidence a realist ontology and (post)positivist research values. Data collection is often rather structured and standardized, the data generated rather concrete, with an applied focus. Within a “discovery” mode, themes are implicitly conceptualized as entities that pre-exist the analysis, and that the researcher elicits from the participants or uncovers through the analytic process (Saunders et al., 2017). Themes are often topic summaries and analytic inputs, reflective of interview questions. Frequency – rather than importance and meaningfulness in relation to the research question – is typically the primary or sole determinant of a theme/code; the rationale for this is rarely explicated.

Although we think it is better for reflexive TA researchers to avoid the concept of data saturation altogether (see Braun & Clarke, 2019b), we recognize that this is not always pragmatically possible. What is vital for quality practice in these instances, is demonstrating *knowing* use, where exactly what saturations means is defined and the researcher is clear about how precisely such saturation was determined.

### ***Statistical Formula***

Fugard and Potts (2015) developed a *quantitative* tool for *prospectively* determining sample size in TA. The tool, which requires researchers to determine the expected population theme prevalence of the *least* prevalent theme, the number of desired instances



of the theme, and the power of the study, attracted critical commentaries from several qualitative researchers (e.g., Braun & Clarke, 2016; Hammersley, 2015). Other equally problematic formulae and tools for prospectively determining sample size or data saturation have also been published (e.g., Galvin, 2015; Trans et al., 2017). Many different aspects of these tools make them conceptually *incompatible* with reflexive TA. Yet there is a risk that those who do not realize the fundamental conceptual differences between different types of TA might be tempted to assess TA research design with such tools (Hammersley, 2015). Therefore, we wish to be very clear that we *do not recommend* the use of any of these statistical tools to determine or imagine the “correct” dataset size for a study that uses on reflexive TA.

### ***Determining and Justifying Participant Group/Dataset Size in Reflexive TA***

There is no failsafe way to justify participant group/dataset size in reflexive TA. We recommend avoiding the (post)positivist temptation to ground size decisions around some idealized notion of “generalizability,” implicitly “buying into” the notion that bigger is better and statistical generalizability is an ideal for all research (but see Smith, 2017, for an important discussion of ways generalizability can be reconfigured in qualitative research). Instead, in terms of a conceptual model for estimating participant group/dataset size (in advance of data analysis, for purposes like ethics review), we find Malterud et al.’s (2016) notion of information power useful. Rather than precise calculations, it invites the researcher to reflect on the “information richness” of the dataset and how that meshes with the aims and requirements of the study – different purposes mandate different approaches to sample size. Using this reflection, a study with:

- A broad aim;

- Non-specific or few inclusion criteria;
- A more inductive and exploratory approach;
- Thinner data generated from each participant or data item;
- Analysis focused *across* a dataset; and
- Analysis conducted by a novice researcher

would generally require a larger participant group/dataset to have adequate information power – that is, to be able to say something (qualitatively) meaningful. And, contrastingly, a study with a narrower aim, a more specific population/dataset focus, perhaps a more deductive approach, and with “thicker” or richer individual data items, would generally require fewer data items. Such aspects should not be seen as operating independently and summatively, but as potentially interacting (Sim et al., 2018). This makes some in situ assessment of the dataset – such as during data collection; following familiarization or even coding – more important than clear determination prior to the research starting.

Another concept that may be useful in thinking about dataset *adequacy* is “theoretical sufficiency.” Developed by grounded theorist Ian Dey (1999) – who described saturation as an “unfortunate metaphor” (p. 257), suggesting completeness of understanding and a fixed point – theoretical sufficiency is intended to capture the notion that data collection stops when the researcher has reached a sufficient or adequate depth of understanding to build a theory. Similar ways of framing this are “conceptual density” or “conceptual depth” (Nelson, 2016). Regardless of whether “building theory” is a goal, these concepts emphasize meaning-richness as key to the validity of the (size of the) dataset. For us, informational or meaning sufficiency seems a useful concept for the point at which to stop data collection in TA, and it is only something that can reflexively be determined in situ.

Despite our acknowledgement that it is difficult to *determine* participant group/dataset size *in advance of* data collection, there often *is* a pragmatic need to provide some indication of participant group/dataset size, to meet the requirements of institutional review boards, research degree committees, funding bodies, and because of the practical need to plan time and resources. We suggest researchers provide a participant group/dataset size *range*, with the final participant group/dataset size determined during data collection or after early phases of the analytic process (see Braun & Clarke, 2013, for some suggestions for student projects).

### **Thinking Ethically for TA Research**

Ethics are a key requirement of all research design and practice, and are both *procedural* (what we do in relation to participants) and more socio-political, related to the politics of research, the power relationship between researcher and participant, and the researcher's values. Research ethics are codified within ethics codes like those from the American Psychological Association (2017) and New Zealand Psychological Society (2012), and applied through *institutional* ethical review. Ethical guidance may change in relation to particular *modes* of inquiry (e.g., online research; Association of Internet Researchers, 2012), research participants (e.g., Indigenous populations [e.g., Smith, 2013]) or children [e.g., Shaw et al., 2011]), or collaborating organizations (some – especially health and medical organizations – may have their own ethical review processes and requirements). A key point to emphasize is that ethics codes represent *minimal* requirements. The British Psychological Society noted (2009) that “thinking is *not* optional” (p. 5; our emphasis), and “no code can replace the need for psychologists to use their professional and ethical judgement” (p. 4).

The use of TA as analytic approach requires little ethical discussion *per se*, but *qualitative* research in general, and *especially* within qualitative paradigms, raises important ethical considerations. Familiarization with the discussions of ethics in the context of *qualitative* research – including the emotional impacts of data on researchers – is important (e.g., Brinkmann & Kvale, 2017; Denzin & Giardina, 2016; McClelland, 2017; Miller et al., 2012). Thinking more broadly about ethics, the design and conduct of qualitative research often involves complex and “fuzzy” ethical and moral considerations such as issues of difference, power (Karnieli-Miller et al., 2009) and control, and how we relate to, and represent, participants (Fine, 1992). While none of these considerations are a *necessary* feature of TA research specifically, we encourage TA researchers to pursue a complex and sophisticated reflexive approach to qualitative research and research ethics. To exemplify *best practice* with regard to relating to and representing participants, especially with regard to questions of difference, and to conduct research that is genuinely inclusive, culturally sensitive and politically astute.

### **Quality Standards and Practice**

The final area we consider for design thinking is quality, which intersects profoundly with, and brings our discussion back to, conceptual thinking. The quality-assurance strategies we discuss here are informed by the theoretical assumptions and values of a Big Q qualitative paradigm (Braun & Clarke, 2013) and focus on encouraging reflection, rigor, a systematic and thorough approach, and even greater depth of engagement, rather than on determining the “accuracy” of coding or theme identification.

Judgements around quality relate to both process and outcome. We focus both on the quality practices that researchers incorporate into their research designs, and on the

quality standards and criteria researchers strive to adhere to, and reviewers, editors and examiners use to assess the quality of qualitative research. It is often assumed that there are universal quality criteria that apply to all forms of qualitative research – we mentioned previously the problematic assumption that coding reliability measures are relevant to all forms of TA. This assumption is typically underpinned by a (limited) conceptualization of qualitative research as commensurate with the study of subjective experience and (post)positivist research values. For example, many quality criteria and standards include “member checking” or “participant validation” as a form of credibility check (e.g., Elliott et al., 1999; Morrow, 2005) – in some cases *without* acknowledgement that this quality practice is not conceptually coherent with *all* forms of qualitative research (Reicher, 2000), or consideration of the practical and pragmatic challenges of implementing this practice (Braun & Clarke, 2013). Thus, quality practices and criteria are another aspect of research design that requires conceptual thinking – researchers should reflect on the theoretical assumptions embedded in particular standards and practices to determine whether they are coherent with their research design and use of reflexive TA.

Even though there are aspects of some of these quality criteria that do not translate well, or at all, to reflexive TA, we nonetheless find Elliott et al.’s (1999) checklist-criteria publishability guidelines, Yardley’s (2015) much looser flexible principles, intended to be applicable to a wide range of qualitative research methods and approaches, Tracy’s (2010) eight “big tent” flexible criteria, and the APA journal reporting standards (Levitt et al., 2018) useful. For reflexive TA – and indeed qualitative research more widely, such “criteria” are not designed to be applied in a rigid way, but as *flexible resources*, open to reinterpretation, for thinking about quality in general, and the appropriate quality standards that should apply to a particular piece of reflexive TA research (Sparkes & Smith, 2009). Elsewhere, we

have provided a 15 point checklist for researchers to reflect on the quality and rigor of their reflexive TA practice (Braun & Clarke, 2006), and guidelines for reviewers and editors evaluating reflexive TA research for publication (Braun & Clarke, 2020).

For reflexive TA, we stress the importance of a deep, engaged, and critically-open reflexivity *throughout* the research process. Of the researcher reflecting on, trying to understand, and interrogating: their values and personal positioning; their assumptions and expectations about the topic of their research; and, in designs with participants, their relationship to and with participants (Wilkinson, 1988, termed this *personal reflexivity*); their design and methodological choices (*functional reflexivity*); and their disciplinary location and standpoint (*disciplinary reflexivity*). And, indeed, how *all* of these intersect with and shape the research process and knowledge produced. In more politically-oriented research, conceptualizations of reflexivity that highlight the power dynamics of research are also important (e.g., Ramazanoğlu & Holland, 2002). Reflexivity to us is best conceptualized as a meshed-in mode of (Big Q) research *practice*; if this is unfamiliar, see Finlay and Gough (2003) for an accessible starting point, and Trainor and Bundon (2020) for an example when doing reflexive TA.

Researcher reflexivity is highlighted in many quality standards and criteria – Elliott et al. (1999), for example, included “owning one’s perspective” in their publishability guidelines, emphasizing the need for researchers to specify their theoretical orientations and personal expectations. In acknowledgement of the incompleteness of reflexivity – full insight is rarely possible – and the multiplicity of our perspectives, we reframe this as researchers *striving* to “own”, in the sense of acknowledging and taking responsibility for, their *perspectives*. Some of these reflections should ideally be included in research

reporting, to render visible to the reader some aspects of the context of the research. However, we recognize, as do others (Levitt et al., 2018), that tight journal word limits constrain the reporting of qualitative research in various ways, and reflexivity is often something that is “sacrificed” to remain within such limits. Reflexivity can also be stylistically confronting for those schooled in “scientific” writing practices, as it brings the voice of the individual researcher into the text.

One quality practice we strongly encourage is keeping a reflexive journal throughout the research process for recording the researcher’s reflections and insights, but also to use the practice of writing as a tool for deepening reflexivity. As discussed, a concern for quality practice is embedded throughout the process of reflexive TA – procedures like organic and open-ended coding, theme review and refinement, and the recursivity of the phases, are intended to sensitize the researcher to the need for prolonged and deep engagement with their data to produce a meaningful, and useful, analysis that exceeds the superficial and the obvious.

### **Best Practice for Reporting Reflexive TA**

To aid discussions of quality for reflexive TA, we now provide a brief discussion of best practice for reporting reflexive TA. This captures how reflexive TA research, in which the researcher has thoroughly engaged with practices of conceptual and design thinking,

would *ideally* be reported.<sup>14</sup> Our aim is to support: a) researchers in producing written reports of reflexive TA to the highest standards; and b) reviewers, editors and examiners in *appropriately assessing* written reports of reflexive TA. Overall, we encourage writing styles that bring the “voice” of the individual researcher into the text – such as the use of the first person in the methodology section. For the purpose of this section, we assume that any claimed conceptual or other positions would be aligned with actual reported practice and analysis.

### ***Introduction***

The introductory section of the paper should provide *contextualisation* and a *rationale* for the research – referencing existing research, relevant theory, and the wider context (e.g., social, cultural, policy, political, media, health) – and may or may not include a conventional literature review. The aim of any synthesis of existing literature is not necessarily to identify “gaps” in knowledge, but to contextualise the current research; for this reason we recommend *introduction* over *literature* review if a section heading is required. A *research question* appropriate to reflexive TA should be clearly articulated, and conceptually aligned with the form of TA reported in the analysis (it may be useful to discuss the refining of an initially broader research question).

---

<sup>14</sup> In the current context at least, these best practice guidelines remain partly aspirational. They diverge in many ways from accepted reporting standards in the wider discipline, which remain oriented to (post)positivist norms and values (see also Levitt et al., 2018).



## **Methodology**

We prefer the heading *methodology* over *method*, to signal a theoretically embedded and reflexive account of the research process. The *conceptual underpinnings of the research* – ontological and epistemological assumptions; any methodological, explanatory and political/ideological theory informing the data analysis – should be clearly identified, and *how* theory informed data analysis should be discussed. The particular *orientation to reflexive TA* – inductive <> deductive; semantic <> latent meaning – should be explicit and explained in a situated manner, by which we mean specific to the project, rather than generically. *Reflexivity* should be evident, through writing style, discussion of reflexive practices (e.g., journaling) throughout the process, and, if appropriate, consideration of the researcher's personal positioning in relation to the topic, and the participants (see Trainor & Bundon, 2020). This latter connects to wider qualitative research *ethics*,<sup>15</sup> including around representation. The participant group/dataset should be clearly and richly situated (without compromising anonymity), and some explanation of, and rationale for, the constitution and size of the participant group/dataset should be provided – without reference to saturation or statistical models (see Braun & Clarke, 2019b). Where appropriate, how the choice of method/data source shaped the research process and the knowledge produced might be considered. The *process of analysis* should be described in a situated and specific – not generic – way, and using the most up to date terminology (see Braun & Clarke, 2019a). How

---

<sup>15</sup> Ethical discussion would also include more standard disciplinary processes, such as formal review processes.

various *quality measures* appropriate to the reflexive TA process (e.g., theme review and refinement) were practiced should be included. With more than one author, *how* each author contributed to the analysis should be discussed.

### ***Analysis***

Our current preferred heading for the *results/findings* section is *analysis* because it avoids evoking both discovery and finality. This section – which often includes the *discussion* in the sense that the analytic observations are contextualised in relation to existing research and theory in the reporting of themes – should start with a brief overview of the analysis to come (a figure or table, or even a simple description or list; this can also be used to convey the relationship between themes). The analytic narrative should explain the meaning and significance of the data, avoiding both paraphrasing and “arguing with” the data. Theme frequency counts should be avoided, especially as a rationale for the analytic content and structure, because reflexive TA does not equate frequency with importance. A large number of participants may say or write things that are not relevant to the research question, while a small number may say or write things that are crucial. Furthermore, the quantification of qualitative data, even in the form of simple frequency counts, is often far from straightforward because data collection is not typically rigidly structured and systematised, with precise comparability across participants or data items.

The themes should form a coherent overall “story” about the data, presented in the order that best tells the overall story. We generally recommend discussing two to six themes (including any subthemes) in any single report; any more themes suggests an *underdeveloped* analysis. Use subthemes judiciously; a overly elaborate thematic structure similarly suggests underdevelopment (see Connelly & Peltzer, 2016). Each theme should be

rich, complex and multifaceted (i.e., consist of more than one analytic observation), with a distinct core meaning or *central organising concept* (themes should *not* be topic summaries). There should be little or no overlap (boundary blurring) between themes. Each *theme name* should convey something of the “essence” of each theme; one-word names should be avoided. The detailed discussion of *each theme* should include a balance of data extracts and analytic narrative (interpretation), regardless of a more illustrative or more analytic use of data extracts.<sup>16</sup> Vivid and compelling data extracts should be drawn from *across* the dataset to evidence patterning; presented data extracts should “fit” (or “evidence”) the analytic claims. Presentation of data extracts in tables should be avoided.

### **Conclusion**

In the final *conclusion* (or sometimes *discussion*) section, analytic conclusions and implications should arise from, or cut across, the themes, reflecting that the themes themselves are not analytic conclusions – theme-by-theme contextualisation should be avoided (it is often a sign that the results and discussion should have been combined). The section should also include evaluative reflection on how design choices shaped (and possibly delimited) the knowledge produced, as well as wider reflection on the limitations of the study, and the overall analysis; claims about a lack of statistical probabilistic generalisability should be avoided.

---

<sup>16</sup> An analytic use of data involves a detailed analysis of the specific features of a particular data extract (see Braun & Clarke, 2013), rather than data extracts being used more generically to illustrate analytic observations.

## Summary

In this paper, we have emphasized that TA is not a single approach with a single theoretical foundation; we outlined three different schools of TA, and focused on reflexive TA. Sound reflexive TA practice depends on deep thinking about the conceptual foundations of the research, and effective planning – processes we term *conceptual and design thinking*.

Specifically, we have emphasized the importance of *coherence or fit* in the different aspects that constitute a qualitative project where reflexive TA is used to analyze data; considering these elements should help produce TA research characterized by methodological *integrity* (Levitt et al., 2018). Given the flexibility of reflexive TA, we noted different types of research questions it can be used to address, different data types it works with, and discussed issues related to dataset constitution and size. In particular, we critically discussed the ubiquitous use of (claimed) data saturation as a rationale for TA dataset size, and the use of statistical models for determining dataset size in advance of analysis. Instead, we emphasized the value of thinking critically and reflexively, and in a located way, about the information richness of data as key in decisions about participant group/dataset size. We also noted the importance of ethicality and the use of reflexive journaling to aid quality practice.

## References

- American Psychological Association (2017). *Ethical principles of psychologists and code of conduct*. American Psychological Association. <https://www.apa.org/ethics/code/>.
- American Psychological Association (2020). *Qualitative Psychology*. <https://www.apa.org/pubs/journals/qua/>
- Arbeit, M. R., Hershberg, R. M., Rubin, R. O., DeSouza, L. M., & Lerner, J. V. (2016). "I'm hoping that I can have better relationships": Exploring interpersonal connection for young men. *Qualitative Psychology*, 3(1), 79–97. <https://doi.org/10.1037/qup0000042>
- Arnett, J. J. (2008). The neglected 95%: Why American psychology needs to become less American. *American Psychologist*, 63(7), 602-614. <https://doi.org/10.1037/0003-066X.63.7.602>
- Association of Internet Researchers (2012). *Ethical decision-making and Internet research 2.0: Recommendations from the AoIR ethics working committee*. Association of Internet Researchers. <https://aoir.org/reports/ethics2.pdf>
- Attride-Stirling, J. (2001). Thematic networks: An analytic tool for qualitative research. *Qualitative Research*, 1(3), 385-405. <https://doi.org/10.1177/146879410100100307>
- Awad, G. H., Norwood, C., Taylor, D. S., Martinez, M., McClain, S., Jones, B., Holman, A., & Chapman-Hilliard, C. (2015). Beauty and body image concerns among African American college women. *The Journal of Black Psychology*, 41(6), 540–564. <https://doi.org/10.1177/0095798414550864>

- Blackie, L. E. R., Colgan, J. E. V., McDonald, S., & McLean, K. C. (2020). A qualitative investigation into the cultural master narrative for overcoming trauma and adversity in the United Kingdom. *Qualitative Psychology*. <https://doi.org/10.1037/qap0000163>
- Bond, L. A., Holmes, T. R., Byrne, C., Babchuck, L., & Kirton-Robbins, S. (2008). Movers and shakers: How and why women become and remain engaged in community leadership. *Psychology of Women Quarterly*, *32*, 48-64.  
<https://doi.org/10.1111/j.1471-6402.2007.00406.x>
- Bowen, G. A. (2008). Naturalistic inquiry and the saturation concept: A research note. *Qualitative Research*, *8*(1), 137-152. <https://doi.org/10.1177/1468794107085301>
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)
- Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57-71). American Psychological Association.  
<https://doi.org/10.1037/13620-004>
- Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practice guide for beginners*. Sage.
- Braun, V., & Clarke, V. (2016). (Mis)conceptualising themes, thematic analysis, and other problems with Fugard and Potts' (2015) sample-size tool for thematic analysis.

*International Journal of Social Research Methodology*, 19(6), 739-743.

<https://doi.org/10.1080/13645579.2016.1195588>

Braun, V., & Clarke, V. (2019a). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise & Health*, 11(4), 589-597.

<https://doi.org/10.1080/2159676X.2019.1628806>

Braun, V., & Clarke, V. (2019b). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*. ONLINE FIRST.

<https://doi.org/10.1080/2159676X.2019.1704846>

Braun, V., & Clarke, V. (2020). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*. ONLINE FIRST.

<https://doi.org/10.1080/14780887.2020.1769238>

Braun, V., Clarke, V., & Gray, D. (Eds.), (2017). *Collecting qualitative data: A practical guide to textual, media and virtual techniques*. Cambridge University Press.

<https://doi.org/10.1017/9781107295094>

Brinkmann, S., & Kvale, S. (2017). Ethics in qualitative psychological research. In C. Willig & W. Stainton Rogers (Eds.), *The Sage handbook of qualitative research in psychology* (2<sup>nd</sup> ed., pp. 259-273). Sage.

British Psychological Society (2009). *Code of ethics and conduct*. British Psychological Society. <https://www.bps.org.uk/files/code-ethics-and-conduct-2009pdf>

Brown, S. D., & Locke, A. (2017). Social psychology. In C. Willig & W. Stainton-Rogers (Eds.), *The Sage handbook of qualitative research in psychology* (2nd ed., pp. 417-430). Sage.

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>

Carter, S. M. & Little, M. (2007). Justifying knowledge, justifying method, taking action: Epistemologies, methodologies, and methods in qualitative research. *Qualitative Health Research, 17*(10), 1316-1328. <https://doi.org/10.1177/1049732307306927>

Cavallerio, F., Wadey, R., & Wagstaff, C. R. D. (2020). Understanding overuse injuries in rhythmic gymnastics: A 12-month ethnographic study. *Psychology of Sport and Exercise, 25*, 100-109. <https://doi.org/10.1016/j.psychsport.2016.05.002>.

Cedervall, Y., & Åberg, A. C. (2010). Physical activity and implications on well-being in mild Alzheimer’s disease: A qualitative case study on two men with dementia and their spouses. *Physiotherapy Theory and Practice, 26*(4), 226-239.  
<https://doi.org/10.3109/09593980903423012>.

Chamberlain, K. (2012). Do you really need a methodology? *Qualitative Methods in Psychology Bulletin, 13*, 59-63.

Chamberlain, K., Cain, T., Sheridan, J., & Dupuis, A. (2011). Pluralisms in qualitative research: From multiple methods to integrated methods. *Qualitative Research in Psychology, 8*(2), 151-169. <https://doi.org/10.1080/14780887.2011.572730>



- Champ, F. M., Nesti, M. S., Ronkainen, N. J., Tod, D. A., & Littlewood, M. A. (2020). An exploration of the experiences of elite youth footballers: The impact of organizational culture. *Journal of Applied Sport Psychology, 32*(2), 146-167.  
<https://doi.org/10.1080/10413200.2018.1514429>
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis* (2<sup>nd</sup> ed.). Sage.
- Clarke, V. & Braun, V. (2014). Thematic analysis. In T. Teo (Ed.), *Encyclopaedia of critical psychology* (pp. 1947-1952). Springer. <https://doi.org/10.1007/978-1-4614-5583-7>
- Connelly, L. M., & Peltzer, J. N. (2016). Underdeveloped themes in qualitative research: Relationships with interviews and analysis. *Clinical Nurse Specialist, 30*(1), 51-57.  
<https://doi.org/10.1097/NUR.0000000000000173>
- Coughlin, S. S., Smith, S. A., & Fernandez, M. E. (Eds.), (2017). *Handbook of community-based participatory research*. Oxford University Press.
- Delgado-Infante, M. L., & Ofreneo, M. A. P. (2014). Maintaining a “good girl” position: Young Filipina women constructing sexual agency in first sex within Catholicism. *Feminism & Psychology, 24*(3), 390–407. <https://doi.org/10.1177/0959353514530715>
- Denzin, N. K., & Giardina, M. D. (Eds.), (2016). *Ethical futures in qualitative research: Decolonizing the politics of knowledge*. Routledge. (Original work published in 2007)  
<https://doi.org/10.4324/9781315429090>
- Devaney, D. J., Nesti, M. S., Ronkainen, N. J., Littlewood, M. A. & Richardson, D. J. (2018). Athlete lifestyle support of elite youth cricketers: An ethnography of player concerns

within a national talent development program. *Journal of Applied Sport Psychology*, 30(3), 300-320. <https://doi.org/10.1080/10413200.2017.1386247>

Devine-Wright, H., & Devine-Wright, P. (2009). Social representations of electricity network technologies: Exploring processes of anchoring and objectification through the use of visual research methods. *British Journal of Social Psychology*, 48(2), 357-373. <https://doi.org/10.1348/014466608X349504>

Dey I. (1999). *Grounding grounded theory: Guidelines for qualitative enquiry*. Academic Press.

Dworkin, S. L. (2012). Sample size policy for qualitative studies using in-depth interviews. *Archives of Sexual Behavior*, 41, 1319-1320. <https://doi.org/10.1007/s10508-012-0016-6>

Elliott, R., Fischer, C. T., & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38(3), 215-229. <https://doi.org/10.1348/014466599162782>

Fassinger, R. E. (2005). Paradigms, praxis, problems, and promise: Grounded theory in counseling psychology research. *Journal of Counseling Psychology*, 52(2), 156-166. <https://doi.org/10.1037/0022-0167.52.2.156>

Fine, M. (1992). *Disruptive voices: The possibilities of feminist research*. University of Michigan Press. <https://doi.org/10.3998/mpub.23686>

Finlay, L., & Gough, B. (Eds.), (2003). *Reflexivity: A practical guide for researchers in health and social sciences*. Blackwell.

- Fletcher, R., & StGeorge, J. (2011). Heading into fatherhood— nervously: Support for fathering from online dads. *Qualitative Health Research, 21*(8), 1101–1114.  
<https://doi.org/10.1177/1049732311404903>
- Flick, U. (2014). *An introduction to qualitative research* (5<sup>th</sup> ed.). Sage.
- Fugard, A. J. B., & Potts, H. W. W. (2015). Supporting thinking on sample sizes for thematic analysis: A quantitative tool. *International Journal of Social Research Methodology, 18*(6), 669-684. <https://doi.org/10.1080/13645579.2015.1005453>
- Fusch, P. I., & Ness, L. R. (2015). Are we there yet? Data saturation in qualitative research. *The Qualitative Report, 20*(9), 1408-1416.  
<https://nsuworks.nova.edu/tqr/vol20/iss9/3/>
- Galvin, R. (2015). How many interviews are enough? Do qualitative interviews in building energy consumption research produce reliable knowledge? *Journal of Building Engineering, 1*, 2-12. <https://doi.org/10.1016/j.jobbe.2014.12.001>
- Gaskell, G. (2000). Individual and group interviewing. In M. W. Bauer & G. Gaskell (Eds.), *Qualitative researching with text, image and sound* (pp. 38-56). Sage.
- Gergen, K. J. (2015). *An invitation to social construction* (3rd ed.). Sage.  
<https://dx.doi.org/10.4135/9781473921276>
- Gibbs, G. R. (2007). *Analysing qualitative data*. Sage.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Sociology Press.

- Gleeson, K. (2011). Polytextual thematic analysis for visual data: Pinning down the analytic. In P. Reavey (Ed.), *Visual methods in psychology: Using and interpreting images in qualitative research* (pp. 314-329). Psychology Press.
- Gough, B., & Madill, A. (2012). Subjectivity in psychological research: From problem to prospect. *Psychological Methods, 17*(3), 374-384. <https://doi.org/10.1037/a0029313>
- Gough, B., Seymour-Smith, S., & Matthews, C. R. (2016). Body dissatisfaction, appearance investment and wellbeing: How older obese men orient to “aesthetic health”. *Psychology of Men and Masculinity, 17*(1), 84-91. <https://doi.org/10.1037/men0000012>
- Grabe, S., Grose, R. G., & Dutt, A. (2015). Women’s land ownership and relationship power: A mixed methods approach to understanding structural inequities and violence against women. *Psychology of Women Quarterly, 39*(1), 7-19. <https://doi.org/10.1177/0361684314533485>
- Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59-82. <https://doi.org/10.1177/1525822X05279903>
- Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Sage. <https://dx.doi.org/10.4135/9781483384436>
- Hagaman, A. K., & Wutich, A. (2017). How many interviews are enough to identify metathemes in multisited and cross-cultural research? Another perspective on Guest, Bunce, and Johnson’s (2006) landmark study. *Field Methods, 29*(1), 23-41. <https://doi.org/10.1177/1525822X16640447>

Hammersley, M. (2015). Sampling and thematic analysis: A response to Fugard and Potts.

*International Journal of Social Research Methodology*, 18(6), 687-688.

<https://doi.org/10.1080/13645579.2015.1005456>

Hancock, M. E., Amankwaa, L., Revell, M. A., & Mueller, D. (2016). Focus group data saturation: A new approach to data analysis. *The Qualitative Report*, 21(11), 2124-

2130. <https://nsuworks.nova.edu/tqr/vol21/iss11/13/>

Hayes, N. (2000). *Doing psychological research: Gathering and analysing data*. Open University Press.

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2016). Code saturation versus meaning saturation: How many interviews are enough? *Qualitative Health Research*, 27(4),

591-608. <https://doi.org/10.1177/1049732316665344>

Ho, K. H. M., Chiang, V. C. L., & Leung, D. (2017). Hermeneutic phenomenological analysis:

The “possibility” beyond “actuality” in thematic analysis. *Journal of Advanced Nursing*, 73(7), 1757–1766. <https://doi.org/10.1111/jan.13255>

Huijg, J. M., van der Zouwe, N., Cone, M. R., Verheijden, M. W., Middelkoop, B. J. C., & Gebhardt, W. A. (2015). Factors influencing the introduction of physical activity interventions in primary health care: A qualitative study. *International Journal of Behavioral Medicine*, 22, 404-414. <https://doi.org/10.1007/s12529-014-9411-9>

Jennings, M., Braun, V., & Clarke, V. (2019). Breaking gendered boundaries? Exploring constructions of counter-normative body hair practices in Aotearoa/New Zealand using story completion. *Qualitative Research in Psychology*, 16(1), 74-95.

<https://doi.org/10.1080/14780887.2018.1536386>

- Joffe, H. (2012). Thematic analysis. In D. Harper & A. R. Thompson (Eds.), *Qualitative methods in mental health and psychotherapy: A guide for students and practitioners* (pp. 209-223). Wiley. <https://doi.org/10.1002/9781119973249>
- Josselson, R. (2009). The present of the past: Dialogues with memory over time. *Journal of Personality, 77*(3), 647-668. <https://doi.org/10.1111/j.1467-6494.2009.00560.x>
- Karnieli-Miller, O., Strier, R., & Pessach, L. (2009). Power relations in qualitative research. *Qualitative Health Research, 10*(2), 279-289. <https://doi.org/10.1177/1049732308329306>
- Kidder, L. H., & Fine, M. (1987). Qualitative and quantitative methods: When stories converge. In M. M. Mark & L. Shotland (Eds.), *New directions for program evaluation* (pp. 57-75). Jossey-Bass.
- King, N. (2012). Doing template analysis. In G. Symon & C. Cassell (Eds.), *Qualitative organisation research: Core methods and current challenges* (pp. 426-450). Sage. <http://dx.doi.org/10.4135/9781526435620>
- King, N., & Brooks, J. (2017). Thematic analysis in organisational research. In C. Cassell, A.L. Cunliffe & G. Grandy (Eds.), *The Sage handbook of qualitative business and management research methods* (pp. 219-236). Sage.
- Komolova, M., Pasupathi, M., & Wainryb, C. (2020). "Things like that happen when you come to a new country": Bosnian refugee experiences with discrimination. *Qualitative Psychology, 7*(1), 93–113. <https://doi.org/10.1037/qup0000127>

Ivey, G., & Sonn, C. (2020). A psychosocial study of guilt and shame in White South African migrants to Australia. *Qualitative Psychology*, 7(1), 114–130.

<https://doi.org/10.1037/qup0000133>

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA publications and communications task force report. *American Psychologist*, 73(1), 26-46.

<http://dx.doi.org/10.1037/amp0000151>

Levitt, H. M., Motulsky, S. L., Wertz, F. J., Morrow, S. L., & Ponterotto, J. G. (2017). Recommendations for designing and reviewing qualitative research in psychology: Promoting methodological integrity. *Qualitative Psychology*, 4(1), 2-22.

<https://doi.org/10.1037/qup0000082>

Madill, A., & Gough, B. (2008). Qualitative research and its place in psychological science. *Psychological Methods*, 13(3), 254-271. <https://doi.org/10.1037/a0013220>.

Madill, A., Jordan, B., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology*, 91, 1-20. <https://doi.org/10.1348/000712600161646>

Malterud, K., Siersma, V. K., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), 1753-1760. <https://doi.org/10.1177/1049732315617444>

Malterud, K. (2013). Systematic text condensation: A strategy for qualitative analysis. *Scandinavian Journal of Public Health*, 40, 795-805.

<https://doi.org/10.1177/1403494812465030>

- Manago, A. M. (2013). Negotiating a sexy masculinity on social networking sites. *Feminism & Psychology, 23*(4), 478–497. <https://doi.org/10.1177/0959353513487549>
- Mason, M. (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum: Qualitative Social Research, 11*(3). <http://www.qualitative-research.net/index.php/fqs/article/view/1428/3027>
- Mauthner, N. S., & Doucet, A. (2003). Reflexive accounts and accounts of reflexivity in qualitative data analysis. *Sociology, 37*(3), 413-431. <https://doi.org/10.1177/00380385030373002>
- Maxwell, J. A. (2012). *A realist approach for qualitative research*. Sage.
- McClelland, S. I. (2017). Vulnerable listening: Possibilities and challenges of doing qualitative research. *Qualitative Psychology, 4*, 338–352. <http://dx.doi.org/10.1037/qup0000068>
- McLeod, J. (2015). *Doing research in counselling and psychotherapy* (3rd ed.). Sage.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2<sup>nd</sup> ed.). Sage.
- Miller, T., Birch, M., Mauthner, M., & Jessop, J. (Eds.). (2012). *Ethics in qualitative research* (2nd ed.). Sage.
- Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology, 52*(2), 250-260. <https://doi.org/10.1037/0022-0167.52.2.250>



- Morrow, S. (2007). Qualitative research in counseling psychology: Conceptual foundations. *The Counseling Psychologist, 35*(2), 209-235.  
<https://doi.org/10.1177/0011000006286990>
- Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research, 5*(2), 147-149. <https://doi.org/10.1177/104973239500500201>
- Morse, J. M. (2000). Determining sample size. *Qualitative Health Research, 10*(1), 3-5.  
<https://doi.org/10.1177/104973200129118183>
- Morse, J. M. (2015). "Data were saturated...". *Qualitative Health Research, 25*(5), 587-588.  
<https://doi.org/10.1177/1049732315576699>
- Nadin, S., & Cassell, C. (2004). Using data matrices. In C. Cassell & G. Symon (Eds.), *Essential guide to qualitative methods in organisational research* (pp. 271-287). Sage.  
<http://dx.doi.org/10.4135/9781446280119>
- Nelson, J. (2016). Using conceptual depth criteria: Addressing the challenge of reaching saturation in qualitative research. *Qualitative Research, 17*(5), 554-570.  
<https://doi.org/10.1177/1468794116679873>
- New Zealand Psychological Society (2012). *Code of ethics: For psychologists working in Aotearoa/New Zealand*. <https://www.psychology.org.nz/wp-content/uploads/2014/04/code-of-ethics.pdf>
- O'Connor, C., & Joffe, H. (2020). Intercoder Reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*.  
<https://doi.org/10.1177/1609406919899220>

- Onwuegbuzie, A. J., & Leech, N. L. (2007). Sampling designs in qualitative research: Making the sampling process more public. *The Qualitative Report*, *12*(2), 238-254.
- O'Reilly, M., & Parker, N. (2012). "Unsatisfactory saturation": A critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, *13*(2), 190-197. <https://doi.org/10.1177/1468794112446106>
- Palomäki, J., Laakasuo, M., & Salmela, M. (2013). "This is just so unfair!": A qualitative analysis of loss-induced emotions and tilting in on-line poker. *International Gambling Studies*, *13*(2), 255-270. <https://doi.org/10.1080/14459795.2013.780631>
- Patton, M. Q. (2015). *Qualitative research and evaluating methods: Integrating theory and practice* (4th ed.). Sage.
- Pilecki, A. (2017). The moral dimensions of the terrorist category construction in presidential rhetoric and their use in legitimizing counterterrorism policy. *Qualitative Psychology*, *4*(1), 23–40. <https://doi.org/10.1037/qup0000037>
- Ponterotto, J. G. (2005). Qualitative research in counseling psychology: A primer on research paradigms and philosophy of science. *Journal of Counseling Psychology*, *52*(2), 126–136. <https://doi.org/10.1037/0022-0167.52.2.126>
- Popper, K. (1959). *The logic of scientific discovery*. Basic Books.
- Ramazanoğlu, C., & Holland, J. (2002). *Feminist methodology: Challenges and choices*. Sage.
- Reicher, S. (2000). Against methodolatry: Some comments on Elliott, Fischer, and Rennie. *British Journal of Clinical Psychology*, *39*(1), 1-6. <https://doi.org/10.1348/014466500163031>
- Reissman, C. K. (2007). *Narrative methods for the human sciences*. Sage.

Rendón, M. J., & Nicolas, G. (2012). Deconstructing the portrayals of Haitian women in the media: A thematic analysis of images in the associated press photo archive.

*Psychology of Women Quarterly*, 36(2), 227-239.

<https://doi.org/10.1177/0361684311429110>

Ritchie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A.

Bryman & R. G. Burgess (Ed.), *Analysing qualitative data* (pp. 173-194). Taylor &

Francis. <https://doi.org/10.4324/9780203413081>

Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology*, 11(1), 25-41.

<https://doi.org/10.1080/14780887.2013.801543>

Robinson-Wood, T., Balogun-Mwangi, O., Weber, A., Zeko-Underwood, E., Rawle, S.-A. C.,

Popat-Jain, A., Matsumoto, A., & Cook, E. (2020). "What is it going to be like?": A

phenomenological investigation of racial, gendered, and sexual microaggressions

among highly educated individuals. *Qualitative Psychology*, 7(1), 43–58. [https://doi-](https://doi.org/10.1037/qup0000113)

[org./10.1037/qup0000113](https://doi.org/10.1037/qup0000113)

Ronkainen, N. J., Watkins, I., & Ryba, T. V. (2016). What can gender tell us about the pre-

retirement experiences of elite distance runners in Finland?: A thematic narrative

analysis. *Psychology of Sport and Exercise*, 22, 37-45.

<https://doi.org/10.1016/j.psychsport.2015.06.003>

Rowley, J., Rajbans, T., & Markland, B. (2020). Supporting parents through a narrative

therapeutic group approach: A participatory research project. *Educational*

*Psychology in Practice*. <https://doi.org/10.1080/02667363.2019.1700349>

Rubin, H. J., & Rubin, I. S. (1995). *Qualitative interviewing: The art of hearing data*. Sage.

- Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). Sage.
- Sandelowski, M. (1995). Sample size in qualitative research. *Research in Nursing & Health*, *18*(2), 179-183. <https://doi.org/10.1002/nur.4770180211>
- Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2017). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity: International Journal of Methodology*, *52*(4), 1893-1907. <https://doi.org/10.1007/s11135-017-0574-8>
- Schegloff, E. A. (2007). *Sequence organisation in interaction: A primer in conversation analysis*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511791208>
- Schnur, J. B., Ouelette, S. C., Bovbjerg, D. H., & Montgomery, G. H. (2009). Breast cancer patients' experience of external-beam Radiotherapy. *Qualitative Health Research*, *19*(5), 668-676. <https://doi.org/10.1177/1049732309334097>
- Shaw, C., Brady, L-M., & Davey, C. (2011). *Guidelines for research with children and young people*. London, UK: National Children's Bureau. <https://www.ncb.org.uk/resources-publications/guidelines-research-children-and-young-people>
- Sim, J., Saunders, B., Waterfield, J., & Kingstone, T. (2018). Can sample size in qualitative research be determined apriori? *International Journal of Social Research Methodology*, *21*(5), 619-634. <https://doi.org/10.1080/13645579.2018.1454643>
- Smith, B. (2017). Generalizability in qualitative research: Misunderstandings, opportunities and recommendations for the sport and exercise sciences. *Qualitative Research in Sport, Exercise and Health*, *10*(1), 137-149. <https://doi.org/10.1080/2159676X.2017.1393221>

- Smith, L. T. (2013). *Decolonizing methodologies: Research and indigenous peoples* (2nd ed.). University of Otago Press.
- Smith, J., & Firth, J. (2011). Qualitative data approaches: The framework approach. *Nurse Researcher, 18*(2), 52-62. <https://doi.org/10.7748/nr2011.01.18.2.52.c8284>
- Sparkes, A. C., & Smith, B. (2008). Narrative constructionist inquiry. In J. A. Holstein & J. F. Gubrium (Eds.), *Handbook of constructionist research* (pp. 295-314). London: Guilford Publications. [https://doi.org/10.1111/j.1744-6570.2009.01160\\_2.x](https://doi.org/10.1111/j.1744-6570.2009.01160_2.x)
- Sparkes, A. C., & Smith, B. (2009). Judging the quality of qualitative enquiry: Criteriology and relativism in action. *Psychology of Sport and Exercise, 10*, 49-497. <https://doi.org/10.1016/j.psychsport.2009.02.006>
- Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in qualitative evaluation: A framework for assessing research evidence*. Government Chief Social Researcher's Office, Prime Minister's Strategy Unit.
- Staneva, A., & Wittkowski, A. (2013). Exploring beliefs and expectations about motherhood in Bulgarian mothers: A qualitative study. *Midwifery, 29*, 260-267. <https://doi.org/10.1016/j.midw.2012.01.008>
- Tebbe, E. A., Moradi, B., Connelly, K. E., Lenzen, A. L., & Flores, M. (2018). "I don't care about you as a person": Sexual minority women objectified. *Journal of Counseling Psychology, 65*(1), 1-16. <https://doi.org/10.1037/cou0000255>
- Terry, G., & Hayfield, N. (2020). Reflexive thematic analysis. In M. R. M. Ward & S. Delamont (Eds.), *Handbook of qualitative research in education* (pp. 428-439). <https://doi.org/10.1177/1077800410383121>

- 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>
- Trainor, L. R., & Bundon, A. (2020). Developing the craft: Reflexive accounts of doing reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*. ONLINE FIRST. <https://doi.org/10.1080/2159676X.2020.1840423>
- Trans, V.-T., Procher, R., Tran, V.-C., & Ravaud, P. (2017). Predicting data saturation in qualitative surveys with mathematical models from ecological research. *Journal of Clinical Epidemiology*, 82, 71-78. <https://doi.org/10.1016/j.jclinepi.2016.10.001>
- Wiggins, S. (2017). *Discursive psychology: Theory, method and applications*. Sage.
- Wilkinson, S. (1988). The role of reflexivity in feminist psychology. *Women’s Studies International Forum*, 11(5), 493-502. [https://doi.org/10.1016/0277-5395\(88\)90024-6](https://doi.org/10.1016/0277-5395(88)90024-6)
- Willcox, R., Moller, N., & Clarke, V. (2019). Exploring attachment incoherence in bereaved families’ therapy narratives: An attachment theory-informed thematic analysis. *The Family Journal: Counseling and Therapy for Couples and Families*, 27(3), 339-347.  
<https://doi.org/10.1177/1066480719853006>
- Willig, C. (2013). *Introducing qualitative research in psychology* (3rd ed.). Open University Press.
- Yardley, L. (2015). Demonstrating validity in qualitative psychology. In J. A. Smith (Ed.), *Qualitative psychology: A practical guide to research methods* (2nd ed., pp. 257-272). Sage.



## Box 1: Overviewing Conceptual and Design Thinking for Thematic Analysis Research

Reflexive TA researchers should consider:

- The type of TA to be used.
- Orientations to TA – experiential/constructionist, inductive/deductive, semantic/latent.
- The philosophical meta-theories (ontologies and epistemologies) underpinning the research.
- Any methodological, explanatory and political/ideological theories informing the research – more loosely, as part of the package of things the researcher “brings” to data analysis, and/or more formally, as the theoretical lens(es) through which the data are interpreted in deductive orientations.
- The research question – both the initial (potentially broader) research question and a more refined/focused question settled on following (some) data analysis.
- The method(s) of data collection, and particular orientations/modalities (e.g., narrative, feminist, video-call interviewing).
- Participant group and dataset matters – selection strategy (e.g., convenience, purposive), constitution (heterogenous, homogenous), size, and recruitment/selection.
- The researcher’s positioning in relation to the topic and any participant group.
- Ethical (and political) considerations (e.g., how consent will be negotiated; the politics of representation).
- Conceptually coherent quality practices and standards.

Reflecting and deciding is not necessarily in this order, and not necessarily just *before* the research, or a particular phase of the research, commences. Some post-hoc reflexive “unpacking” of assumptions and practices may be required for researchers to strive to “own



their perspectives” (Elliott et al., 1999) in the reporting of the research, and explain and defend their choices and practices.

**Table 1: A Typology of Suitable Research Questions for Reflexive Thematic Analysis**

Research question focus	Examples
People’s contextually situated lived experiences and interpretations of subjective phenomena.	Bosnian refugees’ experiences of discrimination in the US (Komolova et al., 2020); South African migrants’ feelings of guilt and shame around leaving their homeland (Ivey & Sonn, 2020).
The views, perceptions, understandings, perspectives, needs, motivations of particular groups, about particular phenomena, in particular contexts (often combined with lived experience questions.)	Public perceptions and symbolic associations of electricity network technologies in the UK (Devine-Wright & Devine-Wright, 2009); African American college women’s beauty and body image concerns (Awad et al., 2016).
The factors or social processes that influence the shape and texture of particular phenomena.	The processes and factors that make interpersonal relationships meaningful to young men transitioning to adulthood and beginning postsecondary education and how these relationships influence their life plans (Arbeit et al., 2016); the factors influencing the introduction of physical activity interventions in primary health care (Huijg et al., 2015).
The things people <i>do</i> in the world – their contextually	How incoherence, a narrative marker of attachment insecurity, is displayed in the talk of families undergoing

<p>situated (variously conceptualized as) behaviors or practices, and their sense-making around them.</p>	<p>bereavement family therapy (Willcox et al., 2019); how new fathers request, offer, and receive social support in an online chat room (Fletcher &amp; StGeorge, 2011).</p>
<p>The (often implicit) contextually situated rules and norms that regulate particular phenomena.</p>	<p>How sporting cultural values and unwritten cultural norms influence the occurrence and experience of overuse injuries in rhythmic gymnastics (Cavallerio et al., 2016); how the organizational cultural experiences of elite youth footballers shape their identity development and behavior (Champ et al., 2020).</p>
<p>The representation of particular “social objects” or phenomena in particular contexts, and the implications or effects of these.</p>	<p>The moral dimensions of the construction of the category “terrorist” in presidential political speeches and implications of these for legitimating counter-terrorism policy (Pilecki, 2017); the representation of Haitian women in mainstream US media (Rendón &amp; Nicolas, 2012).</p>
<p>The social or discursive construction of particular “social objects”, subject positions or other social phenomena in particular contexts and the implications and effects of these.</p>	<p>People’s constructions and meaning-making around counter-normative body hair practices (Jennings et al., 2019); older fat men’s - involved in a weight loss intervention - constructions of their bodies and bodily change (Gough et al., 2016).</p>