WHAT IS THEMATIC ANALYSIS?

TA is a method for systematically identifying, organizing, and offering insight into patterns of meaning (themes) across a data set. Through focusing on meaning across a data set, TA allows the researcher to see and make sense of collective or shared meanings and experiences. Identifying unique and idiosyncratic meanings and experiences found only within a single data item is not the focus of TA. This method, then, is a way of identifying what is common to the way a topic is talked or written about and of making sense of those commonalities.

What is common, however, is not necessarily in and of itself meaningful or important. The patterns of meaning that TA allows the researcher to identify need to be important in relation to the particular topic and research question being explored. Analysis produces the answer to a question, even if, as in some qualitative research, the specific question that is being answered only becomes apparent through the analysis. Numerous patterns could be identified across any data set—the purpose of analysis is to identify those relevant to answering a particular research question. For instance, in researching white-collar workers’ experiences of sociality at work, a researcher might interview people about their work environment and start with questions about their typical workday. If most or all reported that they started work at around 9:00 a.m., this would be a pattern in the data, but it would not necessarily be a meaningful or important one. If many reported that they aimed to arrive at work earlier
than needed so that they could chat with colleagues, this could be a meaningful pattern.

TA is a flexible method that allows the researcher to focus on the data in numerous different ways. With TA you can legitimately focus on analyzing meaning across the entire data set, or you can examine one particular aspect of a phenomenon in depth. You can report the obvious or semantic meanings in the data, or you can interrogate the latent meanings, the assumptions and ideas that lie behind what is explicitly stated (see Braun & Clarke, 2006). The many forms TA can take means that it suits a wide variety of research questions and research topics.

WHY THEMATIC ANALYSIS?

The two main reasons to use TA are accessibility and flexibility. For people new to qualitative research, TA provides an entry into a way of doing research that otherwise can seem vague, mystifying, conceptually challenging, and overly complex. It offers a way into qualitative research that teaches the mechanics of coding and analyzing qualitative data systematically, which can then be linked to broader theoretical or conceptual issues. For much qualitative research, the relationship is reversed. For example, to do DA, the researcher needs to first be familiar with complex theoretical perspectives on language (see Chapter 8 of this volume), which invert the commonsense view of language as a mirroring reality—instead, language is theorized as creating reality. Knowing this background is essential because it guides what the researcher sees in the data, how they code and analyze the data, and the claims that they make. In contrast, TA is only a method of data analysis, rather than being an approach to conducting qualitative research. We see this as a strength because it ensures the accessibility and flexibility of the approach.

TA offers a way of separating qualitative research out from these broader debates, where appropriate, and making qualitative research results available to a wider audience. Its accessibility as a method also suits multimethods research being conducted by research teams, where not everyone is a qualitative expert. TA also has a lot of potential for use within participatory research projects—such as participatory action research (see Chapter 11 of this volume) or memory work (Onyx & Small, 2001)—in which many involved in the analysis are not trained researchers.

FLEXIBILITY AND CHOICES IN THEMATIC ANALYSIS

Linked to the fact that it is just a method, one of the main reasons TA is so flexible is that it can be conducted in a number of different ways. TA has the ability to straddle three main continua along which qualitative research approaches can be located: inductive versus deductive or theory-driven data coding and analysis, an experiential versus critical orientation to data, and an essentialist versus constructionist theoretical perspective. Where the researcher locates their research on each of these continua carries a particular set of assumptions, and this delimits what can and cannot be said in relation to the data as well as how data can and should be interpreted (for a detailed discussion of these positions, see Volume 1, Chapter 1, this handbook). Any researcher doing TA needs to actively make a series of choices as to what form of TA they are using and to understand and explain why they are using this particular form (Braun & Clarke, 2006).

An inductive approach to data coding and analysis is a bottom-up approach and is driven by what is in the data. What this means is that the codes and themes derive from the content of the data themselves—so that what is mapped by the researcher during analysis closely matches the content of the data. In contrast, a deductive approach to data coding and analysis is a top-down approach, where the researcher brings to the data a series of concepts, ideas, or topics that they use to code and interpret the data. What this means is that the codes and themes derive more from concepts and ideas the researcher brings to the data—here, what is mapped by the researcher during analysis does not necessarily closely link to the semantic data content.

In reality, coding and analysis often uses a combination of both approaches. It is impossible to be purely inductive, as we always bring something to the data when we analyze it, and we rarely completely ignore the semantic content of the data when we
code for a particular theoretical construct—at the very least, we have to know whether it is worth coding the data for that construct. One tends to predominate, however, and a commitment to an inductive or deductive approach also signals an overall orientation that prioritizes either participant or data-based meaning or researcher or theory-based meaning. For this reason, inductive TA often is experiential in its orientation and essentialist in its theoretical framework, assuming a knowable world and “giving voice” to experiences and meanings of that world, as reported in the data. Deductive TA is often critical in its orientation and constructionist in its theoretical framework, examining how the world is put together (i.e., constructed) and the ideas and assumptions that inform the data gathered. These correspondences are not given, however, or necessary. Consistency and coherence of the overall framework and analysis is what is important.

Braun and colleagues’ analysis of gay and bisexual men’s experiences of sexual coercion provides a good example of a more inductive, experiential, essentialist form of TA, in which different forms or modes of sexual coercion were identified from men’s reported diverse experiences (Braun, Terry, Gavey, & Fenaughty, 2009). Clarke and Kitzinger’s (2004) analysis of representations of lesbian and gay parents on television talk shows is a good example of more deductive, critical, constructionist TA. This study drew on the concept of heteronormativity to examine how participants in liberal talk-show debates routinely invoke discursive strategies of normalization, emphasizing lesbian- and gay-headed families’ conformity to norms of White, middle-class heterosexuality, as a response to homophobic and heterosexist accounts of lesbian and gay parenting and its impact on children.

Like any form of analysis, TA can be done well, and it can be done poorly. Essential for doing good TA are a clear understanding of where the researcher stands in relation to these possible options, a rationale for making the choices they do, and the consistent application of those choices throughout the analysis (further criteria are discussed later in the chapter). We now provide a worked example that lays out how you actually do TA.

**THEMATIC ANALYSIS: A WORKED EXAMPLE**

We illustrate how to do TA using a worked example from an ongoing project that examines sexuality, gender identity, and higher education (Braun & Clarke, 2009; Clarke & Braun, 2009b). Like many research projects, which evolve not just from identified gaps in the literature but also from topics that grab us and pique our curiosity, this one developed as a result of our experiences and reflections related to teaching and teaching training as well as from intellectual and political questions about sexuality and gender identity in the classroom.

Part of the project involved interviewing 20 LGBT-identified students in New Zealand (10 students) and Britain (10 students) to understand their experiences of university life. Our worked example of thematic analysis uses data from four of the British students. The students varied on race/ethnicity (one British Asian, three White, one born in Europe), class (working or middle class) and age (one middle-aged student), but they were all studying social science subjects. The scope of university life was broadly conceived, including the classroom, the curriculum, and “hidden” curriculum—the norms and ideas implicitly conveyed at university—interactions with course peers and teaching staff, the campus and wider university environment, the local geographic area, and the local gay scene. In the semistructured interviews, which lasted around an hour, participants were all asked about their expectations of university life, whether they were out (open) about their sexuality at university, their experiences of the classroom and the curriculum, their views on LGBT lecturers coming out in the classroom, and, if they were studying a people-based discipline (Ellis, 2009), whether LGBT issues were included when relevant. Experiences and perceptions of the wider campus environment and of student housing, interactions with other students, friendship networks and social life, and the best and worst things about university life as a LGBT student were also covered.

The interviews were audio recorded and then transcribed orthographically, reproducing all spoken words and sounds, including hesitations, false
starts, cutoffs in speech (indicated by a dash; e.g., thin-), the interviewer’s guggles (e.g., mm-hm, ah-ha), laughter, long pauses [indicated by (pause)], and strong emphasis (indicated by underscore). Commas signal a continuing intonation, broadly commensurate with a grammatical comma in written language; inverted commas are used to indicate reported speech; three full-stops in a row (...) signal editing of the transcript. We have mainly edited for brevity, removing any words or clauses that are not essential for understanding the overall meaning of a data extract. There are many different styles of transcription (e.g., Edwards & Lampert, 1993) but if transcribing audio data for TA, this level of detail is more than sufficient. As a general practice, we do not advocate “cleaning up” the transcript (such as making it more grammatical or removing hesitations, pauses, and guggles) when working with data. Depending on your form of TA, such details may be omitted from quoted data (if done, it should be noted); however, because the details can be revealing, we suggest working with a full transcript while doing the analysis.

This topic, research question, and data collection method all suited TA. The research question was experiential and exploratory, so our worked example illustrates a primarily experiential form of TA, within a contextualist framework, which assumes truth can be accessed through language, but that accounts and experiences are socially mediated (Madill, Jordan, & Shirley, 2000). It illustrates a combination of inductive and deductive TA: inductive as we mainly code from the data, on the basis of participants’ experiences (meaning our analytic lens does not completely override their stories); deductive as we draw on theoretical constructs from feminist and queer scholarship like heterosexism (Adam, 1998), compulsory heterosexuality (Rich, 1980), heteronormativity (Warner, 1991), and the hidden curriculum of heteronormativity (Epstein, O’Flynn, & Telford, 2003) to render visible issues that participants did not explicitly articulate. This means that the data are broadly interpreted within a feminist and a queer (e.g., Clarke & Braun, 2009a; Gamson, 2000) theoretical and ideological framework.

A SIX-PHASE APPROACH TO THEMATIC ANALYSIS

The six phases in our approach to TA (Braun & Clarke, 2006) are outlined and illustrated using worked examples. This is an approach to TA and to learning to do TA. More experienced analysts will (a) likely have deeper insights into their data during familiarization, (b) find the process of coding quicker and easier and be able to code at a more conceptual level, and (c) more quickly and confidently develop themes that need less reviewing and refining, especially if working with a smaller data set. Writing is also likely to take a more central place throughout analysis with more experience. The point we wish to emphasize is that certain skills of analysis develop only through experience and practice. Even experienced researchers, however, will draw and redraw lots of thematic maps when searching for themes and will engage in extensive review processes when working with larger data sets. A thematic map is a visual (see Braun & Clarke, 2006) or sometimes text-based (see Frith & Gleeson, 2004) tool to map out the facets of your developing analysis and to identify main themes, subthemes, and interconnections between themes and subthemes.

Phase 1: Familiarizing Yourself With the Data

Common to all forms of qualitative analysis, this phase involves immersing yourself in the data by reading and rereading textual data (e.g., transcripts of interviews, responses to qualitative surveys) and listening to audio recordings or watching video data. If you have audio data, we recommend listening to them at least once as well as reading the transcript, especially if you did not collect the data or transcribe them. Making notes on the data as you read—or listen—is part of this phase. Use whatever format works for you (e.g., annotating transcripts, writing comments in a notebook or electronic file, underling portions of data) to highlight items potentially of interest. Note-making helps you start to read the data as data. Reading data as data means not simply absorbing the surface meaning of the words on the page, as you might read a novel or
magazine, but reading the words actively, analytically, and critically, and starting to think about what the data mean. This involves asking questions like, How does this participant make sense of their experiences? What assumptions do they make in interpreting their experience? What kind of world is revealed through their accounts? We will illustrate this with a brief example from Andreas’s interview:

Andreas: let’s say I’m in a seminar and somebody a a man says to me “oh look at her” (Int: mm) I’m not going “oh actually I’m gay” (Int: mm [laughter]) I’ll just go like “oh yeah” (Int: mhm) you know I won’t fall into the other one and say “oh yeah” (Int: yep) “she looks really brilliant”

Our initial observations included (a) Andreas reports a common experience of presumed heterosexuality, (b) coming out is not an obvious option, (c) social norms dictate a certain response, (d) the presumption of heterosexuality appears dilemmatic, and (e) he colludes in the presumption but minimally (to avoid social awkwardness). Looking a bit more deeply, we speculated that (a) Andreas values honesty and being true to yourself, but (b) he recognizes a sociopolitical context in which that is constrained, and (c) walks a tightrope trying to balance his values and the expectations of the context. These initial observations suggest the data will provide fertile grounds for analysis; reading Andreas’s answer as data reveals the richness that can be found in even brief extracts of text. We did deliberately pick a particularly rich extract, however; not all extracts will be as vivid as this one, and you may have little or nothing to say about some parts of your data.

The aim of this phase is to become intimately familiar with your data set’s content and to begin to notice things that might be relevant to your research question. You need to read through your entire data set at least once—if not twice, or more—until you feel you know the data content intimately. Make notes on the entire data set as well as on individual transcripts. Note-making at this stage is observational and casual rather than systematic and inclusive. You are not coding the data yet, so do not agonize over it. Notes would typically be a stream of consciousness, a messy rush of ideas, rather than polished prose. Such notes are written only to and for you to help you with the process of analysis—think of them as memory aids and triggers for coding and analysis. At most, they may be shared among research team members.

Phase 2: Generating Initial Codes

Phase 2 begins the systematic analysis of the data through coding. Codes are the building blocks of analysis: If your analysis is a brick-built house with a tile roof, your themes are the walls and roof and your codes are the individual bricks and tiles. Codes identify and provide a label for a feature of the data that is potentially relevant to the research question (Exhibit 4.1 shows an example of coded data). Coding can be done at the semantic or the latent level of meaning. Codes can provide a pithy summary of a portion of data or describe the content of the data—such descriptive or semantic codes typically stay close to content of the data and to the participants’ meanings. An example of this is “fear/anxiety about people’s reactions to his sexuality” in Exhibit 4.1. Codes can also go beyond the participants’ meanings and provide an interpretation about the data content. Such interpretative or latent codes identify meanings that lie beneath the semantic surface of the data. An example of this is the “coming out imperative”; this code offers a conceptual interpretation to make sense of what Andreas is saying (see Exhibit 4.1).

Some codes mirror participants’ language and concepts; others invoke the researchers’ conceptual and theoretical frameworks. For example, the code “not hiding (but not shouting)” stayed close to the participants’ use of language (e.g., John said “I don’t make an attempt to hide that I’m gay but at the same time I’m not very forward about it”). In contrast, the code “modifying behavior . . . to avoid heterosexism” invoked our frame of reference: No student spontaneously used the term heterosexism to describe their experiences, but we interpret their accounts through this framework (Adam, 1998).

Codes are succinct and work as shorthand for something you, the analyst, understands; they do not have to be fully worked-up explanations—those come later. Codes will almost always be a mix of the descriptive and interpretative. A novice coder will likely (initially) generate more descriptive codes; as noted, interpretative approaches to coding develop
with experience. This does not mean that interpretative codes are better— they are just harder to “see” sometimes. What is important for all codes is that they are relevant to answering your research question. Coding is something we get better at with practice.

TA is not prescriptive about how you segment the data as you code it (e.g., you do not have to produce a code for every line of transcript). You can code in large or small chunks; some chunks will not be coded at all. Coding requires another thorough read of every data item, and you should code each data item in its entirety before coding another. Every time you identify something that is potentially relevant to the research question, code it. We say “potentially” because at this early stage of analysis, you do not know what might be relevant: Inclusivity should be your motto. If you are unsure about whether a piece of data may be relevant, code it. It is much easier to discard codes than go back to the entire data set and recode data, although some recoding is part of the coding process.

Once you identify an extract of data to code, you need to write down the code and mark the text associated with it. You can code a portion of data with more than one code (as Exhibit 4.1 shows). Some people code on hard-copy data, clearly identifying the code name, and highlighting the portion of text associated with it. Other techniques include using computer software to manage coding (see Volume 1, Chapter 16, this handbook) or using file cards—one card for each code, with data summary and location information listed—or cutting and pasting text into a new word-processing file, created for this purpose (again, ensure that you record where all excerpts came from). An advantage of the latter methods is that you collate your coded text as you code, but there is no right or wrong way to manage the physical process of coding. Work out what suits you best. What is important is that coding is inclusive, thorough, and systematic.

After you generate your first code, keep reading the data until you identify the next potentially relevant excerpt: You then have to decide whether you can apply the code you have already used or whether a new code is needed to capture that piece of data. You repeat this process throughout each data item.
and the entire data set. As your coding progresses, you can also modify existing codes to incorporate new material. For example, our code “modifying behavior, speech, and practices to avoid heterosexism” was initially titled “modifying behavior to avoid heterosexism.” Because students also reported modifying speech and things like dress or self-presentation to avoid “trouble,” we expanded this code beyond “behavior” to make it better fit what participants said. It is a good idea to revisit the material you coded at the start because your codes will have likely developed during coding: Some recoding and new coding of earlier coded data may be necessary.

This stage of the process ends when your data are fully coded and the data relevant to each code has been collated. Exhibit 4.2 provides some examples of codes we generated from our data, with a few data extracts collated for each code. Depending on your topic, data set, and precision in coding, you will have generated any number of codes—there is no maximum. What you want are enough codes to capture both the diversity, and the patterns, within the data, and codes should appear across more than one data item.

Phase 3: Searching for Themes

In this phase, your analysis starts to take shape as you shift from codes to themes. A theme “captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set” (Braun & Clarke, 2006, p. 82). Some qualitative researchers make reference to “themes emerging from the data,” as if their data set was a pile of crocodile eggs and analysis involved watching the eggs until each baby crocodile (theme) emerged, perfectly formed, from within. If only it were so easy. Searching for themes is an active process, meaning we generate or construct themes rather than discovering them. Although we call this phase “searching for themes,” it is not like archaeologists digging around, searching for the themes that lie hidden within the data, preexisting the process of analysis. Rather, analysts are like sculptors, making choices about how to shape and craft their piece of stone (the “raw data”) into a work of art (the analysis). Like a piece of stone, the data set provides the material base for analysis and limits the possible end product, but many different variations could be created when analyzing the data.

This phase involves reviewing the coded data to identify areas of similarity and overlap between codes: Can you identify any broad topics or issues around which codes cluster? The basic process of generating themes and subthemes, which are the subcomponents of a theme, involves collapsing or clustering codes that seem to share some unifying feature together, so that they reflect and describe a coherent and meaningful pattern in the data. In our data, we noticed codes clustering around heterosexism and homophobia. Examining these in more detail, we identified that the codes either focused on experiences of heterosexism and homophobia, or responses to and ways of managing heterosexism and homophobia. We then constructed one theme using all the codes relating to the participants’ experiences of heterosexism and homophobia (e.g., “incident of (naming) homophobia/heterosexism”; “tensions in relating to straight men”) and another using the codes relating to the participants’ management of (actual and feared) heterosexism (e.g., “monitoring/assessing people and the environment for the possibility of heterosexism”; “modifying speech, behavior, and practices to avoid heterosexism”). The code “managing the heterosexual assumption by minimal agreement” (see Exhibit 4.1) appeared to be a variation of the code “modifying speech, behavior, and practices to avoid heterosexism,” and so it was incorporated into that theme.

A lot of codes also clustered around the issue of identity but did not form one obvious theme. In this case, after exploring lots of different ways to combine these codes into themes and drawing lots of thematic maps, we generated two themes: one around coming out and being out, and one around different versions of being a gay man. These provided the best mapping of the identity data in relation to our research questions. A number of codes cut across both themes, such as the notion of good gays (who conform to the norms of compulsory heterosexuality as much as possible by being “straight-acting” and “straight-looking”; Taulke-Johnson, 2008) and bad gays (who are “politically active and culturally assertive”; Epstein, Johnson, & Steinberg, 2000, p. 19). This example is not a case of undesirable
### Exhibit 4.2
**Six Codes With Illustrative Data Extracts (Direct Quotes)**

<table>
<thead>
<tr>
<th>Modifying speech, behavior, and practices to avoid heterosexism</th>
<th>Tensions in relating to straight men</th>
<th>Incident of (naming) homophobia/heterosexism</th>
<th>Fear/anxiety about people’s reactions to his sexuality</th>
<th>Managing the heterosexual assumption by minimal agreement</th>
<th>Monitoring/assessing people/the environment for the possibility of heterosexism</th>
</tr>
</thead>
<tbody>
<tr>
<td>I’m not somebody that goes out looking for trouble . . . (David)</td>
<td>I know if I go into a lecture hall and I’m like on my own without a group some of the lads are a little bit less inclined to sort of sit with you in a way . . . (David)</td>
<td>This one guy drunk just came along and just started telling me to my face I was sick that there was something wrong with me, there was something wrong with us and we should [f**k] the hell out of there . . . (Asha)</td>
<td>I’d just hate to see what my dad would do (Asha)</td>
<td>I realize and notice that I sometimes try to erm not conceal it, that’s not the right word, but erm let’s say I’m in . . . seminar and somebody- a man says to me “oh look at her” I’m not going “oh actually I’m gay” I’ll just go “oh yeah” you know I won’t fall into the other one and say “oh yeah she looks really brilliant . . .” (Asha)</td>
<td>just how much I know them . . . there’s a lot of people I wouldn’t go into great detail with about what I get up to and stuff, whereas other people I would yeah I suppose I like to feel reasonably safe when telling them stuff like that (John)</td>
</tr>
<tr>
<td>so you don’t want to necessarily go down that road, so you sort of make up some- not make up some story, but you only tell sort of half the truth (Andreas)</td>
<td>that’s the old thing that it’s sort of easier in a way to be out with females than with sort of you know blokey blokes (Andreas)</td>
<td>that’s the old thing that it’s sort of easier in a way to be out with females than with sort of you know blokey blokes (Andreas)</td>
<td>I have once seen a group of lads standing outside one of the [gay] bars like jeering and stuff . . . (John)</td>
<td>I do remember being a bit worried about who I’d end up living with because I opted for a student house and that’s five random people thrown with you (Asha)</td>
<td>I just remember him making some kind of comment to me on the bus to London about Earl’s Court and gay art or something and er yeah, and I just didn’t think that he’d be the sort of person that’d be that bothered by things like that you know what I mean (John)</td>
</tr>
<tr>
<td>I would feel fine going clubbing [to a straight club] with my boyfriend but I’d be very wary of making it obvious (John)</td>
<td>I did have quite a- an interesting conversation with one guy . . . at the end of the conversation . . . he goes . . . “you’re an actual really nice guy aren’t you? ‘Cos I wasn’t really over sure about you when we first started, ‘cos you could tell you were gay as soon as you walked through the door” . . . my reaction was “get knotted” sort of thing and just walked off ‘cos I thought you know that shouldn’t be a issue (David)</td>
<td>There’s this one person from work who’s extremely religious, and I don’t mention it [my sexuality] whatsoever, he did mention one story that er gay people were cursed by the god and turned into monkeys (Asha)</td>
<td>I was asked . . . “why did you come from another country to Bristol?” if you go into this erN有关规定 about “oh there was somebody involved” then you’re close to to “who was it then?” . . . you never know how people react (Andreas)</td>
<td>I was asked “what are you doing then in Bristol?” . . . “was it a nice girl?” so you don’t want to necessarily go down that road so you . . . only tell sort of half the truth (Andreas)</td>
<td>you go to a party where you don’t know anybody . . . and “oh let me introduce you to so and so” and then you sort of after a while you start this there’s always testi-testing can I not can I tell that- but I mean what will happen if I tell will people then immediately say “oh sorry mate I need a drink” (Andreas)</td>
</tr>
<tr>
<td>if I’m out with my boyfriend and it’s late at night and we’re sort of walking home and we’ll sort of holding hands and . . . if it’s like mostly girls and stuff and that’s okay but if a group of lads were coming like we would loosen up or go via like a different route (David)</td>
<td>I had a couple of incidents where all of sudden when you then say “I’m gay” then it’s this (pause) you know erm wink wink nudge nudge thing sort of these jokes (Andreas)</td>
<td>I had a couple of incidents where all of sudden when you then say “I’m gay” then it’s this (pause) you know erm wink wink nudge nudge thing sort of these jokes (Andreas)</td>
<td>if I came out there I probably would have been lad bait so I decided to keep it to myself . . . I had an idea of what kind of response I would get and so just sensible decision of just keeping my mouth shut (Asha)</td>
<td>(Andreas)</td>
<td></td>
</tr>
<tr>
<td>with other Asians as well . . . I wouldn’t say probably I would just shut up (Asha)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
overlap between themes; it illustrates that certain concepts or issues may cut across themes and provide a unifying framework for telling a coherent story about what is going on in the data, overall.

Another important element of this stage is starting to explore the relationship between themes and to consider how themes will work together in telling an overall story about the data. Good themes are distinctive and, to some extent, stand alone, but they also need to work together as a whole. Think of themes like the pieces of a jigsaw puzzle: Together they provide a meaningful and lucid picture of your data. In your analysis, one central theme or concept may draw together or underpin all or most of your other themes—for our example, this would be heteronormativity.

During this stage, it can also be useful to have a miscellaneous theme, which includes all the codes that do not clearly fit anywhere, which may end up as part of new themes or being discarded. Being able to let go of coded material and indeed provisional themes if they do not fit within your overall analysis is an important part of qualitative analysis. Remember, your job in analyzing the data, and reporting them, is to tell a particular story about the data, that answers your research question. It is not to represent everything that was said in the data.

How many themes are enough or too many? For our data set, we generated six themes; for brevity, only four are summarized in Exhibit 4.3. Unfortunately, there is no magic formula that states that if you have X amount of data, and you are writing a report of Y length, you should have Z number of themes. The more data you have, the more codes and thus themes, you will likely generate; if you are writing a longer report, you will have space to discuss more themes. But with more themes, your analysis can lose coherence. What is essential is that your themes are presented in sufficient depth and detail to convey the richness and complexity of your data—you are unlikely to achieve this if you report more than six or seven themes in a 10,000-word report. You should end this phase with a thematic map or table outlining your candidate themes, and you should collate all the data extracts relevant to each theme, so you are ready to begin the process of reviewing your themes.

**Phase 4: Reviewing Potential Themes**

This phase involves a recursive process whereby the developing themes are reviewed in relation to the coded data and entire data set. This phase is essentially about quality checking. It is particularly important for novice researchers and for those working with very large data sets, where it is simply not possible to hold your entire data set in your head. The first step is to check your themes against the collated extracts of data and to explore whether the theme works in relation to the data. If it does not, you might need to discard some codes or relocate them under another theme; alternatively, you may redraw the boundaries of the theme, so that it more meaningfully captures the relevant data. If these tweaks do not work, you might need to discard your theme altogether and start again—you should not force your analysis into coherence. Key questions to ask are as follows:

- Is this a theme (it could be just a code)?
- If it is a theme, what is the quality of this theme (does it tell me something useful about the data set and my research question)?
- What are the boundaries of this theme (what does it include and exclude)?
- Are there enough (meaningful) data to support this theme (is the theme thin or thick)?
- Are the data too diverse and wide ranging (does the theme lack coherence)?

You may end up collapsing a number of potential themes together or splitting a big broad theme a number of more specific or coherent themes.

Once you have a distinctive and coherent set of themes that work in relation to the coded data extracts, you should undertake the second stage in the review process—reviewing the themes in relation to the entire data set. This involves one final reread of all your data to determine whether your themes meaningfully capture the entire data set or an aspect thereof. What you are aiming for is a set of
themes that capture the most important and relevant elements of the data, and the overall tone of the data, in relation to your research question. If your thematic map and set of themes does this, good. You can move to the next phase. If not, further refining and reviewing will be necessary to adequately capture the data. A mismatch will most likely occur if selective or inadequate coding has taken place, or if coding evolved over a data set and data were not recoded using the final set of codes. Revision at this stage might involve creating additional themes or tweaking or discarding existing themes.

**Phase 5: Defining and Naming Themes**

When defining your themes, you need to be able to clearly state what is unique and specific about each theme—whether you can sum up the essence of each theme in a few sentences is a good test of this (see Exhibit 4.3). A good thematic analysis will have themes that (a) do not try to do too much, as themes should ideally have a singular focus; (b) are related but do not overlap, so they are not repetitive, although they may build on previous themes; and (c) directly address your research question. Each theme identified in Exhibit 4.3 has a clear focus, scope, and purpose; each in turn builds on and develops the previous theme(s); and together the themes provide a coherent overall story about the data. In some cases, you may want to have subthemes within a theme. These themes are useful in cases in which there are one or two overarching patterns within the data in relation to your question, but each is played out in a number of different ways. Themes 3 and 4, for example, could be

**Exhibit 4.3**

Definitions and Labels for Selected Themes

```text
Theme 1. “There’s always that level of uncertainty”: Compulsory heterosexuality at university. Maps the participants’ experiences of (infrequent) homophobia and (constant) heterosexism and highlights tensions experienced in relating to (straight) others, particularly people who are common sources of heterosexism and overt homophobia (i.e., straight men; members of religious and non-White groups), and feelings, or fear, of exclusion and not belonging. Heterosexism meant participants negotiated their sexual identities in an uncertain environment and experienced constant (but minimized) fear of people’s reactions to their sexuality. They had expected university students to be liberal and open minded and were surprised and disappointed they weren’t. But they felt this applied if you were “straight-acting,” indicating university is a safe space only if you are a “good gay.” Participants’ experienced difficulty coming out at university but also internalized and took responsibility for these difficulties rather than viewing coming out as something that is difficult because of compulsory heterosexuality. Although participants expressed some anger about experiences of overt homophobia, some homophobic and heterosexist “banter” (e.g., antigay humor) was acceptable if from friends—an indication that friends were comfortable with their sexuality but wasn’t acceptable it from strangers. The heterosexual assumption and compulsory heterosexuality were typically framed as a to-be-expected part of normal life.

Theme 2. “I don’t go out asking for trouble”: Managing heterosexism. Outlines the ways the participants modified their speech, behavior, and practices to avoid heterosexism and homophobia and continually monitored people and the environment for evidence of potential heterosexism or homophobia. They constantly weighed whether it was safe to come or be out with a particular person or in a particular space. The participants typically assumed responsibility for managing heterosexism (they don’t “ask” for trouble) and accepted this as a normal part of life. They seemed to lack a sense of entitlement to live free from heterosexism and a political and conceptual language with which to interpret their experiences of heterosexism and homophobia.

Theme 3. “I’m not hiding, but I’m not throwing it in people’s faces”: Being out (but not too out) at university. Focuses on the degree to which the participants were out and open about their sexuality at university and the management of sexual identity amid competing pressures to be a “happy, healthy gay” (comfortable with and open about their sexuality, with a “fully realized” gay identity) and a “good gay” (not too “overt”; not “forcing” their homosexuality on others).

Theme 4. Mincing queens versus ordinary guys who just happen to be gay. Focuses on participants’ resistance to a gay identity as a “master status” (Becker, 1963), an identity that overrides all other identities—they wanted to be seen as an ordinary guy who just happens to be gay. They took responsibility for carefully managing other people’s perceptions of their sexual identity, acutely aware that it takes very little to be judged as “too gay” (a “bad gay”). They felt very limited by popular conceptions of gay men and worked hard to distance themselves from the image of the camp gay man, the “mincing queen,” the Sex and the City gay best friend, the gay style guru . . .
```
described as subthemes of a broader theme of “managing gay identity.”

This phase involves the deep analytic work involved in thematic analysis, the crucial shaping up of analysis into its fine-grained detail. As analysis now necessarily involves writing, the separation between Phase 5 and Phase 6 is often slightly blurry. This phase involves selecting extracts to present and analyze and then setting out the story of each theme with or around these extracts. What makes good data to quote and analyze? Ideally, each extract would provide a vivid, compelling example that clearly illustrates the analytic points you are making. It is good to draw on extracts from across your data items to show the coverage of the theme, rather than drawing on only one data item (this can be frustrating when one source articulates it all perfectly—the analysis in Exhibit 4.4 quotes Asha because he expressed that part of the theme particularly well).

The extracts you select to quote and analyze provide the structure for the analysis—the data narrative informing the reader of your interpretation of the data and their meaning. In analyzing the data, you use it to tell a story of the data. Data do not speak for themselves—you must not simply paraphrase the content of the data. Your analytic narrative needs to tell the reader what about an extract is interesting and why. Throughout your analytic section, you would typically have at least as much narrative surrounding your data as extracts. Data must be interpreted and connected to your broader research questions and to the scholarly fields within which your work is situated. Some qualitative research includes this as a separate discussion section; other research incorporates discussion of the literature into the analysis, creating a Results and Discussion section. Both styles work with TA. An integrated approach works well when strong connections exist with existing research and when the analysis is more theoretical or interpretative. This approach can also avoid repetition between results and discussion sections.

Exhibit 4.4 shows part of the analysis of our theme “managing heterosexism.” It starts with a general summary of the theme’s core issue, and then expands on this by providing specific examples of different aspects of the theme, illustrated using brief extracts. Once sufficient detail has been provided to show the scope of the theme, the longer extract offers rich and evocative detail of what this actually meant for one participant. Analysis of that extract begins by highlighting some data features that provide the basis for our interpretation around a broader practice of minimization and individualization—a pattern across the data set. There is an interweaving of detailed and specific analysis of what happens in a particular data extract, and more summative analysis that illustrates the broader content of the data set in relation to the theme. This reflects our combination of two broad styles of thematic analysis: (a) descriptive, in which data tend to be used in illustrative ways, and (b) conceptual and interpretative, in which extracts tend to be analyzed in more detail, often for the latent meanings on which they draw. Both offer important analyses of data and serve different purposes, but they can usefully be combined, as we show. The latter can be a more difficult form of analysis to grasp because it moves from surface or apparent meanings to latent or implicit meanings; it can take experience to learn to see these in data.

Even when we present a lot of short extracts of data, however, seemingly reporting quite closely what participants said, the analysis always moves beyond the data. It does not just report words—it interprets them and organizes them within a larger overarching conceptual framework. Regardless of what form of TA is done, analysis uses data to make a point. Analysis needs to be driven by the question, “So what?” What is relevant or useful here to answering my question? This process of telling an analytic narrative around your data extracts needs to take place for all your themes. Each theme also needs to be developed not only in its own right but also in relation to your research question and in relation to the other themes. Conclusions can and should be drawn from across the whole analysis. So an analysis needs to make interconnections between themes and say something overall about the data set.

The other aspect of this phase is working out what to call each theme. Naming might seem trivial, but this short title can and should signal a lot. A good name for a theme is informative, concise, and
In common with others (e.g., Taulke-Johnson & Rivers, 1999), our participants described monitoring and assessing people and the environment for evidence of potential heterosexism, weighing up whether it would be safe to come and be out. They decided not to come out when people made overtly antigay comments. Asha, for instance, took the comment “one thing I just can’t understand is gay people” as strong evidence of a potential negative response to his coming out and chose not to. They made decisions to come out when people discussed gay-related issues in a broadly positive way, mentioned gay friends, or expressed “gay-friendly” sentiments (e.g., “want[ing] to be the ultimate personal fag hag,” Asha). This monitoring was sometimes a relatively passive process (“I just picked up tell-tale signs about it,” Asha); at other times, participants actively “test[ed] the waters” (David) and “tr[ied] and manipulate the conversation to head in that direction and see how to respond to it” (Asha). Asha described this rather evocatively:

Asha: just basically erm er, does he have a gay friend? Yes or no, is he alright with a gay friend? Yes or no. This person is alright to go out with- you know to come out with and basically if the answers are different the questions are different and the outcomes would be different . . . you’re just trying to you know answer all the questions to see what the outcome is and it’s kinda a bit of a headache

VC: It sounds exhausting, and stressful

Asha: It is, very much so but it’s kinda something that I have in the back of my mind . . . I find out you know which box they tick, which box they don’t tick and if they tick the right ones or if they tick the wrong ones I know what action to take from there . . .

VC: Yep yep, god that sounds very hard

Asha: Well the thing is it’s almost kinda- I wouldn’t, I don’t know it’s something that just happens in the background you know- I hardly notice it

VC: Yeah like this processing that going on and kinda churning away

Asha: Yeah all these things that you just happens that you’re not even completely aware of but it’s building up and you know you look back at it you see all these point and you say to my- you say to yourself right “I’m gonna tell this person I’m gay” “I’m gonna” you know and yeah

After initially agreeing with the interviewer, VC’s, assessment that this is an “exhausting stressful process” (“it is, very much so”), Asha described it as a more subconscious process, something he “hardly notice[d].” When VC again suggested it sounded “very hard,” he offered no agreement. Despite his detailed and vivid account, Asha appeared invested in framing this as a mundane rather than negative, and therefore “hard,” process. This “minimizing the negative” approach was common: The participants consistently framed phenomena that could be read as evidence of heteronormativity and instances of prejudice (Taulke-Johnson, 2008) as to-be-expected parts of normal life.

Asha earlier vividly described this process in a way that suggested it was negative yet implicitly located the problem within his own psychology rather than the environment:

Asha: constantly monitoring, keeping an eye out, keeping an ear out just you know, the little checklist this worst case- or not a worst case scenario but you’re having a list in your mind of all the possible things that can go wrong and you- you’re always going over that list of all the things that could go wrong I’ve kinda built- well personally for me it builds on my paranoia

In describing himself as paranoid, Asha suggests his response, rather than a heterosexist context, is at fault. All the participants interpreted difficulties they experienced in navigating a heterosexist world in this way. John, for example, associated his difficulties with coming out with his personality (he got embarrassed, and feared getting and looking embarrassed) rather than with the inherent difficulties that can exist around coming out (see DeCrescenzo, 1997; Flowers & Buston, 2001; Markowe, 2002) in heterosexist contexts. In internalizing their response to heteronormative contexts thus, responsibility for change is located within the participants, making it a personal rather than a political issue.

The degree to which students implicitly accepted responsibility for managing heterosexism to avoid “trouble” (David) by constantly modifying their speech, behavior, and other practices was the most striking feature of how they navigated the university climate. They had a strong sense that behaving or speaking in certain ways (being a “bad gay”; Taulke-Johnson, 2008) invited “trouble” and placed the onus on themselves to avoid it and protect themselves: “you have to sort of be very careful how you sort of came across to people” (David). The participants censored their speech and behavior (“tell . . . half of the truth,” Andreas); avoided coming out or making “overt” displays of homosexuality, such as by showing affection to a same-sex partner, being too camp and acting like “a mincing queen” (John), or wearing “obviously gay” clothing; and avoided certain people (“groups of lads,” John) and areas. Campus and city were seen as safe “as long as you took the measures—you know as long as you’re sensible about it you don’t go throwing it in people’s faces you don’t go down to you know places like [predominantly working class/non-White city suburb]” (Asha). [analysis continues]
catchy. The name “mincing queens” versus “ordinary guys who just happen to be gay” (see Exhibit 4.3) is memorable and signals both the focus of the theme—different ways of being gay—and something about the content of the analysis—that participants’ navigate between two different versions of being a gay man. “Mincing queens” is also a direct quote from the data. Using quotes in titles (also evident in Themes 1–3) can provide an immediate and vivid sense of what a theme is about while staying close to participants’ language and concepts.

Phase 6: Producing the Report
Although the final phase of analysis is the production of a report such as a journal article or a dissertation, it is not a phase that only begins at the end. Unlike in quantitative research, we do not complete our analysis of the data and then write it up. Writing and analysis are thoroughly interwoven in qualitative research—from informal writing of notes and memos to the more formal processes of analysis and report writing. The purpose of your report is to provide a compelling story about your data based on your analysis. The story should be convincing and clear yet complex and embedded in a scholarly field. Even for descriptive TA, it needs to go beyond description to make an argument that answers your research question. Good writing comes with practice but try to avoid repetition, paraphrasing, unnecessary complexity, and passive phrasing. In general, qualitative research is best reported using a first-person active tense but check the requirements for your report.

The order in which you present your themes is important: Themes should connect logically and meaningfully and, if relevant, should build on previous themes to tell a coherent story about the data. We decided to use “compulsory heterosexuality at university,” which documents the participants’ experiences of homophobia and heterosexism, as our first theme because these experiences, particularly the constant possibility, and fear, of heterosexism, shaped almost every aspect of the students’ university life and would be referenced throughout the rest of the analysis. From there, it made sense to discuss the participants’ experiences of managing heterosexism. We decided the two identity themes were the logical next step because the theme of coming out and being out closely related to the participants’ fear of heterosexism and the ways in which they managed their practices to avoid heterosexism. The second identity theme—which discussed different conceptualizations of gay identity and the participants’ desire to be perceived as ordinary guys who just happen to be gay—had a less immediately obvious connection to the first two themes but linked well to the first identity theme.

DOING THEMATIC ANALYSIS WELL
These guidelines lay out the process for producing a good TA that is thorough, plausible, and sophisticated. But like any analysis, TA can be done well, and it can be done poorly. Common errors include providing data extracts with little or no analysis (no interpretation of the data that tells us how they are relevant to answering the research question) or simple paraphrasing or summarizing data (see Braun & Clarke, 2006). Using data collection questions as themes is another common error—themes are better identified across the content of what participants say rather than via the questions they have been asked. “Incidents of homophobia” would be a weak theme, for example, because it would involve simply describing different things participants reported in response to an interview question on the topic. “There’s always that level of uncertainty: Compulsory heterosexuality at university” is a much stronger theme because it captures something more complex about how the participants’ constant fear of homophobia and heterosexism shaped their university lives. It also incorporates data from across the whole interviews not just responses to specific questions about homophobia and heterosexism.

On a different level, an analysis can be weak or unconvincing if themes are not coherent or try and do too much. Analysis can also suffer from lack of evidence. You need to provide examples of, and analyze, enough data to convince the reader that this pattern you claim really was evident—consider the balance of data and analysis in Exhibit 4.4. A TA does have to relate to patterns found across your data set. This does not mean every data item has to
evidence each theme, but it has to be more than idiosyncratic. Finally, TA can suffer because of mismatches between the data and analysis, or between the form of TA done, and the theoretical position of the report (for more discussion of these and for a checklist for doing good TA, see Braun & Clarke, 2006). In developing and revising your analysis, make sure data-based claims are justified and that the claims fit within your overall theoretical position (e.g., whether you are using an experiential or critical form of TA).

References


