

At the end of Chapter 2 I raised the question, ‘what does one think one is doing when conducting EFA?’ I offered two possibilities, which I referred to as, i) a ‘*realist*’ answer; and ii) an *interpretivist* answer. The interpretivist response claims that EFA reduces the number of variables in order to tell a meaningful story, which would be a challenge if we were to retain all of Sanderson’s original 70 measured variables. The realist interpretation can perhaps be analysed into two versions, the first of which asserts that we are identifying the structure that underlies the data that we have and no more. This would be an *inductive* understanding of what we are doing: inducing the structure of the data. The second version, however, goes further, proposing that we are discovering something about the world that has generated the data, so that we are seeing our survey as a window onto something that is beyond itself. This second version is perhaps more appropriately referred to as realist, hence the quotes around this term when referring to both versions of i) above. For the moment, I’ll refer to the first version of i) as *data-grounded* to distinguish it from the *realist* version.

If we adopt the data-grounded interpretation of the research then what we have discovered is the pattern of responses by a particular collection of individuals to the survey instrument on the particular occasions and in the particular circumstances of its administration. This may be too limited a claim to persuade people to pay much attention to the work and so we are more likely to argue that our sampling and data collection and analysis strategies entitle us to claim that the structure that we have identified is representative of the pattern of attitudes in a wider population and under more general circumstances and this has implications for policy and practice and so forth. The interpretivist version—ii) above—offers an interpretation of the pattern of schoolchildren’s attitudes to the aesthetic dimensions of dance and asks: ‘suppose you look at it like this, what follows, how does it inspire you to act?’ The realist and interpretivist versions perhaps seem more viable, but I don’t want to dismiss the data-grounded approach just yet.

Indeed, the data-grounded approach is my interpretation of Barney Glaser’s intention in his oft quoted dictum, ‘all is data’ (Glaser, 2007). Glaser is one of the two originators of Grounded Theory (GT), the other being Anselm Strauss (Glaser & Strauss, 1965; 1967). Whilst people generally do not say it to my face, I hear quite a few reports of academics announcing that they don’t like Grounded Theory (GT). Why they won’t tell me personally is not entirely clear as I do not claim to be a Grounded Theorist as such and would not be accepted as one by Barney Glaser: Glaser has claimed that in order to do GT one should have been mentored by a Grounded Theorist. I was not so mentored (my doctoral supervisor was Basil Bernstein who certainly was not a Grounded Theorist). Indeed, Glaser’s claim establishes a necessary lineage of all GT back to its originator, which is to say, to Dr Glaser himself or perhaps his co-originator, Anselm Strauss, though Strauss may have been disqualified as a result of the separation of the ways of the two authors of *The Discovery of Grounded Theory: strategies for qualitative research* (1967. New York. Aldine Publishing Company). The reasons for the rejection of GT seem primarily to include Glaser’s dismissal of the preliminary literature review, but perhaps also his outright condemnation of ‘qualitative data analysis’ (QDA) on the basis of a range of positions that he attributes to QDA and which he claims are irrelevant to GT. Several chapters in the anthology by Martin and Gynnild (2011) refer to the interdiction on the preliminary literature review with Guthrie and Lowe going so far as to suggest that if university protocol insists on a preliminary literature review then the student should produce “a logically plausible (but quite irrelevant) literature review” (p. 61)!

In his chapter, Glaser presents the reason for rejecting the preliminary literature review is to avoid preconceptualising the data. Judith Holton reports an anecdote from Alvin Gouldner that had been recounted by Glaser. The anecdote concerned a student interested in risk-taking behaviour among steeplejacks. The student had found no evidence that seemed to relate risk-taking until, one day, he observed a group of steeplejacks drawing straws. The point of the competition, however, was not for allocation to the least or most risky position on the scaffolding, but for the allocation of the most advantageous position for the purposes of window peeping! Had the student produced a literature review on risk-taking, he would have wasted his effort as the study turned out to be about strategic positioning.

Now I think one needs to appreciate the significance of the 'all is data' dictum in order to understand Glaser's rejection of the preliminary literature review. After all, we should have learned to adopt a critical approach to antecedent research and to avoid being unduly influenced by it. Furthermore, If I was supervising a student interested in researching steeplejacks I would advise them to review the social research on steeplejacks and not to concentrate on one presumed aspect of their activity. This advice, however, is based on the assumption that the student was setting out to study steeplejack culture and such predetermination is not appropriate from a data-grounded—what, in my interpretation, Glaser refers to as a 'grounded theory'—perspective. I am using the expression 'data-grounded' because the expression more frequently used—though not Glaser's elaboration of it—ellides just what the theory is grounded on: it's grounded on the data: 'all is data'! So, using this approach, we can't know what our research is about until we start collecting the data on which our theory will ultimately be grounded, so we cannot lay out the parameters for a literature review. These parameters will derive from the principal interests of the subjects of our study. This being the case, we should avoid leading them in any particular direction, so, if we are conducting an interview, we might begin by asking the subject to describe a recent day—an example of a *grand tour* question—and their principal interests or concerns will emerge in their response and/or in our use of *probes* to generate more information ('can you tell me a bit more about that', 'is that the way it usually happens?', and so on, but generally avoid *prompts*—suggesting responses). Try to focus on specific instances of action rather than asking for generalisations.

Glaser (1998) advises against audio recording of interviews; it creates a false sense of security and delays analysis, which should begin at the same time as the start of the first interview. I have more to say about analysis (as does Glaser), but at this point I want to depart from a rigid adherence to grounded theory, primarily because, in my experience, most researchers are keen to initiate their studies with a particular interest in mind. This disposition would also be important if one were applying for research funding or submitting an application, one of the requirements for which is a research proposal. I am not denying that genuine GT is capable of producing exciting and original and indeed useful research, but I think it's not for everyone. That having been said, we can and should learn a great deal from GT method: this does not entail regarding it as legislation. So, in what follows I will introduce my principles for qualitative data production and analysis acknowledging GT as appropriate, but also marking my deviations from it where necessary. My references in this respect are predominantly to 'classic GT' as advocated and developed by Glaser rather than the alternatives including the perhaps more widely cited version presented by Strauss and Corbin (eg 1980) In the end, the researcher is obliged to produce an argument in support of the claims that they make: the research and methodology as legislation rule seems to be quite

prevalent with citations attached to claims without further comment as if that settled the matter: it does not.

Approaching Data Collection and Analysis

I'll start by clearing up one minor point. At the head of his article, 'All is data', Glaser (2007) asserts that 'Although data is plural "is" sounds better' (no page nos). 'Data' is a *Latin* plural, the singular of which is *datum*. In English the word is generally used as a plural in quantitative research; in qualitative research it is more often used as a mass noun, which does not take a plural. Some publishing houses in the US apparently prefer the plural; I haven't come across such a preference in the UK. The examiner of one of my students failed to appreciate the use of 'data' as a mass noun and, apparently without checking a dictionary (OED or Merriam-Webster), required the student to change the number of all verbs associated with instances of 'data' to plural; there's no accounting for ignorance!

I've put data (mass noun) collection and analysis together because, with Glaser, I believe that in qualitative research the two processes should run together: analysis starts as soon as data collection starts, though is likely to continue well beyond the termination of the latter. In contrast with Glaser, however, I prefer the use of audio recording (don't forget to take spare batteries/battery charger or a fully charged phone, which I did—only once!) and transcription of interviews together with field notes, written during or as soon as possible after the interview (two interviewers helps in this respect particularly with group interviews, where video recording aids transcription). Again, this is because I regard research as a transaction between the researcher's agenda and the setting, whereas classic GT backgrounds the former.

Analysis of any kind of qualitative data in exploratory research should appropriately begin with *open coding*, by which I mean assigning a label—a *code*—to sections of the data to identify their meaning or significance as interpreted by the researcher. Keep going line by line of the transcript. After a while, go back and compare the codes and compare the codes with the data, which will enable you to develop the coding and reduce the number of codes. This *constant comparison* process should continue throughout the data analysis. Start with one interview or observed action or section of text or perhaps start with two sections of data that you think contrast in some way, the precise nature of which is to be determined: I find the use of contrasting segments assists in the formulation of codes. Write a *memo* for each code to elaborate on the meaning that you've given it. This process is facilitated by the use of qualitative data analysis software, such as HyperRESEARCH or Nvivo. You will need to import the data—which can be text or image or audio or video or pdf—into the app. You will then be able to select, label and annotate (memo) sections of the data and easily retrieve the codings. Note, however, it is still you doing the analysis; the machine won't analyse your data for you as statistics apps do (though even with stats apps you will still have to interpret the results).

Having performed preliminary analysis of your first interview(s) or section(s) of data, ask yourself, 'where should I go/what should I observe/whom should I interview next? This GT strategy is *theoretical sampling*. In my opinion it is generally a productive way to build a sample in response to your developing analysis that does not restrict you to one form of data. It is, however, not always going to be possible. If you are conducting research in a school, for example, you may find that the principal will want to know details of your proposed sample at the outset. This is their call, not yours, but you might try persuading them of the value of theoretical sampling and promise to keep them apprised of your activities.

The data collection, open coding, memoing and theoretical sampling continue in parallel until you reach a point at which one or perhaps more than one code (one code in GT) stands or stand out as of central significance: in GT this is your *core category*, in my own work and that of some of my students we prefer to dimension our analysis, so develop multiple core categories—more of this (SAM) later. Subsequent to this the process continues by concentrating on those codes that relate to the core category/ies (*selective coding*) until the codes become *saturated*, which is to say that no new codes are emerging. At this point data collection is, for the time being anyway, complete and analysis proceeds with *theoretical coding*, whereby one develops relations between the codes and, in particular, between the core category and the other codes. The memos that will have been developed during the analysis can be drawn on in producing the write-up of the research. Figure 3.1 illustrates the sequencing of the stages of data collection and analysis that is broadly consistent with classic GT, though the diagram is mine.

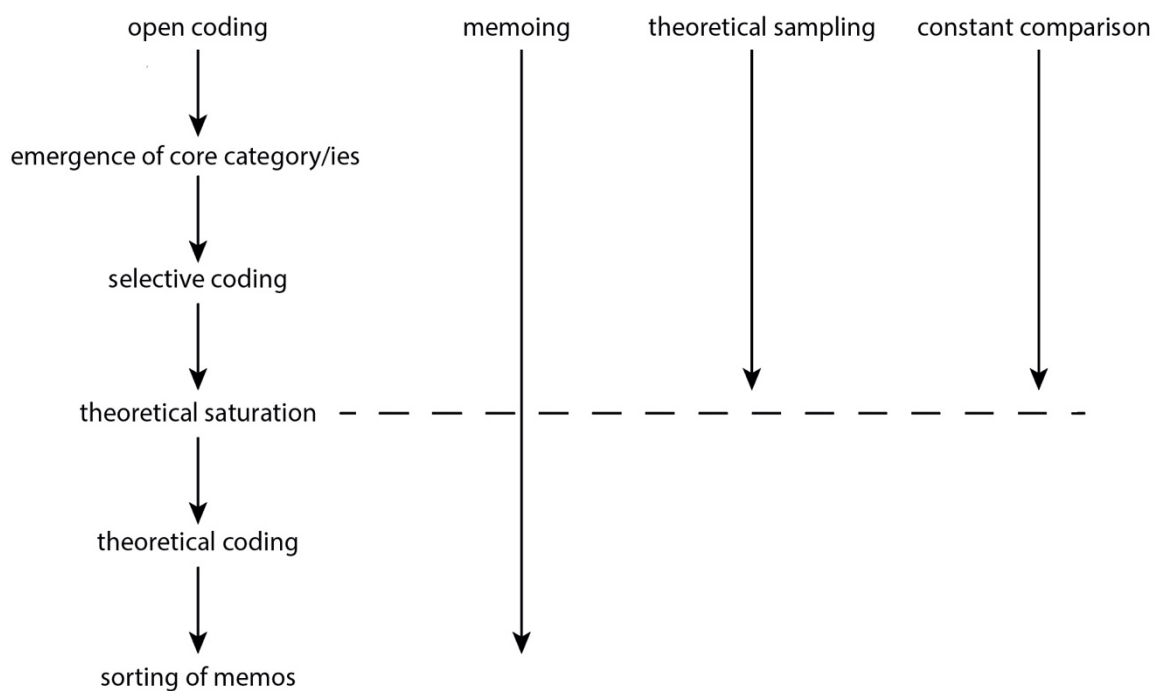


Figure 3.1: Data collection and analysis schedule

Illustrating qualitative analysis: music reviews

Articles reporting qualitative research in general and GT studies in particular are often parsimonious in respect of the inclusion of details of data analysis, so that one can see what they have achieved, but rarely how they have achieved it. Furthermore, the raw data is, of course, generally not available to the audience. Some time ago I was asked by a group of my doctoral students to present an illustration of how constant comparison works. I took advantage of the 12-hour flight from Tokyo to London (via Frankfurt, as it happens) to generate an illustrative analysis of four music reviews that I had found in the online version of *The Guardian*. The reviews were:

- ‘Sinikka Langeland/Trio Mediaeval: The Magical Forest review—a stream of entrancing sound’ (<https://www.theguardian.com/music/2016/aug/25/sinikka-langeland-trio-mediaeval-the-magical-forest-review>)
- ‘Ray BLK’s BBC Music Sound of 2017 win recognises diversity of British pop’ (<https://www.theguardian.com/music/musicblog/2017/jan/06/ray-blk-bbc-sound-of-2017-diversity-british-pop>)
- ‘Rag’n’Bone Man wins 2017 Brits critics’ choice award’ (<https://www.theguardian.com/music/2016/dec/08/ragnbone-man-wins-2017-brits-critics-choice-award>)
- ‘Sleeping Giant: Hand Eye CD review – brilliant performances of exciting new commissions’ (<https://www.theguardian.com/music/2016/mar/30/hand-eye-review-eighth-blackbird-sleeping-giant-timo-andres-andrew-norman>)

These reviews constitute the raw data. They are all still available on *The Guardian* website, which edits, but does not seem to delete anything. I should perhaps mention that the selection is not representative either positively or negatively of my musical preferences, though I’m always impressed by Eighth Blackbird recordings and by their choice of name, which derives from the eighth stanza of Wallace Stevens’s poem, ‘Thirteen ways of looking at a blackbird’ (<https://www.eighthblackbird.org/about/ensemble/>):

I know noble accents
And lucid, inescapable rhythms;
But I know, too,
That the blackbird is involved
In what I know

(<https://www.poetryfoundation.org/poems/45236/thirteen-ways-of-looking-at-a-blackbird>)

Resonant with my approach to data analysis!

Here is my first set of codes (remember, I’m analysing the reviews in *The Guardian*, not the music itself:

composer of music	general diversity
author of performance	genre
source	composition antecedents
biography of performer	performance antecedents
biography of composer	closed technical language
principles of composer	open technical language
principles of performer	metaphorical language
celebrity	evaluation
origin	authenticity
interpretation	awarding institution: evaluator/evaluatee
composition	response to award
performance	collaboration
local diversity	

Scrutinising these codes, combining some, changing the terms where necessary resulted in a reduced set:

composition/composer
performance/performer

biography/principles

reference set: composers: closed/open
reference set: performers: closed/open

language: general: academic/popular—genre
language: local: technical (I+)/figurative (I-) xd

reference context

diversity: general/local

As you can see, some of these codes are now scaled, showing two properties. Continuing with constant comparison yielded. The following further reduction:

3. What is being reviewed?

Composition/Performance as action

- actor (composer/performer)
- history: musical/extra-musical
- celebrity
- quality
- originality

2. How is the review formulated?

Language

- C/F
- I+/I-

The properties under 'language' are theoretical concepts. *C* refers to Bernstein's *classification*, which is concerned with the distinction *between* categories, *F* refers to his *framing*, which is a measure of control *within* categories. I subsequently rejected these concepts on the grounds that they effectively refer to the same variable at different *levels of analysis* (for elaboration of this argument see Dowling, (2009), pp 76-81). I replaced both with *I*, the strength of *institutionalisation* of a practice, which is to say:

... the extent to which a practice exhibits an empirical regularity that marks it out as recognisably distinct from other practices (or from a specific other practice).

Dowling, 2009. P. 81

The empirical regularity is clearly 'within' the practice, but it is this that establishes or maintains the distinction 'between' practices. This simplification of the coding of language, now as strongly/weakly institutionalised (I+/I-) allows me to conceptualise its coding in respect of each of the categories, *composer*, *composition*, *performer*, *performance* in terms of the scheme in Figure 3.2.

Discourse	Level of Analysis	
	Between	Within
I ⁺	<i>generic</i>	<i>musical</i>
I ⁻	<i>idiosyncratic</i>	<i>figurative</i>

Figure 3.2: Coding of composer, composition, performer, performance

This mode of relational analysis is a characteristic of *Social Activity Method (SAM)* (Dowling, 2009, 2013, see also Whiteman & Dudley-Smith, in press, which marks out the distinctions between this mode and other relational schemes). It operates on the basis of taking the cross-product of two mutually independent, binary variables, giving rise to four properties. Because this scheme codes the four categories in the title of the Figure, the method produces four schemes that can potentially be used to analyse the original data. It is important to point out that the properties are not boxes into which the music reviews are to be inserted. Rather, it is expected that each music review will deploy more than one and potentially all four of the properties. In conducting the analysis one must find a *lexia* (a section of the text or image) that can be coded uniquely in terms of one of the four properties in Figure 3.2. Each music review can then be described in terms of the distribution of these properties that it deploys and the reviews compared in terms of these distributions. It may be considered helpful to quantify these distributions, but the researcher should be aware that the *lexia* on which they are based are unlikely to be equivalent in terms of magnitude or significance, so care must be taken in this form of quantification.

It will be apparent that the coding in Figure 3.2 does not exhaust the codes in the earlier lists and it is of course possible to extend the analysis to describe the nature and distribution of 'what is being reviewed'. The work here, however, is intended to introduce an approach to analysis rather than to present a complete (theoretically saturated) analysis of the texts. The example that follows is perhaps closer to theoretical saturation.

Martin's Story: a second illustration

Here I want to present another example of textual analysis. These are the principles that will direct the analysis.

1. The text must be bounded, constituted as a complete object rather than as a field for data collection.
2. Ask: what is put together; what is kept apart: look for oppositions and alliances in the text.
3. Look for trajectories in the text.

These principles will enable the description of *what* is being established in the text and *how* is this being achieved, corresponding to the two questions that were being addressed in the analysis of the music reviews above.

The object text here is a double page spread from an old (c. 1996) edition of the *London Zoo Guide* that is titled 'Martin's story'. It is divided into an upper section, on a white background, and a lower section, which is about twice the height, on a pale green background. The upper section contains verbal text on the lefthand side; this text is headed: 'London Zoo

holds the European studbook for Sumatran tigers and co-ordinates the breeding programme' the text reads (I've retained the line breaks):

There were eight subspecies of tiger, but 3 of these subspecies are already extinct. Tigers are seriously endangered in the wild, partly through disappearance of their habitat, but particularly because they are hunted for parts of their body. Tiger bones, penises, gall bladders and eyes are highly prized as medicines in China.
Martin is a Sumatran tiger. With a population of only 400, it was very important that we breed from him.

Below this is the page title, Martin's Story, in green capitals.

On the righthand side of the upper section is a black and white drawing of a walking tiger in profile and labelled. This drawing has a title at the upper left:

How tiger parts
Are used in
oriental
medicine

The drawing is labelled with the various uses, for example:

Carry a claw in your pocket or wear or as a piece of jewellery and you will possess courage and be protected from sudden fright.

And:

Mix the brain with oil, rub on your body and you will be cured of laziness and acne.

The lower section begins immediately below the page title. On the left is a large colour photograph of a tiger's (presumably Martin's) head and the front part of his body. He is looking out of the frame, slightly to the left and not directly at the viewer. The text on the upper left of this section states:

We were able to incorporate Martin in the European breeding programme because the studbook showed that he was pure bred Sumatran. A Sumatran tiger with Siberian or Bengali ancestors would have been a hybrid and therefore unsuitable for breeding.

Martin's image partially penetrates the righthand side of the lower section, which shows a colour photograph of Martin engaged in tigerish copulation, apparently with Mira (see below). Above and to the right of this photograph is the following text (line breaks altered);

The next stage was to select the perfect mate. We searched the studbook for a female of the right breeding age who was as distantly related to Martin as possible. The computer is able to give each tiger a "mean kinship number" which indicates each individual's degree of relatedness to every other tiger on the programme. From all the information available, the species coordinator selected Mira, a 7-year old from Bremmerhaven Zoo in Germany.

Mira flew into Heathrow in May 1993. London Zoo paid the costs of transporting her, since we were the zoo receiving an animal on loan.

For Martin and Mira it was clearly love at first sight and the Zoo was delighted when they produced Hari in January 1996.

Below this image and text is the following text above and to the left of another colour, group photograph showing a white male with a group of Indonesian males captioned:

London Zoo staff taught tiger husbandry to Indonesian staff.

The text is:

CONSERVATION ^{IN} ACTION — in the wild

London Zoo staff are currently involved in the development of a global strategy for the conservation of Sumatran tigers in the wild and in captivity. In 1992, they attended a captive breeding workshop in Taman Safari Zoo in Java, where wild tigers which would previously have been shot for attacking livestock of people, now form an important founder group to supply zoos with fresh genetic material.

At a further meeting in Sumatra, a database of the wild tigers living in the five reserves on the island was set up and many recommendations were made regarding improving protection, funding and staffing of the reserves.

The most obvious contrast here is between the tiger images at bottom left and top right. The colour photograph exhibits high modality, what I referred to as a strong *visual code of presence* (Dowling, 1998), it is a *real* tiger, which contrasts with the diagrammatic drawing in the top right, not a *real* tiger, because it has been dismembered by its labelling and reportedly in its recruitment in 'oriental medicine'. The 'real' tiger, by contrast is not only real, but anthropomorphised in its ('his'?) love story with Mira. A real tiger, but not a threatening tiger: he's looking gently away from and not directly at us, not about to pounce. This is how we (Europeans?) see the tiger, contrasting with the response of Indonesians, who shoot it and Chinese, who chop it up for medicine and jewellery.

The text at the top left of the page establishes that it is not the Indonesians or Chinese who are in danger, but the tiger as a species, represented here by Martin (and family). Two reasons for this endangerment are mentioned. The first, the disappearance of their habitat for which no subject is presented. The main problem, however, appears to be that tigers are

hunted for body parts by or on behalf of the Chinese. Move over to the right and the diagram of a tiger and we are given illustrations of what is presented as mythical medicinal use. This practice is ridiculed in particular by the, presumably imaginary, caveat appended to one of the illustrations:

To cure a fever caused by ghosts, sit on a tiger skin. Caution—if used too often, you may turn into a tiger!

The hyperbole used in other illustrations adds to this ridicule:

Eyeballs rolled into pills are a sure cure for convulsions

This ridiculing of an ‘oriental’ practice contrasts with the rational science of conservation, which involves the maintenance of a studbook to enable to avoidance of hybrids and breeding between animals that are too closely related. The principal subject of conservation in this text is London Zoo that ‘holds the European studbook for Sumatran tigers and co-ordinates the breeding programme’ and, furthermore, teaches Indonesian staff about tiger husbandry. The Zoo Guide attracts us to conservation practice through the image and anthropomorphising of a beautiful living animal that is apparently under threat and repels us from its opposite via the ridiculing of ‘oriental’ non-science, also contrasted with occidental rationality, a rationality the aim of which—conservation—is never justified, but presented as a taken-for-granted good: I should emphasise that I am not arguing *against* conservation, merely pointing out that this text, as I have bounded it, does not include any case *for* it. The principal opposition in this text is between the oriental, responsible for irrational destruction and the source of the tiger’s problems, on the one hand, and the occidental, responsible for the solution, on the other. There is, however, another narrative present in the text—or, rather, in my reading of it—which is a eugenic narrative: species purity is good, hybridity is bad, again, not explained.

Moving beyond the text as I have bounded it: as I write, we are in the midst of a global coronavirus pandemic. Whereas London Zoo scientists are energetically trying to conserve sumatran tigers as a species, other scientists—including colleagues at my own institution—are engaged in urgent attempts to eradicate the virus, to render it extinct. If they succeed, they will undoubtedly be universally applauded as heroes. Yet the coronavirus is arguably a beautiful entity, as is evidenced by its images online. (see <https://unsplash.com/s/photos/coronavirus> for example). I have to say the if tigers routinely paraded around Russell Square (or South Dock, where I am currently practising social isolation), mauling passers by, I would certainly side with inhabitants of Java, who shoot tigers to preempt their attacks. The coronavirus, however, is indiscriminate, having the temerity to attack Europeans as well as Indonesians and Chinese; no wonder London Zoo is not concerned with its conservation!

A Grounded Theory Example

The two illustrations above are examples of text analysis, but what might be the point were they to be considered more than pedagogic examples? The analysis of music reviews has resulted in a scheme in Figure 3.2 that might be recruited in the analysis of further texts and settings, though without forcing itself on them. Thus it is a contribution to theory, which is

one of the ways in which qualitative data analysis can generalise. That it is a two-dimensional analysis of four musical categories—*composer, composition, performer, performance*—illustrates my motive for avoiding stopping the analysis at a single core category as is recommended in classic Grounded Theory. It is not appropriate to consider either of the dimensions in Figure 3.2 to be fundamental to the other, nor to consider any of the four categories above to have priority. Each application of the scheme to one of these categories presents a particular viewpoint, raises the challenge: *suppose you look at it like this*.

The analysis of ‘Martin’s story’ operates in a different way. It reduces the text to a hybrid of two narratives. One of these narratives constitutes a racist hierarchy of occidental preservation over oriental destruction, the other constructs a political hierarchy of eugenic conservation of the purity of species over contamination and hybridity, sitting uncomfortably, perhaps, in the society of the scientific descendants of Charles Darwin. This analysis contributes to the accumulation of cases, which is the second way in which qualitative data analysis can generalise. The raw data for each of these analyses is readymade.

A substantial component of a good deal of research is data collection, which is rendered trivial by the use of readymades. A more common source of data for qualitative research might be termed the empirical setting. I want to describe a classic GT study by Tina Johnston (2008). Johnston was interested in:

... the issues surrounding the various ability grouping models in middle level mathematics classes, hoping to better understand the practice and problems associated with it.

(Johnston, 2008, 44)

Her research setting comprised middle schools in the Northwest United States, which clearly cannot be constituted as a readymade: it’s far too extensive and it’s dynamic. She needed to sample the setting and collect and analyse data from the sample. A theoretical sampling was adopted. Little information is provided on this, though:

If most of the interviewed parents had children in the high mathematics group, the researcher sought parents with children in at/below grade level classes for contrast. As parents mentioned others who were actively working towards advanced middle level mathematics placements, the researcher contacted them for interviews. Since the same problem developed among several schools in one part of Oregon, the researcher interviewed teachers from wider regions within Oregon and the Northwest United States. Literature was collected and integrated representing research projects across the United States, Canada and United Kingdom as well as from multi-national studies. Finally, as parents with ties to India, China, Korea and Mexico were interviewed, the researcher requested information about how students were grouped in mathematics in those countries.

(Ibid.; 45)

No details are given on how the information about foreign school practices might have informed the analysis, which, in the article, concentrated on interviews with:

Forty-one subjects who worked with or had children attending 13 schools in 10 school districts and three states were interviewed. Sixteen teachers, three administrators, and 23 parents were interviewed.

(Ibid.; 44)

The numbers in the extract above don’t quite add up correctly, but never mind.

As I’ve indicated, classic GT does not begin with a tightly specified research question, because it seeks to identify the principal concerns of subjects in the setting, rather than

imposing its own priorities. However, Johnston did direct her interviewees to her general area of interest rather than using a more open, ‘grand tour question’:

Subjects were informed of the topic of the project. If administrators or teachers did not immediately begin to share their thoughts on ability grouping in middle level mathematics, the researcher asked them to discuss their schools’ mathematics ability grouping arrangement ...
(Ibid.)

Johnston proceeded with the analysis, much along the lines of the scheme in Figure 3.1, ultimately identifying the ‘main concern’ of the subjects as:

A lack of trust in general and, more specifically, a lack of trust in the school ability grouping filter system were identified as the main concern voiced by subjects when asked about their experiences of the substantive area.

(Ibid.; 46)

The researcher named the response to this ‘main concern’ as ‘pushing’:

Pushing is a way for parents and educators to deal with the lack of trust and fear and/or disappointment in the student placement process into mathematics ability groups.

(Ibid.)

This *core category* exhibited three stages: ‘investing’, ‘pressuring’ and ‘lobbying’, each of which had three properties as shown in Figure 3.3, which is my representation of Johnston’s theory.

Substantive GT	Stages	Properties
<i>pushing</i>	<i>investing</i>	<i>personal tutoring</i>
		<i>purchased tutoring</i>
		<i>classroom volunteering</i>
	<i>pressuring</i>	<i>rallying support</i>
		<i>pressure on teachers</i>
		<i>pressure on administrators</i>
	<i>lobbying</i>	<i>positioning</i>
		<i>policy changing</i>
		<i>systemic changing</i>

Figure 3.3: Pushing for Privileged Passage (a Grounded Theory)

It occurs to me that the interpretation of most of the properties—which explain the stages—is straightforward apart, perhaps, from ‘positioning’, which is glossed as ‘positioning [oneself] on school or district committees’ (Johnston, 2008; 50) and this enables influence in respect of policy and systemic change.

I consider that the decisions on the naming of categories are very far from arbitrary and can be quite difficult to achieve. It is important to avoid using terms for theoretical categories that one wants to use routinely in natural language. In addition, one need to avoid using the

same term more than once: my own sociological language as represented in Dowling (2009) contains approximately 200 technical terms and quite a few more have been added by myself and others, so inventing new ones becomes increasingly tricky. In my opinion, the term 'pushing' used by Johnston is perhaps an unwise choice for a different reason. The word was chosen, it seems, because it came up frequently in interviews. In this context it is an empirical or *emic* term, a word used by participants in the empirical setting. The trouble with this is that it carries with it a diversity of denotations and connotations: emic words always mean more than one thing. Theoretical or *etic* terms, by contrast, should be defined uniquely, which precision may tend to be confused by the recruitment of an emic word. Of course, one solution would be to invent neologisms for all theoretical categories. Unfortunately, this would ultimately turn a research article's language into idiolect and render it effectively unreadable. So this is simply a warning to take care over the naming of one's codes.

The theory illustrated in Figure 3.3 is a substantive Grounded Theory. I satisfies both modes of generalisation of qualitative analysis that I have mentioned: it adds both to theory and to the accumulation of cases. In respect of the former, there is a further stage in classic GT, the potential development of a substantive GT into a formal GT. Johnston mentions the situation of 'powerless populations', involving people being cared for or requiring assistance from persons who have or take on the role of guardians who may *push* for the privileged passage of their wards or those for whom they take on concern. Johnston also mentions the area of athletics and sport generally where individuals may be *pushed* to develop their or others' performances. The development of a formal GT proceeds, not necessarily by further data collection and analysis, but perhaps by reviewing available research literature on activity in related fields.

My analyses of the music reviews and of 'Martin's story' are clearly data grounded, grounded on readymade data. Johnston's theory is grounded in her own data, but the empirical setting is open both in respect of mathematics group placement in middle schools and in terms of the development into other fields, which would constitute a formal Grounded Theory, but may not still be data-grounded in the manner in which I introduced this term at the start of this chapter. Data-grounded analysis makes no claim to speak for anything that is beyond the specific data (mass noun) that is analysed, although, to the extent that it constitutes a case and to the extent that it generates theoretical language, it may still generalise.