A reflexive exploration of two qualitative data coding techniques

Erik Blair
Royal College of Physicians, United Kingdom

In an attempt to help find meaning within qualitative data, researchers commonly start by coding their data. There are a number of coding systems available to researchers and this reflexive account explores my reflections on the use of two such techniques. As part of a larger investigation, two pilot studies were undertaken as a means to examine the relative merits of open coding and template coding for examining transcripts. This article does not describe the research project per se but attempts to step back and offer a reflexive account of the development of data coding tools. Here I reflect upon and evaluate the two data coding techniques that were piloted, and discuss how using appropriate aspects of both led to the development of my final data coding approach. My exploration found there was no clear-cut 'best' option but that the data coding techniques needed to be reflexively-aligned to meet the specific needs of my project. This reflection suggests that, when coding qualitative data, researchers should be methodologically thoughtful when they attempt to apply any data coding technique; that they do not assume pre-established tools are aligned to their particular paradigm; and that they consider combining and refining established techniques as a means to define their own specific codes.

Key words: coding; pilot; qualitative data; reflexivity

For many involved in analysing qualitative data there comes a time when they look at their raw data and ask themselves, "What does all this mean?" The 'answers' that lie within their data are not always explicit and methods need to be employed in order to extricate these 'answers'. Robson (2002, p.387) suggests that, before we can interpret our findings, the messages that lie within the data "need careful teasing out", and Morgan and Krueger (1997) report that the tools that are used for analysis and reporting should not be left unexamined and should be scrutinized in the early stages of the research process. It was with this in mind that I decided to pilot-test two coding systems for transcribed data that had been produced during two focus groups. This article discusses how these two methods of data coding—open coding and template coding—were piloted. Here I reflect upon and evaluate the two methods and discuss how using appropriate aspects of both led to my eventual data coding scheme.

The research project that produced the focus group data and my explanation herein are grounded in my interpretivist view: understanding the world through my interaction with others; acknowledging my dynamic relationship with the data, and accepting my place within the research (Greenbank, 2003). Such a perspective addresses the concept of interpretation and subjectivity from the start and does not pretend to be purely "objective" or classically "scientific". Heshusius (1994) suggests that researchers should embrace a

participatory consciousness — recognising that they are not separate from the world in which the data are produced. Here, the qualitative tradition encourages researchers to be aware that they are uniquely situated yet open and receptive to the thoughts of others. This subjectivity reflects a holistic epistemology that replaces the traditional relationship between "truth" and "interpretation" through recognising the productive relationship between all the players in a research project, and in my case this relationship is between myself and the focus group data. Therefore the validity of my results becomes a question of hermeneutics as I interpret my findings according to my position. In exploring the data developed from the two focus groups, my approach was reflexive in that there was constant scrutiny of the relationship between the researched and the researcher through "immediate, continuing, dynamic and subjective self-awareness" (Finlay, 2003, p. 108) something that Lynch (2000) calls being methodologically self-conscious. Being methodologically self-conscious means that the qualitative data analyst should be aware that their particular perspective is likely to influence their choice of coding method—but, in qualitative data analysis, this is not generally thought to be a bias that needs to be 'corrected' rather it is seen to be beneficial that the analyst is able to use their own unique skills, talents and expertise.

Pallas (2001) suggests that it is important for researchers to have an epistemological framework and proposes that addressing this is central to educational research; therefore, before the research project, I took time to reflect upon my epistemology and ontology. In trying to define "epistemology" and "ontology" I was faced with some paradigmatic issues in that how I perceive the world affects how I understand these terms. If I were to see the world and knowledge as fixed entities awaiting discovery I could clearly have defined these terms—as they must have some meaning that is connected with an absolute concept. If I were to have adopted such a position, I could have used a "dictionary" approach and held that epistemology is the study of knowledge and that ontology relates to the nature of reality. Maintaining an interpretivist paradigm, where a person's relationship with the world creates meaning and understanding, means that epistemology is a less tangible term, perhaps one that is connected to the relationship between the researcher and the object under research (Bettis & Gregson, 2001). Ontology in this instance becomes a question about how I use my reflections to construct my understanding. Such a process is validated by the central facets of reflexivity—constant internal scrutiny and questioning. Peshkin (1988) refers to such internal scrutiny as performing a reflective "audit" that allows researchers to explore and acknowledge their particular form of subjectivity and I found that my research in this instance was ontologically subjective in that, at all times, I was forced to question what it was that constituted the identity of the data. At the beginning of the process, I wanted to find a system that would allow me to extricate and interpret the data in

a reliable and useful way; however, taking a reflexive position throughout the piloting process, led me to understand Drake's remark that interpreting research "is not a matter of looking harder or more closely but of seeing what frames our seeing" (2010, p.88).

The Data Source

The research project itself is not the focus of this article but, briefly, it involved gathering the perspectives of university lecturers regarding the value of attending various in-house professional development workshops. The research also sought to relate the participants' responses to key literature in the field of learning theory. Twenty two lecturers were involved in the project; all of them had taken part in an individual semi-structured interview and a focus group discussion. The 22 participants were split into two focus groups each of which lasted around 15 minutes. The semi-structured interviews and the focus groups were recorded and the data transcribed verbatim. I found myself with a lot of qualitative data and needed a method to examine it. As my two areas of data were textual I decided that content analysis would be the most likely analytical tool. Krippendorff (2004, p.18) suggests that content analysis is "a research technique for making replicable and valid inferences from texts" therefore I decided to pilot-test two content analysis techniques on a sample of my data in an effort to establish which would be likely to yield the "best" results. For the purpose of these pilots I decided that using the transcripts from the focus groups would allow for a pragmatic, clearly defined and purposeful evaluation of two methods of data coding. Marshall and Rossman (1999, p.150) describe the analysis of data as "a messy, ambiguous, time-consuming, creative, and fascinating process" and my experience supports this.

Establishing Qualitative Codes

Content analysis is dependent on creating labels (codes) that can be applied to data in order to develop data into meaningful categories to be analysed and interpreted. Stemler (2001) discusses two approaches to the coding of data: emergent coding where codes are drawn from the text and *a priori* coding where codes are created beforehand and applied to the text. In piloting two data coding techniques I was able to examine the benefits and drawbacks of emergent and *a priori* coding and consider which was best suited to my study. Faherty (2010, p.59) reports that there are "no absolute hard-and-fast rules" to coding; therefore, it was important for me to be open to data coding methods and use these pilots as a means of discovery. By positioning myself within this study I chose to embrace the subjectivity of my research, therefore the choice of my research methods was likely to be influenced by my values (Greenbank, 2003) and the two analytical methods piloted were also likely to be value-influenced since my analysis of the

data was not a neutral pursuit but was infused with my epistemological and ontological assumptions (Mauthner & Doucet, 2003). I also recognise that any coding is likely to be a subjective and interpretive process and that the validity of my results becomes a question of hermeneutics as I interpret and translate my findings according to my position.

The first data coding system that was piloted involved "open coding": an emergent coding technique drawn from grounded theory methodology (Glaser & Strauss, 1967; Strauss & Corbin, 1998). The second pilot used "template coding": an *a priori* coding system drawn from template analysis (Crabtree & Miller, 1992; King, 1998). These two coding approaches were selected as prototypical extremes, where one attempts to identify the meaning within a text without any preconceptions and the other uses a purposefully developed framework as a means to draw out meaning. Open coding was identified as a method of generating a participant-generated 'theory' from the data and template coding was identified as a tool for framing data into a coherent construct through the application of an established 'language'.

I applied open coding to the transcript of focus group 1 (FG1) and template coding to the transcript of focus group 2 (FG2). In coding any data it is likely that some things will not fall neatly inside a code and some areas of both transcripts were left uncoded. This does not mean that such data are unimportant—only that they have not been classified, and for this reason, the coding of FG1 and FG2 involved three passes through the text in order to classify the data in as much detail as possible.

Open Coding (FG1)

Open coding is the first stage in the process of creating grounded theory (Strauss & Corbin, 1998). Grounded theory methodology (GTM) is an approach whereby the analysis of text allows the researcher to find the answers within; theory is developed from the data rather than imposed upon it. In GTM, the answers come from repeatedly coding, reviewing and refining the coding process. Strauss and Corbin suggest three stages to the process:
1.   Open coding
2.   Axial coding
3.   Selective coding
Open coding involves applying codes that are derived from the text (emergent codes). There is some debate regarding how this might be done: Glaser (1978, 1992) suggests that this should be done line by line; Corbin and Strauss (1990, p.12) encourage researchers to code "conceptually similar events/ actions/ interactions", and Stalp and Grant (2001) offer a linked framework that guides the first-time open-coder in how to recognise inductive concepts. Glaser (1978) also proposes constant comparisons of data and categories whilst Corbin and Strauss (1990, p.6) suggest that "the research *process* itself guides the researcher" [original emphasis]. But these angels-on-the-head-of-

pins debates are more suited to those firmly affixed within this field and for my purposes I decided that I would code concepts rather than lines of data. I decided against coding lines of data as this seemed rather arbitrary in that the amount of data on each line would be dependent on the size font used and the length of the line rather than the quality of the data.

The codes that are applied during open coding are not *a priori* codes and the researcher should not try to impose his/her own codes. This emergent approach causes some conflict with my epistemological perspective. This made it very difficult to code the data in a detached manner, as I was not merely coding an isolated transcript but, in reading through it, I could hear the "voices" of the participants and supposed myself able to understand what participants were "getting at" when they offered a response. It could be thought that such presumptions are unscientific but it is also important to consider the importance of structures of awareness (reflection) that allow us to interpret what we find (Marton, 1993).

Another issue when applying open coding is that the process implies that there is an actual truth out there awaiting discovery and that by coding and recoding I should be able to find this truth. We should also consider whether any coding system can really be "open" as we are all independently positioned subjects who are likely to start any activity from a certain viewpoint; whether we call this "individual perspective", "practitioner insight", "experience", "common sense", "institutional guidance" or even "theory".

The second and third phases in discovering grounded theory involve axial and selective coding. During axial coding "categories are related to their subcategories to form more precise and complete explanations" (Strauss & Corbin, 1998, p.24), and during selective coding "categories are organised around a central explanatory concept" (p.161) until an "analytic gestalt" (p.144) allows the theory to emerge. Here there is some tension within Strauss and Corbin's position as they move from the logical, systematical relation and intersection of data to the sudden insight of discovery. However, since my approach is reflexive, in that I am positioned within all aspects of the research process, I felt that making such an intuitive leap was possible as it was "grounded" by my insider knowledge. Whilst "the text of the written study is [...] considerably removed" from reality (Holliday, 2002, p.100), the text of my study is brought closer to the reality of its situation through the constant variable of reflexivity, in this way the analysis and the object under study are connected. Of course, embracing this ontologically subjective position means that I should also be aware of my inevitable internal assumptive bias. However, being aware of such bias does not, in this instance, mean that I try to reduce it—since this natural bias is part of who I am. It simply means that I am aware that the procedures and outcomes of my research are not the product of a heteronomous approach but the considered "best try" of a situated, autonomous individual. In this regard, I can only hope to make my methods as

transparent as possible; share my assumptions, and allow the reader to assess the credibility of my research.
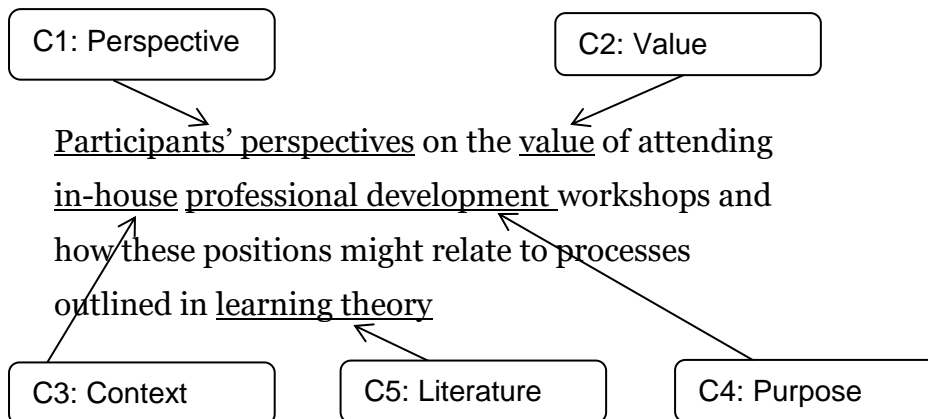
It was not my intention to complete all three stages of GTM during the pilot as the development of a 'theory' was not the focus of this activity. Rather my focus was on the development of a purposive data coding tool. I decided that, if I found the open coding process to be more effective that template coding, I would then return to the process and complete the second and third stages of GTM.

Template Coding (FG2)

The second analytical tool that I decided to pilot was "template coding" (King, 1994). The actual process is not dissimilar to that of GTM but there are two key details that set it apart. Firstly the codes used are defined by the researcher, which involves using *a priori* codes drawn from research, reading or theory. The second area where template coding differs is in its philosophical perspective. Template coding, in recognising the interpretive nature of the researcher, moves away from the positivist/realist paradigm of open coding, suggesting that some researchers are "sceptical of the existence of 'real' internal states which can be discovered through empirical research, and may therefore feel that template coding is more conducive to their position" (King, 1998, p.119). In this regard template coding seemed more in-line with my epistemology and ontology and offered me an analytical method that would allow the data to speak through me rather than at me. As my intention is to tell the story of my participants I felt that template coding could offer specific terms that would give the data a voice.

Template coding also meant that my reflexivity during the process would be addressed (Roberts, 1997) and that the analysis would embrace my position within the study and my personal "code of conduct" (Watts & Ebbutt, 1987, p.33). Chinn and Brewer (2001) suggest that when people analyse data they "construct a cognitive model of the data according to the perspective of the person who is reporting the data" (p.337)—they call this the models-of-data theory and it helps explain how my reflexivity supports my evaluation of the data. Since I am "involved" in the data collection, transcription, reporting and analysis this has an effect upon the "answer" that I then find. Dilley (2000) suggests that researchers should practise being self-reflexive and use their "one voice" (p.154) to analyse interviews — therefore, I felt that template coding would allow me to be honest in my approach and state that I was present throughout this study and I would be discussing things from my perspective. This does not mean that the method adopted is any less valid than others as the responsibility for rigour and validity in this instance lie with the continual self-verification of the researcher rather than an external actor (Morse, Barrett, Mayan, Olsen & Speirs, 2008). Here I claim validity of method—not by offering a faultless technique but through constant self-scrutiny, reflexivity and the analysis of any conclusions I draw.

Figure 1. Interpreting the research problem

```
┌─────────────────┐                    ┌─────────────────┐
│  C1: Perspective │                    │  C2: Value      │
└─────────────────┘                    └─────────────────┘
```

Participants' perspectives on the value of attending
in-house professional development workshops and
how these positions might relate to processes
outlined in learning theory

```
┌─────────────────┐     ┌─────────────────┐     ┌─────────────────┐
│  C3: Context    │     │  C5: Literature │     │  C4: Purpose    │
└─────────────────┘     └─────────────────┘     └─────────────────┘
```

The key problem I discovered with template analysis was in deciding which template to use. Since the codes developed through open coding were drawn from the participants (bottom-up) I was looking for a template that would offer top-down analysis. (At this stage I was hoping to see what was more effective: top-down or bottom-up analysis but, as will be evident later, this was a little naive.) How would I know which was the most relevant template and how would I know which codes would produce the results I was seeking? In an effort to answer this I returned to my research problem and to the key literature that supported the study and broke down the problem into five initial "codes". The irony in trying to develop an *a priori* template was that it first involved scrutiny of the text, through reading and re-reading of the research question. This process allowed the initial five codes to develop and Figure 1 shows the terms that initially emerged as being significant.

These initial five codes were selected as I felt they would allow for coding of specific concepts within the FG2 transcript. Code 1 (C1) was selected as it could be used to identify areas of text where it was clear that participants were talking about their own experiences of where they were offering their interpretation. C2 was selected as a means of highlighting what participants thought were the strengths and weaknesses of attending professional development workshops. I felt that C3 would be useful in showing if participants felt certain types of training were more beneficial than others. C4 was selected as this was central to the research question, so I thought it would be helpful in highlighting where participants discussed this concept. Finally, I selected C5 as I felt this code would be a means of differentiating between were participants spoke from their perspective and where they showed signs of being influenced by literature linked to the topic. Further this last code was central to the second part of my research.

As template coding allows for reflexivity, I also reflected upon these initial codes and considered what I was looking for. Upon reflection, I decided that C1 would not be a useful code as all the responses would

be individual, thus making this code redundant. As I was interested in finding out about participants' perceptions of the worth of professional development in relation to the space where the participants worked I decided to use a template that would code for value (C2) and for context (C3). I also decided that C4 would be a useful tool for highlighting where instances of professional development had been discussed in the focus group. My reflections led me to believe that one code for 'literature' (C5) was rather too simplistic and ignored the varieties of literature. Therefore I decided to break this code into its constituent parts. In an effort to find "literature" codes I identified 12 key themes through the scrutiny of the learning theory literature that had been used to underpin the course design for the various professional development workshops that had been offered to the participants.

Analysis of the workshops' course documentation showed that textbooks by a total of 18 authors had been cited. The textbooks were by Maslow, Honey and Mumford, Bloom, Tomlinson, Renzulli, Secada, Osborne, Ausubel, Bruner, Atkinson and Shiffrin, Festinger, Bandura, Schön, Kolb, Fleming and Mills, Spencer, Skinner, and Fitts and Posner. I applied a descriptive term to each of these texts and used these terms to develop the 'literature' codes. This way I developed codes that would not seek individual authors but would focus on references to the theories themselves. For example, I described Atkinson and Shiffrin's work using the term 'memory' and I described both Honey and Mumford's work and Fleming and Mill's work using the term 'learning style'. In total this left me with 12 codes that were drawn from the literature. I then added these "literature" codes to the three remaining codes drawn from the research question to make a total of 15 codes that made up the template (see Figure 2).

Figure 2. Developing qualitative codes

Initial codes	Revised Codes	Codes from literature	Final 15 codes (Template)
(C1) Perspective			
(C2) Value	Value	→	Value
(C3) Context	Context	→	Context
(C4) Purpose	Purpose	→	Purpose
(C5) Literature	→	Human needs	Human needs
		Learning styles	Learning styles
		Domains of learning	Domains of learning
		Personalisation	Personalisation
		Critical thinking	Critical thinking
		Teaching styles	Teaching styles
		Advance organisers	Advance organisers
		Memory	Memory
		Cognition	Cognition
		Social learning	Social learning
		Reflection	Reflection
		Learning cycle	Learning cycle

The Coding Process

Finlay (2002, p.224) discusses how reflexivity can offer a "confessional account of methodology" and, in offering an exploratory account of two data coding techniques, there is scope for some discussion on how my "position" influences the data coding process—since my interpretivist paradigm sees knowledge as a product of interaction, the data coding (before, during and after) supports an episteme of reflective engagement and personal preference. As well as considering intra-rater reliability, I also considered the possibility of having a second researcher code the data as a means of addressing subjectivity. However, there is very little guidance on how researchers can work together to form inter-coding groups (Campbell, Quincy, Osserman & Pederson, 2013) and there is no clear evidence that the inter-coding of qualitative data is ever truly dependable (Zhao, Liu & Deng, 2013). Therefore, upon reflection, I was not convinced that having a second (or third) researcher co-code the transcripts would improve the validity of the codes developed. In the end, I felt that my coding methods (in both FG1 and FG2) were validated by my constant attempts to establish causal inference that would best capture the data's "imageric meaning" (Glaser, 2002a) and by Kvale's (1994, pp. 166-167) suggestion that, "validity pertains to whether a method investigates what it purports to investigate"—in this regard I can claim my methods to be valid as I have set out clearly what I have done and, throughout, I have focused my methods on assessing whether either coding system has the ability to produce analyses that are in keeping with my reflexive perspective.

I decided to code both focus group transcripts manually using highlighter pens as opposed to using computer software such as NVivo. I decided to manually code the data for four reasons: firstly there were pragmatic issues to consider, such as the time it would take for me to become proficient in using coding software (Welsh, 2002). Secondly "leading programs were developed on the back of a specific approach—coding according to grounded theory" (Flick, 2006, p.353) which might mean that my analysis of template coding could be affected by external factors. My third reason for rejecting computer coding software was that Basit (2003) had found that using this approach might be more suited to those who wished to ultimately quantify their data; therefore this approach did not match with my objective of using these pilots as a stepping stone to further qualitative analysis. Finally, I rejected using computer software for reasons of personal preference—I felt more at ease with highlighter pens and paper spread out over my desk and floor. This was the way that I had always worked; using this method gave a tangible quality to my research; I began to know exactly which sheet held which comment, and I felt this approach gave me an overview (cognitively and literally) of the data and allowed for connections to be made.

Coding and Reflexivity

FG1 was selected for open coding based on the toss of a coin. In an effort to "focus on the meaning" (Charmaz, 2000, p.510) that the data contained I tried not to have any pre-determined terms, groups or codes (for this reason I also coded FG1 before FG2 so that it would not be influenced by an existing template). Mruck and Mey (2007) highlight the tension that can be felt between GTM and reflexivity and propose that not all grounded theory methodologies allow the researcher to embrace their situatedness. My reflections during the coding process illustrate this. I had hoped that the themes would emerge from the data in an almost organic manner but I soon became aware that my position within the research led me to define particular emergent codes. Qualitative research is textually mediated by the author and personally mediated by the reader and in this instance, as author and reader/researcher, I found it impossible and ultimately reductive not to apply my reflexivity to the coding process. That is: I was there during the focus group discussion; I transcribed the focus group data, and now I have coded and analysed it—it is quite clear that much of this is mediated by me and I would argue that this can only help develop an authentic account of my participants' perspectives. From working with the participants, it became clear that I was not an objective researcher reviewing data but a subjective teacher-researcher attempting to get to the heart of what his student-participants were trying to share. An example of this 'insider perspective' can be seen in Figure 3 where I coded participant D's comment as 'Aware of issues' despite its fragmented nature. In this example, I felt able to code D's response in this way as I was present during the focus group and therefore aware of the subtleties of expression and manner that helped give meaning to these remarks. I had also got to know participant D quite well and had a 'reading' of his meaning that was developed through this relationship.

Figure 3. Example of open coded transcript

D:	You need to be aware of how you are treating people. They want to learn and you want to teach... [B: true] ...but there are different ways of learning and teaching.	**Comment [A12]:** Self -awareness **Comment [A13]:** Teaching/learning styles
E:	...I think you want to/you need to/it's about identifying learning styles...	**Comment [A14]:** Learning style inventories
A:	...what their names? The learning style theorists...I always get them confused...	
D:	...the names are not important but what they say is more important...	**Comment [A15]:** Aware of issues
B:	...yeah/yeah...I agree but then you might start grouping individuals and that's like...well you're giving them names and that...	**Comment [A16]:** Labelling individuals

During the open coding phase, individual participant responses were coded for emergent key words. The initial coding of the transcript from FG1 was difficult and I felt a tension between "allowing the data to speak for itself" and my assumptions as the researcher. In the second phase (axial coding) repetitions of codes were searched for and drawn together. Codes were merged where substantial overlap was identified, and coded aspects were analysed for sub-themes. This process led to the development of ten codes that were drawn and refined from the transcript of FG1 and these ten codes were then revisited upon the transcript to highlight relevant areas of data. If I were to have followed GTM into the third stage I would then have started to look for a key theme/theory to emerge and arrange the other codes around this central concept. However this process was not required for the benefit of the pilot.

The template coding methodology involved three passes of the data, using the 15 pre-established codes. I decided on having three passes as Marshall and Rossman (1999) suggest that qualitative researchers can become "intimate" with the data through "reading, reading and reading" (p.153). Each pass happened independently so as to limit the bias from previous passes. I decided that in the three passes I would code for concepts rather than line-by-line as this was in keeping with the open coding method and would allow for side-by-side analysis with FG1 coding. In the end, the template coding pilot involved the application of 15 pre-defined codes to the FG2 transcript and these codes were able to highlight relevant areas of data for further scrutiny.

Analysis

As predicted I felt my epistemological framework was at odds with the open coding pilot. I constantly had to force myself not to pre-label responses and not to assume I knew anything beforehand. This was rather 'fake' as the transcription process meant that I had previously spent hours listening to the tape recording of FG1. A second 'issue' in attempting to use open coding was that it was never designed as a stand-alone process. The initial themes that emerged needed to be manipulated so that they could allow one part of the transcript to relate to another part of the transcript. This meant that I could not stop after the open coding stage but was obliged to draw the initial codes together through axial coding. While Strauss and Corbin (1998) suggest that through GTM the answer/theory will emerge from the data, I did not have an "analytical gestalt" because I held back and did not try to use selective coding. From this, it became clear to me that I was not fully applying a coding methodology, nor was I able to simple select part of a methodology – I was personalising the process to find an approach that worked for me. In total the application of the ten codes led to FG1 being coded 81 times—a number that is not significant in itself but does point to 81 specific areas where subsequent analysis would be likely to explicate participants' subjective understandings.

I had presumed that I would prefer the template coding approach as it would give my data a language with which to speak but I often felt constrained by its pre-definedness. Applying the template sometimes felt like using a tool that was not specifically designed for this purpose. I had assumed that the template would filter the data but there was a chance that it filtered out some key information. By choosing codes I was creating a system whereby some data was bound to be highlighted and some ignored—this seems at odds with my rather egalitarian (Labaree, 1998) research position. Having a template meant that I was specifically searching for things that might not be there and was faced with numerous decisions about whether a concept met the template definitions. If a statement did not form a match with any of my codes I was left with two choices: to leave the data uncoded or to shoehorn it in to a code that it did not quite fit! Although FG2 was coded 78 times, the template tended to highlight large chunks of data rather than specific key information. The template did not produce the thick description that I hoped it would, instead it highlighted broad areas of data and offered general areas that warranted further analysis.

Findings

These two pilot studies attempted to discover a reliable and useful way of coding the transcribed focus group data. I have discussed the factors that were at play throughout these pilots and their relationship with my epistemological and ontological assumptions. The purpose of these pilots was not to analyse the focus group data but to establish an appropriate data coding technique that could be used in analysing all the qualitative data developed during this research project. The open codes were able to highlight specific items/instances (possibly because they were developed from the focus group and therefore more likely to relate to its analysis) and the template codes tended to highlight large areas of data. In applying both these coding methods I was aware of personal tension. During the open coding phase, I found it difficult to be both reflexive and allow the data to speak for itself. Glaser (2002b) reports that researchers may subjectively interpret data in their attempts to reify it and that such bias can be addressed through constant correction but I found that each of my readings of the data was just as subjective as the last and I was never convinced that the interpretations ever became objective. I also found tension in the template coding process where template coding seemed to be dependent upon the selection of the "correct" template but I was never certain that the template I had developed was the right one.

The open coding system did not lead to a "Eureka" moment but the ten coded aspects that were developed through it are clearly participant-based and, since I hoped to seek their perspective, these codes seem likely to be useful tools. I feel sure that the codes drawn from open coding are unlikely to be completely unbiased, and I have already suggested how my reflexivity might enhance the research

project, but they are clear, relevant and useful. Not all the methods of GTM are suitable for my research but the first stage, open coding, has produced a device that is useful for highlighting aspects of professional development that are discussed by participants.

The application of the template to FG2 did not produce any really useful 'answers', and I suggest this is through a mismatch between template and data. The problem here seems to be that I was never comfortable in knowing what the 'best' template would be and that my final codes seemed too broad: which led to them coding broad areas of text.

Conclusion

I had originally set out to evaluate two data coding techniques in order to find which was more appropriate for my research. The open coding system helped develop a bottom-up device that reflects key concepts that were found in the participant data. The template coding system involved a more top-down classification of the data using concepts drawn from my consideration of what was important in regard to the research problem and the key literature. In the end there was no 'best' method and I adopted a combined approach as this seemed more pragmatic, more suited to the job, and because I was more comfortable with what I felt it offered. Using this combined approach also means that confirmatory bias is reduced as the bottom-up and top-down templates speak to, and counter, one another—leaving results that are neither wholly mine nor wholly objective. Figure 4 shows what I found to be 'my answer':

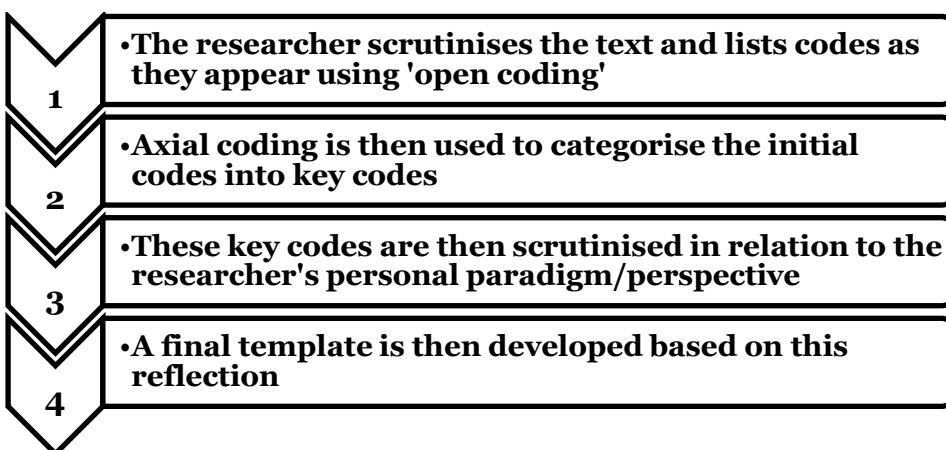*Figure 4.* Combined approach used to develop the data coding template

Figure 4 is not proposed as a generalisable answer but is shown to illustrate how I developed a tool, through reflexive inquiry, that seems most likely to tease out any meaning within the transcribed data developed through this research project. Further, this method of developing a qualitative data coding technique may not even be

applicable to other research projects that I am involved in. What is proposed as a generalisable answer is that to merely apply data coding techniques as they are outlined in the literature is flawed. In attempting to find meaning in qualitative data it is important that the right tool is used but such a tool might not already exist. Therefore, when analysing qualitative data, researchers should be methodologically thoughtful in their use of any data coding technique; that they do not assume that pre-established tools are aligned to their particular research paradigm, and that they consider combining and refining established data coding techniques to ensure that their coding system matches their paradigm/methodology. The time taken to pilot and develop such tools should not be seen as 'wasted time' but 'time well-spent', since the outcome is likely to be a data coding technique that is reflexively-aligned and better suited to finding the meaning the researcher seeks within the raw data.

Author notes. Erik Blair, Education Department, Royal College of Physicians, 11 St Andrews Place, Regent's Park, London, NW1 4LE, erik.blair@rcplondon.ac.uk

References

Basit, T. (2003). Manual or electronic? The role of coding in qualitative data analysis. *Educational Research, 45*, 143-154.

Bettis, P. & Gregson, J. (2001). The 'Why' of research: Paradigmatic and pragmatic considerations. In E. Falmer & J. Rojewski (Eds.) *Research pathways: writing professional papers, theses and dissertations in workforce education.* Lanham, MD: University Press of America.

Campbell, J., Quincy, C., Osserman, J. & Pedersen, O. (2013). Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement. *Sociological Methods & Research*, 42, 294-320.

Charmaz, K. (2000). Grounded theory: objectivist and constructivist methods. In N. Denzin & Y. Lincoln (Eds.) *Handbook of qualitative research, Second edition.* Thousand Oaks: Sage Publications.

Chinn, C. & Brewer, W. (2001). Models of data: A theory of how people evaluate data. *Cognition & Instruction*, *19*, 323-343.

Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education, Sixth edition.* London and New York: Taylor Francis and Francis Routledge.

Crabtree, B. & Miller, W. (1992). *Doing qualitative research.* Newbury Park: Sage Publications.

Dilley, P. (2000). Conducting successful interviews: Tips for intrepid research. *Theory into Practice*, *39*, 131-137.

Drake, P. (2010). Grasping at methodological understanding: A cautionary tale from insider research. *International Journal of Research & Method in* Education, *33*, 85-99.

Faherty, V. (2010). *Wordcraft: Applied Qualitative Data Analysis (QDA): Tools for Public and Voluntary Social Services.* Thousand Oaks, CA: Sage Publications.

Finlay, L. (2002). Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. *Qualitative Research*, *2*, 209-230.

Finlay, L. (2003). Through the looking glass: Intersubjectivity and hermeneutic reflection. In L. Finlay & B. Gough (Eds.) *Reflexivity: a practical guide for researchers in health and social sciences*. Oxford: Blackwell Science.

Flick, U. (2006). *An introduction to qualitative research*. London: Sage Publications.

Glaser, B. (1978). *Advances in the methodology of grounded theory: Theoretical sensitivity*. Mill Valley, CA: Sociology Press.

Glaser, B. (1992). *Basics of grounded theory analysis: emergence versus forcing*. Mill Valley, CA: Sociology Press.

Glaser, B. & Strauss, A. (1967). *The discovery of grounded theory*. Chicago: Aldine.

Glaser, B. (2002a). Conceptualization: On theory and theorizing using grounded theory. *International Journal of Qualitative Methods, 1*, 23-38.

Glaser, B. (2002b). Constructivist grounded theory? *Forum Qualitative Sozial Forschung/Forum: Qualitative Social Research*, 3.

Greenbank, P. (2003). The role of values in educational research: the case for reflexivity. *British Education Research Journal*, *29*, 791-801.

Heshusius, L. (1994). Freeing ourselves from objectivity: Managing subjectivity or turning toward a participatory mode of consciousness? *Educational Researcher*, *23*,15-22.

Holliday, A. (2002). *Doing and writing qualitative research*. London: Sage.

King, N. (1998). Template analysis. In C. Cassell & G. Symon (Eds.) *Qualitative methods and analysis in organizational research*. Thousand Oaks, CA: Sage Publications.

Kvale, S. (1994). *Focus groups: a practical guide for applied research*. London: Sage Publications.

Labaree, D. (1998). Educational researchers: living with a lesser form of knowledge. *Educational Researcher*, *27*, 4-12.

Lynch, M. (2000). Against reflexivity as an academic virtue and source of privileged knowledge. *Theory, Culture & Soci*ety*, 17*: 26–54.

Marshall, C. & Rossman, G. (1999). *Designing qualitative research, Third edition*. Thousand Oaks, CA: Sage Publications.

Marton, F. (1993). Our experiences of the physical world. *Cognition and Instruction, 10*, 227-237.

Mauthner, N., & Doucet, A. (2003). Reflexive accounts and accounts of reflexivity in qualitative data analysis. *Sociology*, *37*, 413-431.

Morgan, D. & Krueger, R. (1997). *The focus group kit*. London: Sage Publications.

Morse, J.M., Barrett, M., Mayan, M., Olsen, K. & Spiers, J. (2008). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods*, 1, 13-22.

Mruck, K. & Mey, G. (2007). Grounded theory and reflexivity. In: A. Bryant & K. Charmaz (Eds) *The SAGE Handbook of Grounded Theory*. London: Sage Publications.

Pallas, A. (2001). Preparing education doctoral students for epistemological diversity. *Educational Researcher*, *30, 6*-11.

Peshkin, A. (1988). In search of subjectivity, One's own. *Educational Researcher*, *17,* 17-21.

Roberts, C. (1997). Transcribing talk: Issues of representation. *TESOL Quarterly*, *31*, 167-172.

Robson, C. (2002). *Real world research, Second edition.* Oxford: Blackwell Publishing).

Stalp, M. & Grant, L. (2001). Teaching qualitative coding in undergraduate field method classes: an exercise based on personal ads. *Teaching Sociology*, *29*, 209-218.

Stemler, S. (2001). An overview of content analysis, *Practical Assessment, Research & Evaluation*, *7.* 137-146.

Strauss, A. & Corbin, J. (1998). *Basics of qualitative research, Second edition.* Thousand Oakes, CA: Sage Publications.

Watts, M. & Ebbutt, D. (1987). More than the sum of the parts: Research methods in group interviewing. *British Educational Research Journal*, *13*, 25-34.

Welsh, E. (2002). Dealing with data: Using NVivo in the qualitative data analysis process. *Forum Qualitative Sozial Forschung/Forum: Qualitative Social Research, 3.*

Zhao, X., Hao, J. & Deng, K. (2013). Assumptions behind intercoder reliability indices. In C. Salmon (Ed.) *Communication Yearbook 36* (New York: Routledge).