

Using Criteria of Significance to Make Sense of Data: Implications for Qualitative Research

Daryl Morrison
University of New Brunswick

Mark Hirschhorn
University of New Brunswick

For many qualitative researchers, the task of dealing with huge amounts of data can be overwhelming. In many qualitative research methodologies, procedures for making sense of large amounts of data are often intentionally unclear and open to interpretation due to the wide range of variability of data and research context. This can be problematic for novice and experienced researchers alike as they consider what parts of their data to feature, exemplify and draw conclusions from. This article puts forth a construct that makes explicit the logics of two researchers using what they label as “criteria of significance” to make sense of their qualitative data. The Criteria of Significance (CoS) serves as a defensible set of criteria by which data is given increased or decreased value regarding its use in the final analysis and conclusions drawn from a study. This paper examines two qualitative studies (Hirschhorn, 2008; Morrison, 2018) and explores how CoS was used to differentiate the data used in their findings.

Keywords: criteria of significance, qualitative research, grounded theory methodology, verisimilitude

INTRODUCTION

For many qualitative researchers, the task of dealing with huge amounts of data can be overwhelming. Regardless of the topic being investigated, the research process can become stalled when researchers consider how to distil the data into the components that best answer the research question while reflecting the data generated from various methods, particularly those that are characteristic of conversations between participants and the researcher like interviews and focus groups. In many qualitative research methodologies, procedures for making sense of large amounts of data are often intentionally unclear and open to interpretation due to the wide range of variability of data and research context. This can be problematic for novice and experienced researchers alike as they consider what parts of their data to feature, exemplify and draw conclusions from. The absence of a procedural orthodoxy can sometimes place the researcher in a vulnerable situation when having to defend or substantiate their knowledge claim. Attending to how decisions are made about the data or what Piantanida et al. (2004) refer to as “logic-of-justification” (p. 335), this article puts forth a construct that makes explicit the logics of two researchers using what they label as *Criteria of Significance* to make sense of their qualitative data. The Criteria of Significance (CoS) serves as a defensible set of criteria by which coded data is given increased or decreased value regarding its contributions to the conclusions resulting from data analysis.

This paper begins by introducing its purpose and continues with some literature about the challenges of qualitative data analysis. Next, each author offers a reflective recount of their struggles when they faced voluminous amounts of data underscoring the need to derive a defensible means to make sense of their data, and ultimately choose which aspects of their findings to centralize in their conclusions. Finally, this paper concludes with the implications of the authors' use of CoS for other researchers seeking to differentiate and apply significance to their own data.

PURPOSE

As researchers who have successfully used grounded theory as a methodology, we acknowledge the dual challenge researchers face of making sense of the methodology itself while dealing with the large amounts of data generated when conducting qualitative research. Although there exists some methodological resources (Charmaz, 2003; Cranton & Merriam, 2015; Creswell, 2014; Crotty, 1998; etc.) that discuss the how's of collecting data using particular methodologies, they stop short of offering the reader specific ways they might make sense of the enormous amounts of data that results.

Against this background, we sought with this paper to introduce a construct known as criteria of significance (CoS) and how these criteria were used in two studies (Hirschhorn, 2008; Morrison, 2018) to sort, prioritize, and attribute significance to the data. In short, we are seeking to help researchers navigate their way through the research process, especially if they feel overwhelmed and uncertain about how to deal with huge amounts of data, by showing how we were successful in differentiating our data in an open and defensible manner using the CoS.

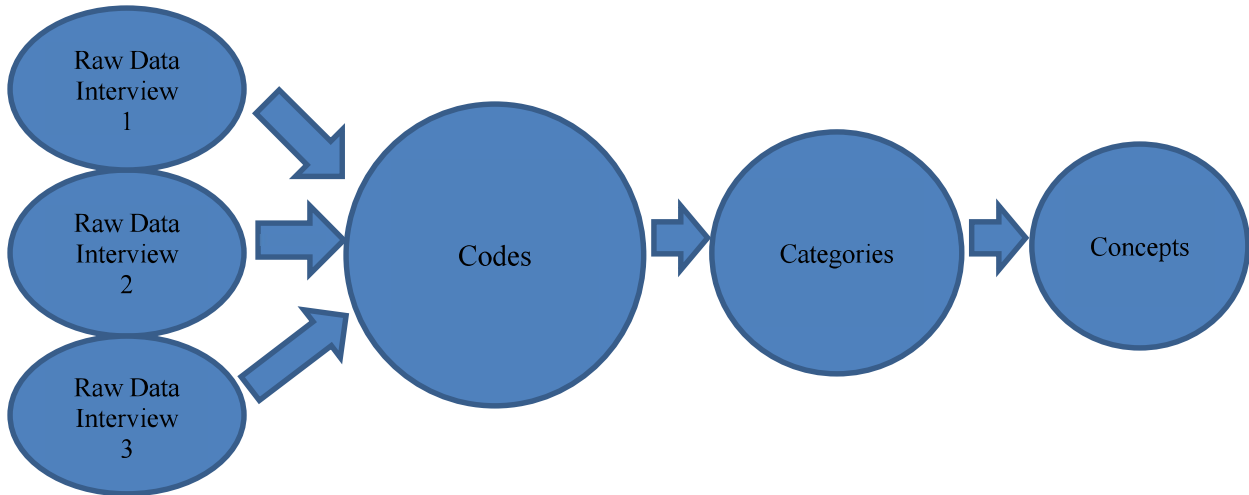
Secondly, this paper will show that by using CoS, the verisimilitude of qualitative studies is strengthened by making the researcher's decisions explicit in relation to how the data was used to find categories, themes, and patterns that emerged from the data or, in the case of grounded theory methodology, in the generation of a theory. This paper puts forth the notion that although all qualitative data is important and significant to one extent or another, researchers must make hard choices regarding what data to feature as the core foundations for the findings from their studies. As Cranton and Merriam (2015) note: "The task of the researcher is to cull from the many bits of data those that are representative, and to shape accurate and meaningful conclusions" (p. 110). We suggest that using criteria of significance is a tool researchers can use for this purpose.

TYPICAL CODING PROCESSES IN QUALITATIVE DATA ANALYSIS AND THE NEED FOR SOMETHING MORE

For many researchers the how's of analysing research are derived from graduate course work, mentors/supervisors, reviewing the methodological descriptions of qualitative research in the literature and by referencing methodology texts such as *Research Design* (Creswell, 2014), *A Guide to Research for Educators and Trainers of Adults* (Cranton & Merriam, 2015) or *Qualitative Research in Education* (Lichtman, 2013). From this array of sources, the process of coding qualitative data is learned. Typically, researchers will use a deductive or an inductive recursive coding process to narrow their data into categories that originates from the data and represents the data.

Figure 1, an adaptation from Lichtman (2013, p. 252), captures the foundation procedure for making sense of your data using reductive coding processes.

FIGURE 1
THE REDUCTIVE CODING PROCESS



Other sources describe the coding process using a range of different descriptors such as a progression of open, axial and selective coding (Strauss & Corbin, 1990), yet the process remains largely the same – large amounts of data being filtered through increasingly encompassing categories to find the concepts that best describe the collected data being considered with regards to the research question. As researchers we have utilized this approach with success, however, there remains one nagging dilemma – how do you weight the categories you derive in the coding process to best align with the emphasis or resonance of the research participants? What criteria are you using to help the reader understand why you have featured the concepts you do at the end of the reductive data analysis process?

Other researchers have expressed concern over the contextual complexity of qualitative data analysis given the necessary lack of canon regarding the derivation of meaning from research data. Thorne (2000), for example, notes: “Unquestionably, data analysis is the most complex and mysterious of all the phases of a qualitative project, and the one that receives the least thoughtful discussion in the literature” (p. 68). She contends that even for consumers of qualitative research, “the language of analysis can be confusing” (p. 68) and knowing how the researcher derived meanings from the data is not always made explicit. She claims:

In describing their processes, some authors use language that accentuates this sense of mystery and magic. For example, they may claim that their conceptual categories “emerged” from the data-almost as if they left the raw data out overnight and awoke to find that the data analysis fairies had organized the data into a coherent new structure that explained everything! (p. 68)

Similarly, Colley and Diment (2001), in a discussion of alternative data analysis methods, echo the challenge of making sense of their data. They contend:

Both of us have encountered difficulties in our research, particularly in making sense of our qualitative data using methods of analysis that dominate educational research and are advocated in the majority of research training courses and textbooks. (p. 2)

These authors explored “the manner which data is analysed” (p. 5) and “believe this is an important research issue for practitioners, because analysis is the means by which we make sense of data” (p. 6). They concluded that “when confronted with a large amount of discursive data, it can be extremely difficult to

make sense of that whole, and arrive at a coherent interpretation” (p. 6). Other researchers like Piantanida et al. (2004) acknowledge the confusion around the language used in data analysis in their quest to understand grounded theory methodology (GTM). Faced with interpreting multiple approaches to GTM, as a result of the original authors’ methodological departure from one another (sometimes referred to as *Glaserian* or *Straussian*), these authors note that “it took considerable time and many discussions to gain some perspectives on the threads of discourse embedded in the literature” (p. 328). This included not only understanding the varied philosophical underpinnings of GTM but related methodological constructs (e.g. constant comparative analysis, theoretical sampling, theoretical sensitivity, saturation) as well. Upon realizing that grounded theory, like other qualitative methodologies, does not rely on a procedural orthodoxy, the authors quickly concluded the “importance of methodological flexibility” (p. 329) and how approaches to research are continually evolving across research contexts. This is an important realization for qualitative researchers. Like Piantanida et al. (2004), moving from a narrow focus on the procedures of grounded theory (or any other qualitative methodology) to a broader focus on making explicit the theorizing process or what they refer to as “*logic-of-justification*” (Piantanida et al., 2004, p. 330), liberates the researcher to organize, select, and prioritize their data. This means that “by approaching method as logic-of-justification, the grounded theory researcher makes explicit the connections among research paradigms, strategies and techniques” (p. 331). Therefore, it is the responsibility of the researcher to make explicit their decisions about how they sorted, prioritized, and placed significance on the data in the derivation of their findings. Cranton and Merriam (2015) succinctly state: “The hard work of analysis, of figuring out how all data might be linked and what it all means, still resides with the researcher” (p. 141). Therefore, both novice and experienced researchers alike need to avoid presenting their findings in a manner that loses the reader in methodological language and terms and be clear about their decisions in relation to making sense of the data. “Researchers are therefore encouraged to articulate their findings in such a manner that the logical processes by which they were developed are accessible to the critical reader, the relation between the actual data and the conclusions about data is explicit, and the claims made in relation to the data set are rendered credible and believable” (Thorne, 2000, p. 70).

As the authors of this paper, we will each share a brief narrative of the specific dilemma we each encountered as we sought to make sense of the voluminous amount of data we had each collected. In the end, we each decided to use a variation of what Hirschkorn (2008) labelled as the Criteria of Significance (CoS) to offer the reader a means to determine why we attributed meaning to our data in the manner that we did. We will follow these narratives with a section delving into the meanings and applications of CoS, so that the reader has a sense of how they may use CoS in their own work.

BEGINNING TEACHERS’ STUDENT TEACHER RELATIONSHIP EXPERIENCES – DR. MARK HIRSCHKORN

I would like to begin my story with an excerpt from a journal I kept as I was doing my Ph.D. that focused on the evolution of student teacher relationships (STR) for beginning teachers and how these STR’s affected the development of the beginning teacher.

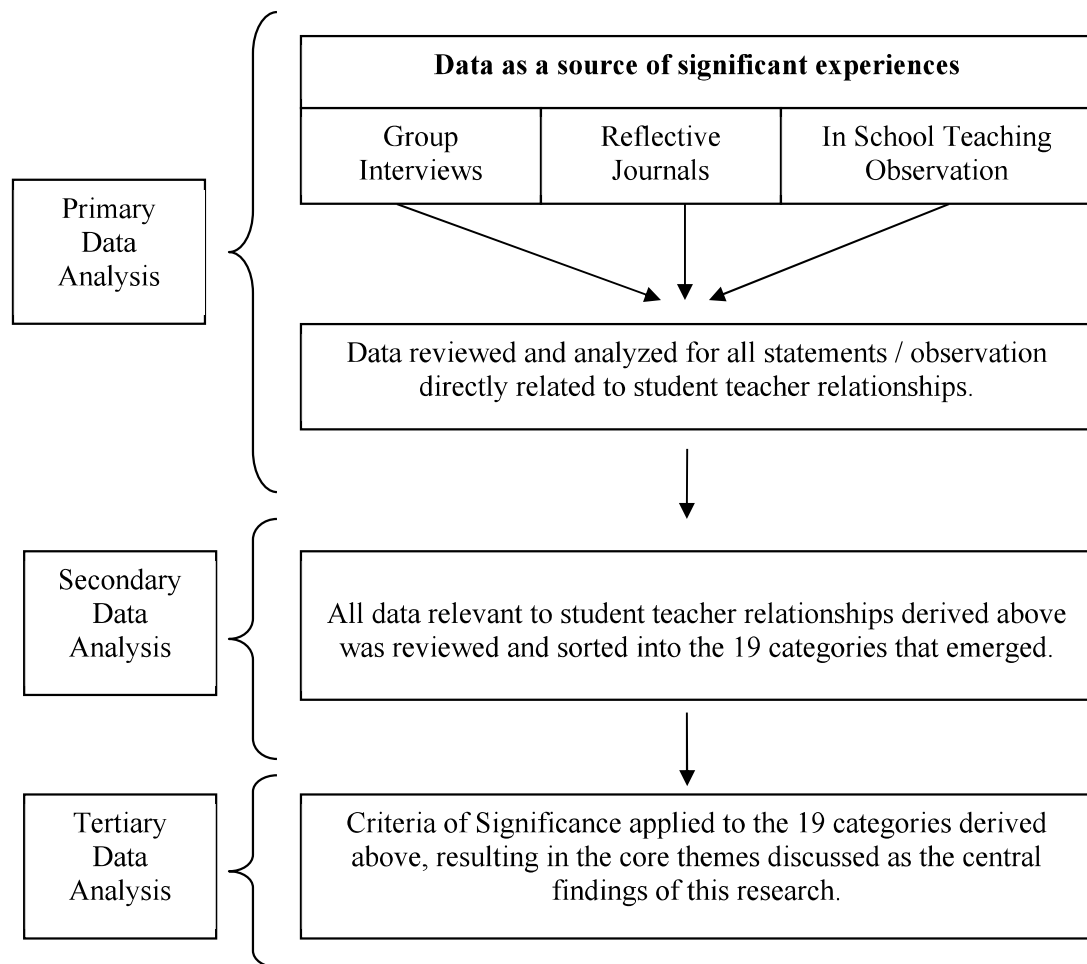
Sigh... , will you look at this mountain of data on my desk. Hours and hours of transcripts from our group meetings, three semesters of journals from my 13 participants and volumes of my notes from the classroom visits I made over the last 2 years. This is going to take a while. I need to make sure I retain a clear idea of the means I am using to sort through all of this, but the time has come to figure out what is relevant and significant in all of this.

I feel almost guilty as I begin – what arrogance for me to visit and revisit the worlds of these beginning teachers and to suppose I have an insight that others would not have. I know that if this stack was handed to any other academic they would likely focus on a different set of quotes or instances, but I am no rookie. I have done my homework and have read a LOT on student teacher relationships. As I now begin to review the data it is almost

as though some of the moments from the last 2 years scream “this is significant; say something about me!” I do not have to have the same eyes as others, but that is ok, I just need to make sure that the eyes I am using remain apparent to the people that read this; That is, they can figure out why I have come to the conclusions that I have.

For my Ph.D. research, I used group interviews, classroom visits and participant journals to collect my qualitative data, analyzing them through a combination of narrative and grounded theory methodologies. I worked with 13 beginning teachers through the last year of their B.Ed. programs and all the way through the first year of in-service teaching. However, since the participants were the direct source for much of the data (journal entries and interviews), the data relevant to the student teacher relationship information I was after inevitably lay embedded within a jungle of emotionally charged commentary. The valuable data was mixed with thoughts on other aspects of their experiences other than the student teacher relationship, and even value judgments evaluating the ‘good’ and ‘bad’ moments. It became necessary to apply an analysis framework that allowed me to remain focused on the student teacher relationship while presenting what I had learned in a manner that suggested the significance and relevance of the data. Not everything that emerged in this research was equally represented, had equal value to the participants (or seemingly, the academy) and thus some representation of its relative significance would need to come through in the analysis. My conceptualization of this analysis framework is summarized in Figure 2 below.

**FIGURE 2
DIAGRAM OF DATA ANALYSIS MODEL**



INFLUENCE OF EACH DATA COLLECTION METHOD ON THE TYPE OF DATA THAT RESULTED

Even though I had a sense of grounded theory as I was collecting the data for my research, it became apparent during the data analysis that the different methods I had chosen were resulting in an unanticipated ‘skewing’ of the data each method collected. For example, the participants’ reflective journals offered some insight on the participants’ intentions for and perceptions of a situation – which at times contradicted observations of their actions during a classroom visit. Whether these differences in the data each method produced were due to when, who and how each method was completed – the different methods did favour slightly different aspects of the student teacher relationship experience. This was not to the exclusion of common findings between the different instruments, but patterns became evident. This became important in my findings, as I was able to use these patterns as one of the criteria of significance I derived to attribute meaning to the patterns I was observing. Commonalities in the data that emerged from different data collection methods was one means I used to attribute more significance to one data category over another.

Group Meetings

Group meetings were particularly strong at evoking resonant proclamations of support or denial for the issue being discussed (and often periods of revelations as well). However, this had a tendency to mask the originator of the insight as well as funnelling the meeting toward a particular topic theme (for better or for worse). Thus, the data that tended to come out of the group meetings were by their nature already distillations of social negotiation. In a sense, the data was filtered by the reactions and body language of the participants. Additionally, the particular make-up of the group (which participants attended at any particular meeting) had a tendency to nudge the topics toward the concerns of the participants willing to hold their own in the social milieu the groups became. Interestingly these patterns of social engagement became evident to the group as well, and eventually there began a form of self-policing that had participants prompting other, quieter, members for their insights on the topic at hand. It is worth noting, the group meetings were the aspect of the research that the participants confided to me was the component they most valued. They had begun to need the group meetings as a means to unpack their anxiety and to sound out what they were thinking and what worried them. They found support and comfort in being able to talk about their experiences with trusted peers who were going through a similar period of development.

Classroom Visits

The classroom visits were characterized by action and outcome. As an observer in the back of the room, I was most often not privy to the reasons for what the participant was doing, or the personal reactions these interactions may have created beyond some minor interpretation of the body language being represented. However, the data collected in the classroom was rooted in the pragmatics and messiness of the classroom - actual student teacher interactions, and the results of these interactions. Classroom visits were the only data collection venue that afforded me, the researcher, some awareness of the relationship in action – I got to see the teacher *and* the student half of the student teacher relationship. Additionally, this data source was particularly good at providing the context for the reflections I was receiving from the other sources. This was also the data collection method in which my presence as the researcher was the most evident. Beyond the reaction of the students (and participant) to my presence in their class, the data collected in this venue was wholly a product of what *I* was seeing and interpreting. Right from its origin, these data were a product of my own experiences as a classroom teacher and what meaning I learned to attribute to what I was observing.

Reflective Journals

The reflective journals were the source for the majority of the insights offered by the participants in which they intentionally focused on the student teacher relationship. As a result, they make up the largest part of the data and are represented most heavily in the insights drawn from the data. In general, they represented a type of data that was much more personal and reflected the motivations and emotional

reactions that were not as evident in the other two forms of data collection. This data collection method was particularly valuable since it was the only venue for the participants to speak directly as they intended without the filter of the group or my perceptions to mask their intent (even though some stated they felt some pressure to edit their reflections, knowing that the University Facilitator and potentially their mentor might read their journals). There were additional strengths to using journals as a data source. The journals reflected the immediacy with which they were written – they were relatively immediate reactions to the moments that created the insight (they didn't have to wait for the next meeting or classroom visit to 'get it out'). The insights present within the journals tended to reflect what each participant actually wanted to share and was not subject to the misrepresentation that can occur with the in-the-moment immediacy and fallibility of conversation (as the cliché goes, "I always think of what I wanted to say after the conversation was over"). It is worth noting that this also was the data collection device least liked by the participants. In an already insanely busy day, they were asked to add this to their list of things to do – most described it as "*having* to do my journal", as opposed to some derivative of *wanting* to. The best evidence for this was that when the obligation to keep a reflective journal as a part of the universities practicum program ended, most of the participants dramatically reduced their journal writing, and all ceased their writing when the research period finished. Goodlad (1990) observed a similar pattern of declining numbers of reflections from his participants and suggested that to expect prospective teachers to become both reflective students and reflective practitioners in the midst of surviving their first few years of teaching may be unreasonable.

DATA ANALYSIS

As detailed at the outset of this article, as a qualitative researcher employing grounded theory and narrative methodologies, my first steps in my data analysis was largely a process of coding and categorizing my data. Two years worth of group meetings, classroom visits and journals resulted in a lot of data, but I followed the analysis protocol typical of qualitative coding. Primary analysis amounted to reviewing all of my data as it was collected, sorting the relevant data from the irrelevant, and attributing a few words to each data moment. Secondary analysis continued the typical coding protocol of grouping similar data utterances under headings that originated from the data while capturing a meaning that contributed to the whole of the study. The result was 19 categories that I concluded encompassed my data, and despite subsequent returns to the primary codes, the data did not result in any further categories. Examples of these categories include "Length of Practicum", "The Mentor", and "Influence of Personal Qualities". Thus, as is part of the grounded theory process, I concluded I had reached saturation.

DERIVING THE CRITERIA OF SIGNIFICANCE

The fundamental assumption of my Ph.D. study was that the experiences associated with student teacher relationships are significant to beginning and in-service teachers. Over the course of a two-year data collection process, I had listened to, collected and recorded many stories and experiences from the participants. After applying the primary and secondary analytic coding processes, 19 categories resulted which described and grouped the data. However, the presentation and discussion of the 19 categories with no attempt to sort or attribute significance to them would not do the data, the participants, or the study justice. Qualitative coding canon would suggest at this point a process of creating broader codes that subsumes and describes multiple categories under common headings is the next step in the GTM analysis chain (Lichtman, 2013). However, as a beginning researcher it was not clear to me how to start combining, and elevating in significance, certain categories over others. Even though all the categories originated with the data, they were not equal in how they represented the participants, or the relative importance attributed to them by the participants. Some of the key categories contained the voice and insights of all the participants and were evident in almost all aspects of their experiences, whereas other categories were more situation or participant specific. Initially, I intended to apply criteria to act as justification for removal of some of the least prominent categories. Upon reflection of my sample size, however, I decided that if a category emerged with only a single person from my 13 participants, that still potentially represented a

significant number of beginning teachers who may be sharing a similar experience and thus, warranted a brief discussion as well. It became necessary to establish criteria that I could use to demonstrate how and on what basis I had narrowed my 19 categories into the core findings of the study.

Group Resonance

As the first criterion of significance, group resonance is based on the collective reaction of the participants to an experience. This emerged in one of two forms. First, the participant discusses and reflects upon a student teacher relationship (STR) issue in their journal or in the discussions we had following a classroom visit, that is also present in the reflections of other participants - independently. Second, a STR insight was volunteered by a participant at one of our group meetings that subsequently provoked a resonant reaction within the other members of the group - as evidenced by the verbal and non-verbal reactions of the group to the story. For example, the majority of the participants discussed in their journals how the atmosphere within the school has a large impact on their relationship with their students, and thus qualifies as an example of group resonance. Similarly, the group as a whole reacted quite strongly to a discussion during one of our group meetings on whether they are treated as teachers in their schools while student teaching – a topic directly tied to teacher legitimacy for them. Even though this did not occur to many of the participants before the meeting, that is, there was no mention of this in their journals or classroom visits, it resonated with them in such a way as to provoke a strong reaction from them, and thus qualifies as a form of group resonance as well. In Table 1 below, group resonance is represented by the ‘Percentage of Participants’ column – as it indicates the number of participants (divided by the total) that had at least two different instances in which they discussed something attributable to that category. I chose a minimum of two to remove the incidentals – if a participant returns to a category more than once it suggests more than just a passing thought.

Device Resonance

The second criterion of significance, device resonance, is when different data collection devices, despite favoring forms of data over others (actions vs. motivations for example), were still the source for data that fell under common categories. A good example of this, is how classroom observations, group meetings and the reflective journals all brought out that how the participants ‘see’ or approach student teacher relationships has a large influence on the relationships that eventually emerged. Yet, the category ‘Peer/Socialization Pressure’ did not appear at all in the data collected from classroom visits, even though it did emerge briefly during our group meetings and in the journals. Thus, using this criterion, ‘how the students see student teacher relationships’ is more significant than ‘peer/socialization pressure’. In Table 1, device resonance is reflected in the ‘Data Breakdown’ column, as it indicates the number of instances in which the participants offered insights through each of the different data collection venues.

Individual Resonance

The third criterion of significance, individual resonance, uses the relative salience of the participants’ described relationship experiences to the participant. If a participant repeatedly focuses on a single aspect of their experiences, that aspect would seemingly be more important than other aspects to that person. For example, if Bill repeatedly describes in his journal and group interview conversations the issue of not trusting his students, and this lack of trust possibly stems from a betrayal he experienced early in his first practicum, that betrayal would be data important to prioritize since it had such a profound effect on his relationships with students. Thus, if a participant presents an experience as a ‘critical incident’ in that teacher’s experience with student teacher relationships, its inclusion in the study received a higher priority. In Table 1, the ‘Feature People’ column represents individual resonance as it demonstrates which individuals repeatedly returned to that category of insights. Given the total number of insights originating with each participant (and the group), five or more separate instances in which they discuss an experience attributable to that category seemed a reasonable cut-off to demonstrate a focus on that category.

Finally, the three criteria of significance are themselves not of equal importance. Group resonance was given the highest priority since it represents the relative value of the category to all of the participants

regardless of data collection method or place in their teacher progression. Device resonance was weighted second in importance due to the unique triangulating value that having a category emerge from different data collection methods represents. It does not necessarily represent all of the participants, but significance is suggested when an issue emerges during a classroom observation and in a personal reflective journal or in the group meeting. Individual resonance is the least significant criterion since it has the potential to be the source for the most anomalous or idiosyncratic data – despite being a wonderful source of in-depth discussion of the insight and why it may be relevant to a beginning teacher.

In summary, the Criterion of Significance applied to the 19 categories are:

- **Group Resonance (GR)** - Which experiences were shared and described by many of the participants, or provoked the most resonant reaction in the group?
- **Device Resonance (DR)** - Which experiences emerged in all the data collection methods?
- **Individual Resonance (IR)** - Which experiences did a single participant return to repeatedly?

As a data discussion tool, applying the criteria to the categories is helpful to illustrate which categories are the most salient to the participants and thus, also the most salient to this study. This does not suggest that the rank of the categories is rigid – even the label attached to each category was coined by me from what I was seeing in the data, but it is indicative of the amount of focus the participants directed toward each category. A summary of the result of applying these criteria is in Table 1.

TABLE 1
HIERARCHY OF DATA CATEGORIES AFTER APPLYING THE CRITERIA OF SIGNIFICANCE

| Rank | Category Description | Percentage of Participants (GR) | Data Breakdown (DR) C=Class Visit M=Meeting J=Journal | Feature People (IR) |
|-------------|---|--|--|---|
| 1 | How do participants ‘see’ Student Teacher Relationships (STR) | 85 | 6C/7M/52J | Eric 12; Jenna 7; Ray 6; Jason 5; Mike 6; Jon 5; Carrie 5 |
| 2 | Influence of personal qualities | 85 | 11C/1M/42J | Carrie 11; Jenna 11; Jason 6 |
| 3 | Nature of the community / school | 77 | 13C/2M/31J | Imran 5; Jenna 5; Lincoln 9 |
| 4 | Teacher legitimacy | 69 | 13C/7M/30J | Carrie 6; Lincoln 7; Jenna 5 |
| 5 | Efficacy | 69 | 0C/6M/35J | Jenna 6; Sincere 5; Lincoln 5 |
| 6 | Content insecurity and the STR | 62 | 11C/1M/24J | Eric 8; Lincoln 6; Jenna 5 |
| 7 | The mentor | 62 | 10C/2M/25J | Eric 5; Sincere 6 |
| 8 | Relationships are good; managing is separate | 54 | 7C/1M/31J | Eric 9; Carrie 5; Jenna 7 |
| 9 | Genuine vs. contrived relationships | 54 | 6C/2M/22J | Camille 6; Jason 6 |
| 10 | Student push limits; know the line | 54 | 8C/1M/18J | Jason 5 |
| 11 | Length of practicum | 54 | 5C/3M/19J | Jenna 5; Eric 5 |

| | | | | |
|----|---|----|-----------|-----------|
| 12 | Technical approach vs. the STR | 54 | 6C/1M/19J | Lincoln 7 |
| 13 | Exhaustion | 46 | 5C/2M/11J | None |
| 14 | Peer / Socialization pressure | 38 | 0C/3M/19J | Lincoln 6 |
| 15 | Pragmatics | 23 | 0C/4M/18J | None |
| 16 | Maintaining the teacher image | 38 | 0C/0M/14J | None |
| 17 | Students crossing the 'unprofessional line' | 15 | 3C/2M/5J | None |
| 18 | Where does the STR happen? | 23 | 0C/0M/11J | None |
| 19 | Universities role / influence | 8 | 0C/0M/8J | None |

A critique of Grounded Theory studies is that the backgrounds and bias of the researcher doing the study heavily influence the coding and categorization process. This is acceptable. As Lictman (2013) notes: "All information is filtered through the researcher's eyes and ears and is influenced by his or her experience, knowledge, skill, and background" (p. 21). Yet, the results of the reductive coding process are often presented as though there was no alternative to the codes and categories that result. This was off-putting to me as a relatively new researcher at the time, because I was very aware of how much of me was present in what I was presenting as important within the data. Thus, the three criteria of significance I derived were ultimately one of the mechanisms I applied to demonstrate to the reader the basis for why I privileged some of my data over other data. By using what I believe to be a defensible set of criteria, I believed I could demonstrate transparent reflexivity and verisimilitude. I could make it apparent to the reader where my primary findings originated and why they were the findings I discuss as the take-away's of significance in my research.

HOW TEACHERS AND ADMINISTRATORS USE SCHOOL IMPROVEMENT PLANS TO MAKE CHANGE IN THEIR SCHOOLS – DR. DARYL MORRISON

My Initial Fear

What am I going to do with all this data? This was the central question I faced as I began to analyse my data. Irritability, frustration, nervous and anxious feelings over not knowing exactly how to do GTM or even where to begin; overwhelmed by the amount of data that needed analysis; terrified of how I was going to get through all of the transcribed interviews-pages and pages of it; losing control of the process I envisioned; being lost in the data; feeling defeated and losing self-confidence. These were some of the feelings I experienced as the data I collected waited for my analysis.

Waking up countless times in the night trying to figure out what to do; listening to interviews repeatedly while driving; losing focus when attempting to analyse the data; increased heart rate and probably high blood pressure coupled with feelings of anxiety, helplessness and doubting myself. These were the physiological realities of my experience prior to analysing my data.

My Mountain of Data

These emotions and sense of self were my lived experience as I grappled with GTM and faced piles of data waiting for my analysis. My data comprised of transcripts from 27 interviews (9 participants) that were usually 30 minutes to one hour in length, elicited responses from teachers and administrators, and pages of extant texts from each of the three schools examined. These were in the form of school improvement plans, school review reports, tables of achievement data and pages of teacher perception survey results. To get a sense of what the transcribed interview data looked like, I went back and looked at my first interview. It was approximately 30 minutes in length. After being transcribed, there were 192 multiple line exchanges between the participant and myself as the researcher. A multiple line exchange meant one or several lines of transcribed text that captured either what the researcher said (co-constructivist role) or what the participant shared. Each response was numbered with the corresponding transcribed text. To get a sense of

the voluminous amount of transcribed text from the first four participants (12 interviews), there were 1,876 responses. To put this in perspective, as part of the audit trail of my data analysis from my dissertation, I included a sample interview in the appendix. This showed 120 single multiple line responses. It was nine pages in length. Therefore, it is easy to conclude that 27 interviews generated a significant amount of data and by far the most data of all the collection methods used. In addition, the forms of extant texts collected also resulted in a plethora of print material to organize, review, and analyse. To a lesser extent, the elicited responses did not generate as much data as I had hoped for. With teachers and administrators being busy people, it is likely they had little time to respond to my questions.

Choosing Grounded Theory as My Methodology

Prior to my study beginning, I was excited about selecting a methodology that I felt would best answer my research question. I was drawn to GTM for a few reasons. First, I was attracted to the notion of developing a substantive theory as a product of my methodological choice. In the absence of theories about how teachers and administrators utilize school improvement plans to make change in their schools, I was eager to create one from the data analysed. Secondly, I was intrigued by what I was reading about GTM in relation to how theories could be generated from categories of data. What I did not realize though, was how much data my collection methods would generate. It was more than I anticipated, and I became overwhelmed with how I was going to analyse it.

Beginning Data Analysis

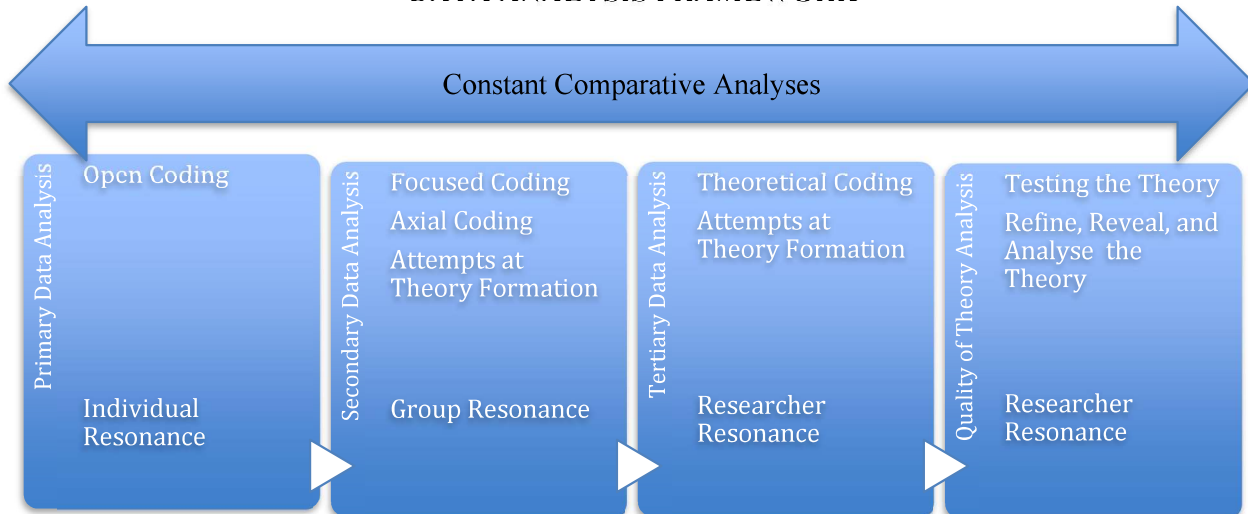
In hindsight, at that point in time, I viewed the methodology as linear as opposed to recursive. This meant that I viewed data collection as one step of the research process as opposed to what now seemed to be a non-stop, recursive activity. While I was aware of some of the terms associated with GTM, like *constant comparative analysis*, *theoretical sensitivity*, *reflexivity* and *theoretical saturation*, I did not really understand their meaning, their role in the data analysis or even in the generation of the theory. Nor did I fully understand that data analysis began at the point data collection started and continued even after all the data was collected. This was in contrast to what I originally perceived as data collection being a single event-a one time get it done activity. It felt like data analysis was occurring all the time, at least in my head, as I was audio recording and engaging with participants during the interview process or while listening to interviews over and over again. In hindsight, I should have been more strategic, precise, and consistent with writing impressions of what I was hearing and collecting at the moment it was happening. Memo writing took place more commonly after the data was collected as opposed to during the data collection process itself. Even so, the notes I made about the data were not organized and simply scribbled on the side of the transcribed interviews or on other forms of data I was reviewing. I knew I needed to uncover what the data meant and work towards my goal of creating a substantive theory in relation to my research question. Amassing huge amounts of data from the interviews, elicited responses, and several forms of extant texts, I needed to move the study along and make sense of what I was listening to (interview transcripts), analysing, and reading. At the same time, I got stuck in my interpretation of what the research methodology was advocating. In short, I concluded that data analysis began with *open* coding, moved to *focused* coding while drawing out properties and dimensions through *axial* coding, and culminated with *selective or theoretical* coding for the purpose of theory generation. These coding terms and procedures were all new to me, and the expectation of *constantly comparing* data with emerging categories that were grounded in the data in the hopes of generating a theory was unnerving.

Reading about research methodologies and engaging others in conversations about doing GTM the *right* way, I knew I had to look at my data line by line and apply codes to the data that answered questions like *what is the data saying? And what does it mean?* (Charmaz, 2006). I did not know at that time how I was going to sort, prioritize, or place significance on some data and to a lesser extent on other data. On top of that, I was stressing about how credible my research would be and immediately defaulted to two constructs from the positivist paradigm – validity and reliability – that seemed pervasive in the research methodology even in relation to qualitative approaches. Discovering terms more aligned with the interpretive paradigm and GTM itself (e.g. fit, grab, rigor, utility, integrity, work, etc.) a sense of relief

occurred one day when I discovered that the verisimilitude of qualitative research can be explained in a number of ways least of which had to do with validity and reliability. The task of the qualitative researcher was to render their findings with enough detail so that the reader can better understand and believe the context and phenomenon under study. Making explicit my decisions about how I used the data would take care of my fears about validity and reliability and ensure that the integrity of my theory could be supported by the data and my decisions about how it was created. Not only would these decisions be made explicitly by me as the researcher but the option for others to scrutinize my findings existed as well. I knew that my interpretive truthfulness could only be accomplished by a rigorous, logical, and organized way of looking at the data. Looking back, I was struggling with how to show the verisimilitude of my study.

I quickly discovered the need to have a data analysis framework one day when I discussed the challenge of being lost in the data with one of my supervisors, Dr. Mark Hirschhorn. He empathized with how I felt and shared his experience when he faced a similar challenge. He explained how he developed and used criteria of significance (CoS) to sort, prioritize, and place significance on his data. This made sense to me, so I adapted how he used CoS (Hirschhorn, 2008) and created a four-phase data analysis framework (Figure 3) to sift through the huge amounts of data that I had to deal with.

**FIGURE 3
DATA ANALYSIS FRAMEWORK**



FINDINGS: PRIMARY AND SECONDARY PHASES OF DATA ANALYSES

From the primary and secondary phases of data analyses using three coding techniques characteristic of grounded theory methodology (open, focused, and axial) along with two criteria of significance (individual and group resonance) to make sense of the data, 15 categories emerged from the data. Table 2 includes the original 15 focused codes and those that were merged with other focused codes because of similar properties and dimensions resulting in eight focused codes or categories of data.

TABLE 2
ORIGINAL FOCUSED CODES WITH MERGED FOCUSED CODES

| ORIGINAL FOCUSED CODES | MERGED FOCUSED CODES | REMAINING FOCUSED CODES |
|-------------------------|---|---------------------------|
| Accountability | Accountability | Accountability |
| Agency | Focus Agency Engagement Seeking Engagement | Agency |
| Conflict | Conflict Having Issues with the SIP Power | Disconnectedness |
| Context | Context | Context |
| Engagement | Merged with Agency | |
| Expectations | Expectations | Expectations |
| Focus | Merged with Accountability | |
| Having issues with SIP | Merged with Conflict and Power | |
| Monitoring | Using Data | Monitoring and Using Data |
| Narrative | Recognizing Differences | Narrative |
| Networking | Networking | Networking |
| Power | Merged with Conflict and Having Issues with the SIP | |
| Recognizing differences | Merged with Narrative | |
| Seeking engagement | Merged with Agency and Engagement | |
| Using data | Merged with Narrative | |
| 15 | TOTAL | 8 |

Discussion of Findings: Primary and Secondary Phases of Data Analyses

As mentioned earlier, with the use of constant comparative analysis permeating all phases of data analyses, various coding procedures and criteria of significance were applied to the data. Beginning with the primary data analysis phase of this research, open coding was applied to the interview data collected. Rich (2012) stipulates that “the idea of open coding is once again tied to the notion of grounded-ness, of letting concepts emerge from the data instead of force-fitting the data to *a priori* theory” (p. 3). Open coding can be done line by line, incident by incident, or by entire section. The main question to ask when applying open codes is: *what is happening in the data?* (Charmaz, 2006). To establish which data were used more centrally in the derivation of my findings, I expanded Hirschkorn’s (2008) use of criteria of significance to

make explicit how I “sorted and prioritized the categories and thus the data” (p. 73). Using *resonance* as a central theme to make sense of the data, during the primary phase of data analysis I applied individual resonance to the data. That is, I examined the data and asked myself which experiences did participants’ give particular meaning or importance to that affected them in a personal, professional, or emotional way as they attempted to make change in their school.

Over 700 open codes were generated from the interview data (N=12) from the first 4 participants. In order to proceed with the data analysis of the remaining interviews (N=15), not to mention the other forms of data collected, I had to find a way to make the data analysis of the remaining interviews more manageable and less time consuming. I needed a more efficient way of coding the data given the voluminous amount of transcribed text and the results of the number of open codes that emerged after the line by line analysis of the interviews from the first four participants (N=742) took place. Attaining this objective would reduce the large amount of open codes to allow the data analysis of the remaining interviews to move along at a faster rate.

I returned to the data and each open code and asked myself- *What was happening in the data?* and *What did it mean?* (Charmaz, 2006). I then assigned a focused code to each open code. Charmaz (2006) defines focused coding:

Focused coding is the second major phase in coding. Focused coding means using the most significant and/or frequent earlier codes to sift through large amounts of data. Focused coding requires decisions about which initial codes make the most analytical sense to categorize your data incisively and completely. (p. 57)

The focused code was either one word or a few words that I thought could shed some light on the properties and dimensions of the focused codes. Properties comment on descriptive questions. Dimensions “answer questions about the variance of such properties” (Rich, 2012, p. 5). An example of a dimension might be the ‘extent of use’, whereas a property might simply be ‘use’. In short, “dimensions measure whereas properties describe” (p. 5). Specifying the properties and dimensions of a category is known as *axial coding*. Proponents indicate that axial coding can strengthen the analysis of a category. Additionally, during the secondary phase of data analysis, I applied group resonance as another criteria of significance to help prioritize and sort the data. This means that as I examined the data I asked myself which experiences among the participants were similar in nature, shared or common.

By returning to the data and extending open codes to focused codes and by using the second criterion of significance, *group resonance* (participant responses that were common experiences among participants), I reduced the number of open codes (N=742) from the first four participant interviews (N=12) to a more manageable number of focused codes (N=96). As noted by Charmaz (2006), “focused coding requires decisions about initial codes which make the most analytical sense to categorize your data incisively and completely” (p. 57).

Although the 96 focused codes were still large in number, the application of the second criterion of significance, group resonance, revealed numerous similarities. For example, I assigned *agency* to all of the open codes that had something to do with the actions taken by teachers or administrators when working with the SIP. The data showed that *agency* had numerous variations (N=33) and was used repeatedly as a focused code but with slight differences. Against this background, other focused codes also revealed variations within each of them as well. This resulted in 15 focused codes or categories of data. Upon closer examination of these focused codes and discovering the relationships that existed within these categories (though axial coding) eight focused codes emerged in the data.

FINDINGS: TERTIARY DATA ANALYSIS

At this point, the identification of eight focused codes with corresponding properties seemed to be waiting for me to make sense of them. Looking closer at the merged focused codes (Table 2) and their properties and searching for connections within and between them in the hopes of moving towards theory

development, another coding technique, known as theoretical coding, was applied to the data. This resulted in five theoretical codes that emerged in the data. Table 3 summarizes this finding and shows the focused codes and their meaning that resulted in the development of the theoretical codes.

TABLE 3
MERGED FOCUS CODES, THEIR MEANING AND THEORETICAL CODES

| MERGED FOCUS CODES | WHAT IT MEANS | THEORETICAL CODES |
|---|---|--------------------------|
| Agency Expectations Monitoring and Using Data | Knowing what to do and empowered to act | Sense of Agency |
| Context Narrative | Knowing why it is important | Sense of Purpose |
| Networking | Knowing what your responsibilities are; knowing how to help one another; keeping each other accountable | Sense of Interdependence |
| Disconnectedness | Knowing how to deal with the challenges that are external to the classroom | Sense of Other |
| Accountability | Knowing how to account for change | Sense of Change |

Discussion of Findings: Tertiary Data Analysis

This phase of data analysis (tertiary) used theoretical coding that led to theoretical development – that is using codes to develop a theory. At this point, I also used the final criterion of significance, known as *researcher resonance*, to continue to make sense of the data. Before discussing the findings further, a brief explanation of *researcher resonance* and *theoretical coding* is needed.

Beginning with *researcher resonance*, I was interested in knowing if I could link what the participants were sharing with me to my collective experiences with school improvement planning. And could these experiences link to the generation of a theory about how teachers and administrators use SIPs to make change in their schools? *Researcher resonance* was based on how the experiences from both the teachers and administrators involved in this study resonated with me as the researcher and if these experiences caused me to think of similar ones from my work as a district administrator with school improvement planning.

Theoretical coding “is a sophisticated level of coding that follows the codes you have selected during focused coding” (Charmaz, 2006, p. 63). She refers to theoretical codes as “conceptualizing” (p. 63). Theoretical coding refers to the process of examining data in theoretical rather than descriptive terms to raise the level of abstraction in the emerging theory. According to Charmaz, “theoretical codes specify possible relationships between categories you have developed in your focused coding” (p. 63). She concludes:

Theoretical codes are integrative; they lend form to the focused codes you have collected. These codes may help you tell an analytic story that has coherence. Hence, these codes not only conceptualize how your substantive codes are related, but also move your analytic story in a theoretical direction. (p. 63)

Discovering the connection between categories, properties, and dimensions leads to *theory development*. The practice of making connections and finding patterns in the data can be facilitated by the

use of models (Strauss & Corbin, 1998). Rich (2012) notes that these models are “diagrammatic displays” (p. 5) that “offer a way to condense what are possibly pages of explanation into a simple figure” (p. 5). Against this background, I re-examined the focused codes many times attempting to develop an emerging theory grounded in the data I analyzed.

Making Sense of My Scribbles

Although not a diagram, I began with a series of scribbled lines from each category that led to a new ‘cluster’ category in the form of a box. In addition, I re-examined the transcribed interviews and selected several responses that supported each category of data. This was the first step I took in shaping my theory. What resulted were five boxes with focused codes, scribbled notes, and selected responses from participants that supported what I thought the data meant. While Table 3 makes explicit the theoretical codes that began to shape a theory about how teachers use SIPs to make change in their schools, the focused codes showed variances within the categories in relation to their properties and dimensions. With this in mind, I chose *extent of use* as a dimension of each theoretical code because the data showed that not all schools, teachers, or administrators participated, acted upon, believed, or supported particular aspects of the SIP. In other words, a central dimension among various properties and therefore its associated theoretical code referred to the *extent of use* – whether something occurred or not. Table 4 identifies the theoretical codes that described what was happening and if change was occurring. The theoretical codes leading to change were renamed to reflect instances when the data showed that change was not taking place.

**TABLE 4
THEORETICAL CODES AND CHANGE**

| Happening and Leading to Change | | Not Happening and Not Leading to Change | |
|---------------------------------|---|--|--|
| Theoretical Codes | What it Means | Theoretical Codes | What it Means |
| Sense of Agency | Knowing what to do | Sense of Complacency (Excuse) | Not knowing what to do-let it pass |
| Sense of Purpose | Knowing why it is important | Sense of Disillusionment | Confusion; how does it fit |
| Sense of Interdependence | Knowing what your responsibilities are; knowing how to help one another; keeping each other accountable | Sense of Isolation | Content to work in their classroom doing what they have always done |
| Sense of Other as Navigable | Knowing how to deal with the challenges that are external to the classroom | Sense of Other as Defeatist | Too much change; can't keep up; this too shall pass attitude |
| Sense of Change | Knowing how to account for change | Sense of Perpetual Motion (Resistance to Change) | Always done this way; not concerned about change; job is to teach and cover the curriculum; keep my house in order |

FINDINGS: QUALITY OF THEORY ANALYSIS

The final level of data analysis took place after the remaining data sources were analysed. These included the remaining interviews, several types of extant texts such as school improvement plans, school

review reports, data from school improvement surveys, and elicited responses, as well as achievement data. No new categories of data emerged as saturation was confirmed.

Discussion of Findings: Quality of Theory Analysis

While the interview data identified five key categories (*sense of purpose, sense of agency, sense of interdependence, sense of the other as navigable, and the sense of change*) to reflect changes taking place along with five additional categories (*sense of complacency, sense of disillusionment, sense of isolation, sense of other as defeatist, and sense of perpetual motion*) to reflect the absence or resistance of change not taking place, along with the meanings from all sources of data used in this study, the plethora of codes used to analyze the data, the connections made among categories, and the logic of justification for the decisions made about the data, a substantive theory about how teachers and administrators use SIPs to make change in schools emerged.

KEY FINDING: THE THEORY

School improvement plans can be used as a means to make change in schools. When teachers and administrators work towards the development of a quality school improvement plan, or when they use it for their work with students and one another, or when they work interdependently while dealing with influences external to their work, changes, especially those relating to the context of school, can take place.

Discussion of Key Finding

In hindsight, my trepidation with the newness of GTM and how to deal with a large amount of data eased largely because of creating this data analysis framework, infusing one or more GTM coding procedures and one or more criteria of significance in each phase of data analysis. This accomplished a few things. First, it gave me a path to follow to apply the coding procedures advocated in the GTM literature. Secondly, by using the CoS I was able to sort and prioritize my data by placing significance and relevance on the data according to the three criteria I used to answer my research question. Finally, by using CoS, my decisions about how I made sense of the data were made explicit. This contributed to the verisimilitude of my study. Most importantly, it was a turning point in my research and provided me with the kick-start I needed to continue. I gained confidence as a researcher and began to develop and trust my interpretive voice which allowed me to generate a theory grounded in the data.

CONCLUSION

In this paper we have acknowledged the challenge qualitative researchers sometimes face when confronted with huge amounts of data to analyze. Making sense of the data can be an overwhelming task particularly at the beginning of the data collection process and for researchers new to inductive, categorical data analysis. We have given voice to this problem by sharing our experiences and showed how we used Criteria of Significance in slightly different ways to make sense of our data. We feel that these narrative portrayals reflect some of our interpretive logic and that by using CoS in qualitative studies, researchers might come to realize, as we have, its benefits.

The first benefit relates to the pragmatics of doing qualitative research. We acknowledge the dual challenge researchers face of making sense of the methodology itself while dealing with the large amounts of data generated when conducting qualitative research. The application of CoS as a strategy to sort, prioritize, and make sense of large amounts of data might provide the necessary direction needed to begin data analysis. Our experiences using CoS provided the impetus for us to feature, exemplify, and draw conclusions from huge amounts of discursive data. Making explicit our decisions about how we made sense of the data leads to another benefit of using CoS.

The second benefit is the CoS conveys to the reader what the researcher counts as relevant data in their pursuit of answering the research question which enhances the transparency of the study by making explicit the researcher's decisions about their data and how they were used in the derivation of their findings. We

feel that the CoS serves as a defensible set of criteria by which data is given increased or decreased value regarding its use in the final analysis and conclusions drawn from a study. The relationship between what and how the data was used, and the conclusions drawn from the data provide consumers of the research insight into how they might apply a similar mechanism in their own data analysis.

The final benefit is to underscore that although all qualitative data is important and significant, researchers must make hard choices regarding what data to feature and draw from the most when deriving conclusions from their data. We posit that using the CoS is a way to make the researcher's decisions about how they made sense of their data explicit. We acknowledge that doing qualitative research and analysing data is not an exact or absolute discipline. As such, we offer CoS as one strategy for the qualitative researcher's toolkit to use to help them make sense of their data and to allow the readers of their research insight into the path and decision-making process when moving from raw data to conclusions.

REFERENCES

- Charmaz, K. (2003). Grounded theory: Objectivist and constructivist methods. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 509-536). Thousand Oaks, CA: Sage.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London, UK: Sage.
- Colley, H., & Diment, K. (2001, December 5-7). *Holistic Research for holistic practice: Making sense of qualitative research data*. Paper presented at the Learning and Skills Development Agency Conference, Research: Making an Impact on Policy and Practice. Cambridge.
- Cranton, P., & Merriam, S.B. (2015). *A guide to researchers and trainers of adults* (3rd ed.). Malabar, FL: Krieger.
- Creswell, J.W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage.
- Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process*. Thousand Oaks, CA: Sage.
- Goodlad, J.I. (1990). Studying the education of educators: From conception to findings. *Phi Delta Kappan*, 71(9), 698.
- Hirschhorn, M. (2008). *Beginning teacher's student teacher relationship experiences* (Unpublished doctoral dissertation). University of Alberta, Edmonton, Alberta.
- Lichtman, M. (2013). *Qualitative research in education: A user's guide* (3rd ed.). London, UK: Sage.
- Morrison, D. (2018). *A grounded theory study of how teachers and administrators use school improvement plans to make change in their schools* (Unpublished doctoral dissertation). University of New Brunswick, Fredericton, New Brunswick.
- Piantanida, M., Tananis, C.A., & Grubs, R.E. (2004). Generating grounded theory of/for educational practice: The journey of three epistemorphs. *International Journal of Qualitative Studies in Education*, 17(3), 325-346.
- Rich, P. (2012). Inside the black box: Revealing the process in applying a grounded theory analysis. *The Qualitative Report*, 17, 1-23.
- Strauss, A.L., & Corbin, J.M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Thousand Oaks, CA: Sage.
- Strauss, A.L., & Corbin, J.M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
- Thorne, S. (2000). Data analysis in qualitative research. *Evidence-Based Notebook*, 3, 69-70.