Practical strategies to avoid the pitfalls in grounded theory research

Naomi Elliott and Joanne Jordan offer ways to develop integrated concepts and practice-based theory

Abstract
How can grounded theory researchers develop integrated conceptual analysis? In this paper, the authors use lessons learnt during a study of clinical judgment in nursing to offer practical strategies that avoid some of the pitfalls commonly encountered during grounded theory research.

Introduction
Grounded theory is a research method frequently used by qualitative researchers (Bryant and Charmaz 2007). In nursing, it has made a major contribution to research and the development of practice-based knowledge (Morse 2001), although methodological problems can arise that limit the research outcomes. The critical literature on grounded theory has identified a number of pitfalls that have the potential to diminish the integrity of grounded theory concepts and associated theory (Morse 2004, Seale 2004, Atkinson and Delamont 2005, Cutcliffe 2005, Rennie and Fergus 2006, Dey 2007). Despite numerous publications on how to undertake grounded theory research (Glaser and Strauss 1967, Glaser 1998, Clarke 2005, Charmaz 2006, Corbin and Strauss 2008), there are no guidelines for researchers on how to avoid these pitfalls or practical strategies to resolve such problems. The purpose of this paper is to offer practical strategies for novice and experienced researchers on how to avoid the pitfalls of grounded theory research. The three pitfalls addressed here are:

- 'Forcing' analysis in the early stages (Glaser 1998, Cutcliffe 2005).

Failing to ensure that the analysis moves beyond narrative description to generate theoretical concepts (Morse 2004).

In view of these challenges, the imperative becomes to identify practical strategies that address the particular risks posed, enable the generation of meaningful concepts and thus enhance the overall validity of analysis.

**Avoid forcing analysis in the early stages**

The overall aim in grounded theory is to generate theory inductively by gathering data about a phenomenon, identifying the key elements and then categorising the relationships of those elements to each other. The first step in analysis – open coding – is critical to induction and the development of concepts because it focuses the analysis on the data. It involves asking three questions of data: ‘What is this a study of? What category does this incident indicate? And what property of what category does this incident indicate?’ (Glaser 1998).

The term ‘open coding’ implies a relatively simple process that is free from the influence of preconceived ideas or predetermined conceptual frameworks, and assumes that researchers adopt a naïve position. For this to happen, Glaser (1998) advised that: ‘Grounded theory’s very strong dicta are a) do not do a literature review in the substantive area and related areas where the research is to be done, and b) when the grounded theory is nearly completed during sorting and writing up, then the literature search in the substantive area can be accomplished and woven into the theory as more data for constant comparison.’

Throughout his various methodological guidelines on grounded theory, Glaser (1992, 1998, 2005) emphasised the importance of avoiding a literature review in the early stages of research. Despite Glaser’s (1998, 2005) assertions, the timing of the literature review in grounded theory has generated much debate and there is no consensus on when it should be carried out (Schreiber 2001, Heath 2006, McGhee et al 2007). In contrast to Glaser, Charmaz and Mitchell (2001) were critical of researchers who delay the literature review.
arguing that 'they are apt to be steeped in specific literature for a variety of purposes beyond a specific research project' since they are active scholars. Moreover, as Holton (2007) and Charmaz (2006) pointed out, there are sources of a priori knowledge other than the literature; for example, researchers carry into the analysis accumulated experiences and preconceptions arising from their discipline or profession.

Irrespective of the intellectual arguments concerning the timing of the literature review, funding and research governance agencies demand a high degree of transparency so the majority of research cannot avoid some form of preliminary literature review. As Cutcliffe (2005) highlighted, a comprehensive literature review is typically necessary to satisfy the requirements of local ethics research committees who insist on a well-defined research question so they can make informed decisions regarding ethical approval. Similarly, research funding bodies need to be satisfied that studies merit financial investment. This requires researchers to present a strong argument supported by relevant literature to establish the need for research.

Notwithstanding the potential for the researcher to obtain a range of sources of contradictory knowledge during a literature review (Charmaz 2006, Holton 2007), there is a risk that knowledge gained from the review colours analysis, influences open coding and consequently weakens any claims to inductive conceptual development.

According to Gilovich (1991), a natural tendency in human reasoning, when examining evidence relevant to a given phenomenon, is an inclination to see what is expected. In other words, people readily accept evidence that validates pre-existing ideas and are found to be less responsive to the implications of new information.

Despite the fact that Glaser’s (1998) ‘open questions’ can guide early analysis, their adoption cannot completely rid the analysis of a priori knowledge gained from a literature review. It is important, therefore, to consider other analytic strategies that can be used in conjunction with open coding to resolve this problem and protect the principles of inductive validity. The strategies suggested in this context are the adoption of in vivo coding, particularly in the early stages of open coding, and starting constant comparative analysis early in the analytic process.
Adoption of in vivo coding

In vivo codes, as defined by Polit and Tatano Beck (2008), use participants' words or are derived directly from the language of the substantive area. Charmaz (2006) said that in vivo codes are used to preserve participants' use of 'specialised terms'. While this is often the case, using in vivo coding only to capture specialist terminology may risk failing to make proper use of the 'ordinary' or 'day-to-day' language used by participants in communication. Such use of in vivo coding to include participants' 'ordinary' words is important because it focuses the analysis on what participants say about issues of importance to them (Box 1). In vivo codes therefore play a crucial role in the early stages of analysis by ensuring that concepts are developed from codes drawn directly from the data.

An over-reliance on in vivo coding can limit the level of analysis to description. In its early stages, in vivo coding is descriptive, but this is crucial to ensuring that subsequent conceptual development is firmly embedded in the narratives provided by participants. Once this conceptual development begins, it is equally crucial that in vivo codes are superseded by categories developed by the analyst that use these codes as the essential building blocks of theory.

Starting constant comparative analysis early

Grounded theory upholds the use of three different types of comparative analysis (Glaser and Strauss 1967):

- Incident to incident.
- Concepts emergent from further incidents gathered in new data.
- Concept to concept.

Each type serves a different purpose and is used at different stages of analysis. Using incident-to-incident comparison (Box 2) in the early stages of analysis can

Box 1. Worked example: using in vivo coding

When coding one of the early interviews, the excerpts: 'her getting to trust me, that she will be able to tell me' and 'I think it's building on the trust and for her to know that in some ways I'm on her side' were analytically related and coded as 'building-up trust in the relationship'. Further on in the same interview the excerpt, 'we will work together through this' was coded as 'working with the patient'.
Excerpts from interview, ‘What I really want to do... is to cement the therapeutic relationship that is going to exist... it’s a bonding alliance’, ‘it protects the relationship... for me it’s absolutely paramount’ and ‘the most important aspect of meeting somebody is establishing a therapeutic relationship, establishing the relationship first and then getting it therapeutic...’ were analytically related and coded ‘building rapport’.

The excerpt, ‘The first thing you do is say hello... give people your time and eye contact and listen to them’ was coded ‘giving individual attention’. In another interview, the excerpt ‘I usually try to work with them on some area that they feel is important even if I don’t think is a priority they might feel it is so. So we try and work around that... and we still have a chat, a social ...and it’s kept fairly superficial’ was coded ‘starting with social chit-chat’.

It was noted that all these codes were related to each other as they indicated different properties of developing the nurse-patient relationship. Eventually, these were subsumed into the concept of ‘intacting therapeutic relationship’, which conveys the meaning of keeping the therapeutic relationship intact.

help to avoid ‘forcing’ since it ensures that open coding does not progress in a simple, linear process but involves the researcher interpreting the meaning of a set of words (one incident) in the context of the whole narrative account (other incidents). In this way, analyst-inspired codes continue to remain true to participants’ overall interpretation of their experiences as they are based on an analysis that identifies issues of importance to participants and establishes how these issues relate to one another. Rennie and Fergus (2006) explored this process, highlighting the dynamic nature of qualitative data analysis. The integrity of the analysis is thus achieved by the researcher working back and forth between the data and the new codes, trying to get a sense of how well the emerging analysis fits with the participants’ accounts from which it is derived.

Importantly, such constant comparative analysis provides a further means of correcting distortions associated with an early literature review. According to Glaser (1998), ‘however accurate or inaccurate the literature data might be, it will be constantly corrected, put in perspective and proportioned in relevance by the constant comparative method. And when the researcher gets into the study and starts generating a grounded theory, the theory will be much stronger than the earlier review because of grounding. It will put the review in its place’.

**Box 2. Worked example: early comparative analysis – incident to incident**

Excerpts from interview, ‘What I really want to do... is to cement the therapeutic relationship that is going to exist... it’s a bonding alliance’, ‘it protects the relationship... for me it’s absolutely paramount’ and ‘the most important aspect of meeting somebody is establishing a therapeutic relationship, establishing the relationship first and then getting it therapeutic...’ were analytically related and coded ‘building rapport’.

The excerpt, ‘The first thing you do is say hello... give people your time and eye contact and listen to them’ was coded ‘giving individual attention’. In another interview, the excerpt ‘I usually try to work with them on some area that they feel is important even if I don’t think is a priority they might feel it is so. So we try and work around that... and we still have a chat, a social ...and it’s kept fairly superficial’ was coded ‘starting with social chit-chat’.

It was noted that all these codes were related to each other as they indicated different properties of developing the nurse-patient relationship. Eventually, these were subsumed into the concept of ‘intacting therapeutic relationship’, which conveys the meaning of keeping the therapeutic relationship intact.
It is difficult to determine how much a priori knowledge has to be engrained that it influences data analysis in ways that are more subtle and more difficult to make apparent and ‘repair’. However, using these two strategies helps to ensure that awareness of the literature can fade as data analysis begins to yield codes, properties and concepts, and these become the major influences on subsequent development of theory.

**Moving beyond fragmentation and non-integrated coding**
The second pitfall commonly accounted to grounded theory research is that the analysis over-relies on coding and the concomitant development of an index of fragmented codes rather than an overarching integrated concept (Seale 2004, Atkinson and Delamont 2005, Rennie and Fergus 2006, Dey 2007) – in other words, there is a risk that researchers limit data analysis to simply accumulating fragmented material and sorting it under thematic headings. The strategies suggested to avoid fragmented analysis are:
- An ability to live with uncertainty in analysis.
- Comparing emerging concepts to further incidents.
- A readiness to revise emergent concepts.

**Living with uncertainty**
It is certainly the case that open coding produces a long list of discrete codes. However, this list is actually a valid stage in the development of a fully integrated coding scheme and subsequent theoretical framework. In its early stages, analysis can be like trying to make a jigsaw without knowing what the individual pieces look like, how they fit together or what the final picture will look like. It is important to recognise that the feeling of ‘uncertainty’ this lack of knowledge produces is a necessary part of the process of inductive theory generation. Consequently, there should be no rush to bring open coding to a premature close – the codes generated at this stage are crucial building blocks in the generation of integrated concepts and, as such, merit extended consideration.

**Comparing emerging concept to incident and to concept**
As outlined earlier, the iterative process of constant comparative analysis begins by comparing incident to incident. But as it progresses, it is the increasingly
abstract process of comparing concept to incident and concept to concept that further integrates coding. The over-riding aim in grounded theory is to generate a theory about a concept, which highlights the point that conceptualisation involves more than the naming of concepts – it involves integrating the component concepts into a theory.

An important step in grounded theory analysis, therefore, is to re-examine the data, looking specifically for links between concepts and incidents and different concepts so that the relative positions of concepts in the emerging overall theory are identified. It is through the process of comparing concept to incident (Box 3) that the researcher can check to see if further incidents fit with the newly developed concepts and, in so doing, ensure that the concepts are capable of accounting for all related incidents in the data. The next stage of comparative analysis, concept-to-concept (Box 4), is particularly important to enable the researcher to identify the relationships between concepts as well as their relative positions in the emerging overall theory.

Readiness to revise emergent concepts
While the technical procedures involved in constant comparative analysis are made explicit in grounded theory textbooks (Glaser and Strauss 1967, Glaser 1998), less attention is paid to the other necessary attributes of analysis, particularly those associated with the researcher. Throughout the analytic process, the researcher needs to be ready to revise and/or discard conceptual categories.

**Box 3. Worked example: comparative analysis – emergent concept to incident**

The emergent concept of 'intacting therapeutic relationship', keeping the therapeutic relationship intact, and the category 'avoiding break-up-able moments' were used when analysing later interview accounts to see if the concept and the category were supported as further data were gathered.

These excerpts from later interview accounts, 'I can’t push [treatment] too far because the patient would just disengage' and 'I had to make a judgment as to how far to explore this with this client and not push the client away' and 'you get a feel that they are becoming hostile when you reach that particular subject or that area and then you pull back a bit, and go around it in another way' provided further examples of incidents that added support to the category 'avoiding break-up-able moments'.
Box 4. Worked example: comparative analysis – concept to concept

Intacting therapeutic relationship

Situated patterning

Theoretical memo: relationship between the concepts of ‘intacting therapeutic relationship’ and ‘situated patterning’.
- Gradual assessment of the patient as the therapeutic relationship develops.
- Allowing trust to develop over time before completing patient assessment.
- Moderating the pace of patient assessment to keep a therapeutic relationship with the patient.
- The bringing to the surface of sensitive information from the patient is helped by allowing time for the therapeutic relationship to develop.
- Surfacing information about the patient problem is easier when a therapeutic relationship is established.

already developed in the light of what is subsequently revealed. This underlying attitude of staying open to possibilities in the data – so as to be able to identify and respond to new ideas emerging from the iterative process of data collection and analysis and revise concepts accordingly – is a more subtle, yet equally important component of constant comparative analysis.

Such requirements are implicit in Glaser’s (1978) notion of developing an ‘emergent fit’. In grounded theory, researchers must be prepared to modify generated concepts so that new data, including data that challenges ongoing conceptual development, can be successfully adapted into the emerging theory. In other words, early codes emerging from the initial data analysis may not be substantiated by data collected during subsequent interviews and, ultimately, may be dropped from concept development. Glaser (1978) described this process as concepts earning their way into grounded theory. On this basis, it is clear that grounded theory’s use of constant comparative analysis is fundamental to the credibility of the emergent theory (Glaser 1998, Seale 1999) and to the development of integrated concepts.

Moving beyond description to theoretical conceptualisation

The third pitfall commonly ascribed to grounded theory is failing to ensure that the analysis moves beyond narrative description to generate theoretical
concepts. The main strategy suggested to ensure that analysis avoids this pitfall and succeeds in generating theoretical conceptualisation is theoretical coding.

There are two different views on the status of theoretical concepts. For Glaser (2001), 'the most important property of conceptualisation for grounded theory is that it is abstract of time, place and people'. This statement implies that grounded theory concepts should be removed from their original context, so increasing their applicability to other contexts. Morse (2004) presents an alternative view that endorses the need for concepts to be abstract yet kept in context because '...we must always be able to define it [the concept] and give examples to illustrate its meaning in a particular context'. Morse thus argues that concepts cannot exist in a vacuum and need to be constructed in ways that enable them to be positioned in a hierarchy of relationships to form a general and significant theory. One way of reconciling these two viewpoints is through the use of theoretical coding - the process of bringing together the constituent parts of an analysis into one conceptual 'whole' (Walker and Myrick 2006, Stern 2007, Holton 2007, Polit and Tatano Beck 2008).

**Theoretical coding**

Glaser (1978, 2005) has developed sets of theoretical codes that can be used by researchers to advance conceptualisation. According to Glaser (1978),

<table>
<thead>
<tr>
<th>Box 5. Theoretical coding – examples of applying the ‘six Cs’ to coding the patient-nurse relationship data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contexts:</strong></td>
</tr>
<tr>
<td><strong>Causes:</strong></td>
</tr>
<tr>
<td><strong>Contingencies:</strong></td>
</tr>
<tr>
<td><strong>Consequences:</strong></td>
</tr>
<tr>
<td><strong>Covariances:</strong></td>
</tr>
<tr>
<td><strong>Conditions:</strong></td>
</tr>
</tbody>
</table>
Avoid forcing analysis in the early stages
- Use in vivo coding in the early stages of coding.
- Comparison of incident to incident.
- Delay reviewing related literature until the core category has been developed from the data analysis.
- Wait until the core category is identified before starting theoretical coding.

Moving beyond fragmentation and non-integrated coding
- Constant comparative analysis: concept to incidents.
- Constant comparative analysis: concept to concept.
- Be prepared to revise ‘emergent’ concepts.

Moving beyond description to theoretical conceptualisation
- Develop the core category first and then use theoretical coding by applying one of Glaser’s set of codes.

Using one of these sets provides an analytical framework that enables researchers ‘... to weave the fractured story back together again’. The first step involves researchers familiarising themselves with the sets and selecting that which they consider most appropriate to the analysis. The second step involves applying the selected code set and re-analysing the data. In our analysis, this entailed looking specifically for information relating to the ‘six Cs’ theoretical codes. Applying these to data analysis meant that it focused on the relationships between contexts, causes, contingencies, consequences, covariances and conditions identified in the data, and provided an internal, structural framework that further served to link component parts of the analysis together (Box 5).

The timing of theoretical coding is important. If theoretical coding is started after the core concept has been developed – for example, by identifying its component concepts, categories and properties – it does not overly direct the generation and integration of concepts. In this context, it provides a kind of post hoc review process, helping to ensure a complete all-encompassing analysis.

By contrast, if theoretical coding is used earlier in data analysis, it risks producing an overly formulaic, procedural and ultimately forced interpretation. This is because the analysis would have been ‘directed’ by what the theoretical set of codes suggested was important, rather than being led by the immediate content of data.
Conclusion
Grounded theory provides researchers and practitioners with a viable means of generating theory that is grounded in everyday clinical practice, although the challenge is to apply grounded theory in a way that responds to current criticisms of the method and demands arising from the research practicum. The practical strategies summarised in Box 6 provide a means of enabling researchers to avoid the potential pitfalls of grounded theory research and thus to develop integrated concepts and practice-based theory.

Naomi Elliott PhD, MSc, BNS, RGN, RNT is a lecturer at the School of Nursing and Midwifery, Trinity College Dublin
Joanne Jordan PhD, BA(Hons) is a lecturer at the School of Nursing and Midwifery Research Unit, Queen's University Belfast

This article has been subject to double-blind review and checked using antiplagiarism software

References


Online archive
For related articles and author guidelines visit our online archive at www.nurseresearcher.co.uk