

Initial Experiences of Using Grounded Theory Research in Computer Programming Education

Enda Dunican

Computing Department,
Institute of Technology Carlow, Ireland,

`Enda.Dunican@itcarlow.ie`

Abstract. This paper presents to the reader the issues encountered to date¹ on a grounded theory study conducted in the area of computer programming education. In this regard, it presents a list of factors namely, underlying issues, methodological clarity and context that are viewed as critical considerations whilst undertaking qualitative research. Furthermore, each of these factors are discussed in detail in terms of how they impinged on the research conducted thus far. This discussion not only highlights the extent to which they impinge on the research but will also show that they are interlinked. Finally, given the many overlaps between grounded theory and other qualitative research methods, it will be suggested that the factors uncovered in this research may be applicable to other qualitative methods and it is hoped that these may be helpful for prospective qualitative researchers embarking on similar types of projects.

1. Introduction

Novice computer programming education in Irish third level institutes is an area of research that has become more popular in recent years. To date, this research mainly depended on quantitative methods [2], [5], & [8]. The research described in this paper has deviated from the norm by engaging in a qualitative research study using the grounded theory research method. In particular, this paper presents an interim report that is based on a large qualitative study engaged upon with 30 participants from four Irish third-level institutions. The research title of this project is *'Learning Computer Programming in Irish Third-Level Institutions: A Study of First Year Students' Expe-*

¹ This set of issues is not being presented as exhaustive, but rather pertains to those factors that manifested emerged during the early stages of a grounded theory study with 28 participants across four separate research sites

riences'. The aim of this paper is to describe in detail the practical issues encountered by the author in terms of study preparation, data collection and initial analysis. To this end, it presents the prospective qualitative researcher with a list of factors that focus on issues of research question, context and underlying methodological concerns, all of which were found to be critical during the early stages of a grounded theory study. It should be noted that this paper describes issues encountered at the early stages of this grounded theory research project. Emergent theory from the study will be presented by the author in subsequent papers. It is hoped that discussion of these factors will in some small way help prospective qualitative researchers at the beginning of their research when they are trying to find their feet and tie down their ideas.

2. Grounded Theory

Grounded theory is an inductive qualitative research method. Using this method, the grounded theory researcher begins by collecting data in the field and lets the theory emerge or emanate from the data. In this regard, it is postulated that the theory is actually *grounded* in the data. Data is usually in the form of interview transcripts or observational notes. Research subjects are chosen using *theoretical sampling* which is based on their potential for contribution to the development of theory. Conducting grounded theory research entails a number of levels of coding and analysis.

Open coding examines the text for items of interest, with the ultimate aim of accumulating codes into categories. The technique used in the open coding process is that of *microanalysis* as described by Strauss and Corbin [7]. Here, the researcher analyses the interview transcripts on a line-by-line basis and inserts initial codes into the margins of the text. Strauss and Corbin [7] point out the researcher must analyse the data as opposed to reading it in a general way. During this analysis the researcher uses the *constant comparative* approach where they constantly compare new instances of the category with those already encountered until he/she saturates the category (i.e. no new insights in the category can be gained from the data).

On completion of open coding, the researcher engages in the *axial coding* process. Axial coding entails relating categories to their sub-categories around the axis of a central category, based on linkages between their properties. According to Goulding [15], axial coding entails moving to a higher level of abstraction by identifying relationships between categories and identification of a core category around which other concepts revolve and these higher level categories (phenomena) form the basis for the construction of theory. To aid this axial coding process the coding *paradigm* developed by Strauss and Corbin [7] is used. This paradigm enables the researcher to analyse a phenomenon (i.e. higher level category) from a number of perspectives in terms of its context and its relationship with its sub-categories. Once again the use of memoing is an essential part of this process. Borgatti [3] advises the researcher to generate a *frame* (see table 1) comprising the various elements of the paradigm model and associated descriptions in the form of memos.

Element	Description (Memo)
Phenomenon	
Causal Conditions	
Context	
Intervening Conditions	
Action/Interaction Strategies	
Outcome/Consequences	

Table 1: Paradigm Frame (Borgatti, 2004)

Finally, *selective coding* entails identifying a central phenomenon and relating central categories to it using statements of relationships. Very often, in selective coding, a ‘storyline’ is generated that narrates the categories and their relationships [7]. The net outcome of grounded theory research is a theory that contains a central phenomenon, its causal conditions, its intervening conditions and its consequences [7].

3. A Pragmatic Approach to Qualitative Research

The approach to qualitative research described in this paper is based on pragmatism i.e. dealing with issues in a common sense and logical manner as they arise. In this regard, fig 1 illustrates the four key factors (uncovered in this research so far) that the prospective qualitative researcher might consider before, during and after a research project. Firstly, it is important that the researcher to be aware of any *underlying issues* that are associated with the chosen qualitative research method, that if not addressed and understood clearly may weaken the final research product, be it theory or phenomena. Secondly, *methodological clarity* relates to a clear and precise understanding of the chosen methodology and its implementation is seen as essential in the quest for sound research. Finally, the researcher must always take the *context* of his/her research into consideration in terms of nature of participants, unavoidable constraints, realistic means of data collection and any other pertinent issues given the fact that one size can never fit all.

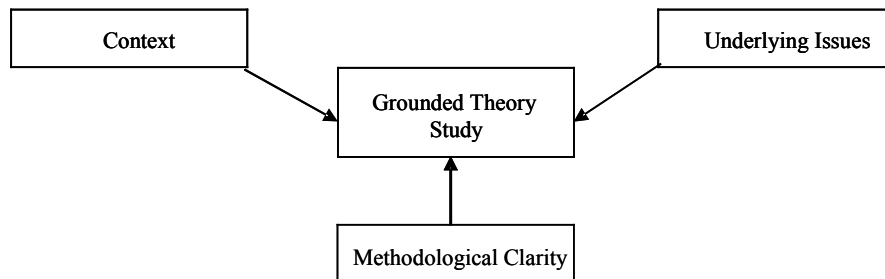


Fig 1 Pragmatic Approach to Qualitative Research

Before the research is undertaken it is vital that the research question is clear and unambiguous as it drives all the other activities that will subsequently take place (the nature of a research question in grounded theory research will be discussed later in this paper). Furthermore, at this stage issues pertaining to underlying methodological issues, methodological clarity and context must be addressed as ignoring them may only serve to result in them coming back to haunt the researcher at a later date. In software development, this is analogous to what stage in the life-cycle a defect is uncovered, the earlier it is found, the cheaper it is to correct. During the data collection and analysis, the researcher must constantly engage in self-check mechanisms to determine if they have deviated from the research question, prescribed methodological implementation or underlying issues identified at the proposal stage. An example here would be the researcher falling unwittingly into the trap of theory forcing, which is an underlying issue that can manifest itself at both the data collection stage and the analysis stage of a grounded theory study.

Reflection Type	Concept
Technical Rationality	Examining one's use of skills and behaviours used in the classroom
Reflection-in-Action	Dealing with on-the-spot professional problems as they occur. Thinking can be recalled and shared subsequently
Reflection-on-Action	Recalling one's teaching after the class. Educator gives reasons for his/her actions or behaviours
Reflection-for-Action	Proactive thinking in order to guide future action

Table 2. Approaches to Reflection (Farrell 1998)

Finally, after the data collection and analysis is complete, it is advisable that the researcher reflect on their methodological application in order to ascertain if they have deviated from any of the four dimensions in such a way that would threaten the legitimacy of the research product. In fact, the experience gained so far in this research suggests that the qualitative researcher can benefit from reflection. Table 2 illustrates

the different types of reflection as described by Farrell [9]. Closer inspection of each of these suggests that they have considerable resonance with the researcher engaging in qualitative research and attempting to utilize the model highlighted in figure 1. For example, engaging in reflection-for-action may aid the researcher in refinement of their data collection mechanisms for subsequent rounds of qualitative interviews. Likewise, engaging in reflection-on-action may help the researcher determine whether they have complied with conventional methodological implementation in grounded theory. If this isn't the case they may document how and why they have deviated from prescribed implementation, e.g. their approach to sampling may not be theoretical in its purest sense and there may be very good rationale for this which should be documented. The following sections will now look at each of these factors in turn.

3.1 Underlying Issues

As well as understanding the methodological issues involved in grounded theory, it is also important for the researcher to be aware of the existing variants of the method. The original version of grounded theory [13] was developed as an inductive, theory-generating method that differed from conventional deductive, hypothesis testing approaches. Since the original version of the method was developed two main variants have emerged that are based on different directions taken by its originators Glaser [12] and Strauss [7]. Babchuck [1], in a clear and succinct manner, summarises the essential differences between these variants. Essentially, the Glaser approach takes the view that the informants world should emerge naturally from both data collection and analysis without being constrained with overly-prescriptive collection and analysis processes in the Strauss and Corbin model. Babchuck [1] describes this as very much a 'laissez-faire' approach to process. In particular, it espouses total flexibility and advises the researcher against unnecessary constraints like tape-recorders for interviews or prior reading of relevant literature (for fear that it would cloud their analysis with preconceived constructs).

On the other hand, the Strauss & Corbin [7] approach is characterized as a more prescribed approach that contains more formal models and procedures to generate theory appears to be 'concerned with producing a detailed description of the cultural scene' ([1], p.3). In fact, Glaser [12], [14] refers to this as a 'contextual description'. This detailed description is achieved by using the paradigm model [7] where the researcher attempts to describe a phenomenon in terms of its causal conditions, context etc (as depicted in table 1). Glaser [12] is vociferous in his vehement objection to this paradigm model :

'In actuality it teaches the analyst to make a full conceptual description on data with no questions about whether the links are relevant to any emerging theory that really explains how the participants process their main concerns. And the more the analyst practices the use of this model, the more he will exclude forever his ability to respond to any theoretical code that may

emerge and become relevant. He will always just see a condition or consequence irrespective of relevance and stake his professional identity on it' (p. 61)

Glaser goes on to assert that the Strauss and Corbin approach is based on Strauss forcing the data into his own 'pet framework' (p.64). Furthermore, Glaser [14] describes what he means by 'contextual description' :

'One concept is generated and then the researcher spends the rest of his time describing it and describing it with incident after incident. There is little or no constant comparative work to generate conceptual properties of the category based on the inter-changeability of indices and conceptual saturation' (p.21).

In this regard, Glaser presents the reader with his vehement distinction between conceptual description and conceptual analysis. In particular, he asserts that using the Strauss and Corbin model the researcher continually over-describes categories with a distinct absence of analysis and ends up with story after story being forced into the concept. In this regard he refers to *theory forcing* as opposed to *theory emergence*.

The issue of the research question is also a cause of divergence between the two approaches. According to Glaser [12] 'the research question in a grounded theory study is not a statement that identifies the phenomenon to be studied' (p.25) but rather emerges out of open coding and sampling based on data that wasn't forced with pre-conceived theory forcing questions whereby the researcher begins his/her study "with the abstract wonderment of what is going on that is an issue and how it is handled' (p. 22). Strauss & Corbin [7] on the other hand, assert that having a research question allows the researcher to stay focused in the midst of masses of data and in a qualitative study this question will be broad and open-ended. More importantly, they contend that whatever the source of the research question that it is 'important that the researcher have enthusiasm for the subject because he or she will have to live with it for some time' (p. 53).

Given these two distinct variants of grounded theory it is necessary for the researcher to make it clear at the outset which version they are using and to remain aligned with the tenets of that model. In fact, Goulding [15] contends that this is a crucial issue and states that 'care should be taken to decide which method best suits the researcher's personality and preferred modes of working, before embarking on the research. (p.48). Whilst this contention holds merit, the experience gained to date in this study suggests that it doesn't go far enough. In particular, choice of method should be determined by the nature of the research question and the context of the research (context here refers to the research setting, previous work (if any), characteristics of participants engaging in the study etc).

In the context of the discussion above, it is important that the qualitative researcher define clearly which variant of grounded theory is being used and clarifies why it is being used. This research project has utilised the Strauss & Corbin [7] variant of

grounded theory. After careful consideration, this variant was chosen because it appeared to be more aligned with both the research question and the context of the research. As mentioned in the introduction, the research question in this project relates to the experiences of first year programming students in Irish third-level institutions. The reasons for the choice of the Strauss and Corbin in relation to this project will now be outlined.

Firstly, as Babchuck [1] suggests, whilst in the Glaser approach 'the informant's world should emerge naturally from the analysis with little or no attention to process' (p. 3) the Strauss and Corbin approach provides insights into the realities of the participants it also provides 'a detailed description of the cultural scene' (p.3). In this regard, the Strauss & Corbin approach may be viewed as more aligned to answering the research question where the outputs of the research containing both theory and detailed cultural description may prove invaluable for improving programming teaching in the future (which happens to be one of the main objectives of this research). In particular, the paradigm model presented in table 1 will enable educators to not only see what phenomena are at play in learning computer programming but will also give them a detailed insight into the context and perceived causal conditions of these phenomena. This is considered by the author as an important part of improving and refining programming education. Secondly, the detailed attention to process in this method seemed as a good starting point for a novice grounded theory researcher. Thirdly, within the context of a an Education Doctorate study where the researcher is required to possess an specific research question at the thesis proposal stage, once again one witnesses greater alignment with the Strauss & Corbin [7]. Within this method, the researcher is recommended to have a clear research question that provides the focus and clarity. In contrast, within the context of a doctoral study, whilst the Glaser [12] approach of standing back in abstract wonderment and letting the research question emerge undoubtedly has pure inductive qualities, it is unlikely to be a luxury that can be afforded by an education doctorate researcher who has a limited time-frame within an academic year to interview students and other non-avoidable logistical constraints.

In summary, this section has highlighted the differences between the two main variants of grounded theory methodology. It has shown that the Strauss & Corbin [7] approach was chosen for its conceptual alignment with both the research question and the context within which the research took place. Finally it must be stated that analysis of literature on both methods indicated to the author that both variants have their own merits and that either one could be used in a study of this nature. Just because this research study is more aligned with the Strauss and Corbin approach, it should be taken to in no way detract from the Glaser approach. In fact a useful exercise in the future would be to use the Glaser approach in the same type of study and discuss the issues that arise in terms of methodological application.

3.2 Methodological Clarity

Methodological clarity, as proposed in this paper has two dimensions. Firstly it relates to the researcher possessing a clear and unambiguous understanding of the chosen research method and its methodological implementation. Secondly, it relates to the requirement on the researcher to report in a clear and precise manner the exact methodological steps they undertook in their data collection and analysis. Each of these aspects will now be discussed.

Given the background of grounded theory and its two main methodological variants, it is not surprising that some grounded theory projects may present a mix of both variants. Whilst it may appear obvious which version of grounded theory is being used i.e. utilisation of the paradigm model demonstrates alignment to the Strauss and Corbin approach, some researchers for whatever reason, may engage in what has been described in other research as ‘method slurring’ [4] whereby techniques from differing qualitative research methods are combined in a non-formal manner. If this is undertaken without due regard to the core principles of the respective technique the result may be a ‘sloppy mismatch’ ([16], p.15). In this particular context the differing methods are actually variants of one method i.e. grounded theory. Method slurring, in this context might entail combining parts of each variant without stating why the chosen combination addresses the research question in a better or more meaningful way than sticking to one variant. Whilst the term ‘slurring’ has negative connotations, there may indeed be certain situations where using concepts from another variant can result in a more reliable study. For example, in this project whilst the issue of avoiding theory forcing is not pushed very much in the Strauss and Corbin variant, it is very much an issue with Glaser. In the project undertaken, time was taken to study the issues relating to theory forcing as described by Glaser and this was seen as strengthening the project given that both versions subscribe to inductive theory generation where neither hypotheses nor theory are generated *a priori*. Predetermination of categories or specific questioning to crudely saturate a category could have the negative impact of potential forcing of theory. Familiarity with Glaser’s stance on this enabled the forcing issue to be addressed and avoided. It was decided in this research that slurring of methods other than careful consideration of the forcing issue might leave the resultant theory open to criticism of the nature of ‘methodological transgression’ as highlighted by Goulding[15].

Practical experience gained in this research has found that some common aspects of each variant can be served whilst adhering to the methodological requirements of one. For example, both variants espouse the fundamental issue of inductive theory generation whereby the theory generated ‘fits’ the actual data collected and is not derived from bias or other non-data-related sources. Despite the fact that Glaser harshly criticises the paradigm model of Strauss and Corbin, the work conducted in this research to date suggests that once the data presented in the paradigm model emanates from the data and is not a result of forced interview questioning then it conforms to

the inductive requirement. As an illustration of this issue table 3 highlights a paradigm model representation of an abstracted category ‘solution strategies’ that was developed during the pilot study of this research.

Element	Description (Memo)
Phenomenon	Solution Strategies
Causal Conditions	Student encounters a difficult issue in programming that doesn’t make sense straight away.
Context	Understanding a programming concept like this is a prerequisite for progression in a project/assignment. For example the ability to reference a cell position in a two-dimensional array.
Intervening Conditions	Lecture materials, i.e. notes and code examples will be given to the student illustrating how the concept works, accompanied by a relevant explanation by the lecturer.
Action/Interaction Strategies	Look back through notes to find an illustrative example. Persist in analysing the examples on an iterative basis until the concept is understood. Seek assistance from another individual who understands the concept.
Outcome/Consequences	Student either understands the concept or decides it is too difficult and looks to use another concept in its place if possible (i.e. a concept that they do understand).

Table 3 Paradigm Model

In this example in table 3 we have a situation where students discussed their approach to dealing with the computer programming subject. In an open-ended interview it is normal for the researcher to ask questions on when and how a certain situation manifests itself, e.g. ‘when do you find this?’ or ‘how do you deal with this?’. Answers to these types of questions can be presented in the paradigm frame as in table 3. Whilst being aware of Glaser’s vehement objection to this paradigm model, this research has found that presenting categories uncovered in the data in this format gives educators an interesting insight into learning experiences in computer programming in terms of how, when and why they occur. In the author’s opinion this can serve as useful material in programming education reform. Furthermore, once the information presented in the paradigm frame is extracted from the data (as a result of how and when questions being asked) then it cannot be construed that it is forced. If the structure of the paradigm model forced the researcher into asking specific paradigm-specific questions to conform to its structure then one could argue the forcing issue. However, if

we look at the last three elements in the paradigm frame we can see that their contents can be a product of natural unforced questioning in an in-depth interview as illustrated in table 4. This table presents an example of the first aspect of methodological clarity as stated earlier in this section, i.e. that of the researcher having a clear and precise understanding of their chosen variant and confidence in the fact that they are not deviating from inductive theory generation. As is clear from an inspection of table 4, none of the questions listed could be construed as leading or forcing, but rather stem out of a desire to get a more in-depth understanding of the participant's experience.

Intervening Conditions	When do you find this is the case?
Action/Interaction Strategies	When in this situation, how do you deal with it?
Outcome/Consequences	How do you find this works for you?

Table 4 Natural unforced questioning

In terms of the second issue of methodological clarity, this research has found that it is a useful exercise for the researcher to constantly reflect back on their methodology on an iterative basis in a way not too dissimilar to reflection-on-action as highlighted in table 2. It has found to be a useful exercise to document each stage of methodological significance and to reflect on its correctness with the purpose of reporting this in a clear and precise manner in thesis or research reporting work. For example, with the Strauss and Corbin approach, the central category is usually an abstract category (a higher order category that describes a number of concrete categories uncovered in the data e.g. see fig 3 where the higher order category 'solutions strategies' subsumes the lower level categories 'illustrative examples', 'assistance' and 'persistence' all of which were uncovered in micro-analysis of the data. Data analysis conducted to date has found that the central category may in fact be an in-vivo code with the data. An in-vivo code is a code based on a verbatim term uncovered in one or more data sources. For example, an in-vivo code that might replace 'assistance' might be 'look for help'. During the early analysis of the first sets of data, potential in-vivo codes uncovered were 'struggling', 'skip-over' and 'sink-in time'. Furthermore, early analysis has suggested that it may be possible that an in-vivo code uncovered in the data may rise to the surface to the highest level of abstraction once it can pull all the other categories together. In terms of issues encountered like this, the novice researcher is advised to constantly return back to the methodology chapters in their relevant textbook.

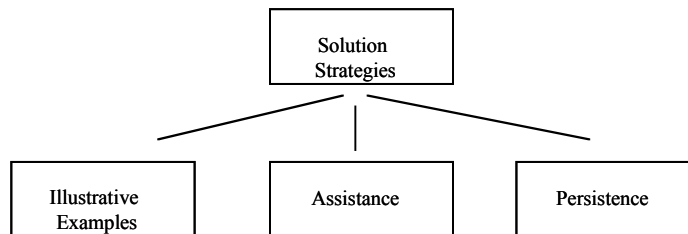


Fig 2. Abstract Categories

3.3 Context

The final factor presented in fig 4 is that of context. As with the other factors presented in this paper, the context of research is an issue that cannot be ignored. In fact Strauss and Corbin [7] explicitly state that ‘by context we mean the conditional background or situation in which the event is embedded’ (p.106). In this project the context of the study was first year programming students in Irish third-level institutions. It has been found that the primary issue in this research relating to context related to data collection. In particular, the nature of the participants, their location and logistical constraints are contextual issues that cannot be ignored by the researcher and must be embraced and dealt with as effectively as possible. One example of context in this research pertained to the nature of the participants, who were primarily school leavers for whom the open-ended interview they undertook may have been their first ever interview. In this regard, great care was taken to make the interviewing environment as non-intimidating as possible. For example, a deliberate decision was taken to visit each research site in person to present the nature of the project to the students. This is recommended by Seidman [17] where he states that ‘building the interviewing relationship begins the moment the potential participant hears of the study’ (p.39). It was found that this approach had positive implications for the project. In specific terms, students who volunteered to take part in the study were those who felt comfortable with the researcher and the nature of data collection in the project as it was described to them in the first contact session. It was also found that this comfortableness extended to the interview and focus-group sessions where the purpose and the ‘setting at ease’ issues were re-iterated to the participant. Deliberate action was taken to inform the participant that the interview questions were not being framed as knowledge-level acquisition questions but rather were borne out of a genuine need to explore the participants’ own personal experience of computer programming. Furthermore, given the age and background of most of the participants deliberate action was taken to ensure that potential off-putting activities like displaying an air of aloofness, or reacting incredulously to certain answers, were avoided.

Furthermore, in order to get the optimal recounting of the participants' experiences, all sites were visited in the space of a couple of weeks. Furthermore, due care was taken not to interview at a time not too close to examinations and not too early in the term where participants wouldn't have had exposure to the full first year programming experience. In this context, the researcher should also make clear the impact the context of the study had on the data collection and analysis. In grounded theory, theoretical saturation requires the researcher to continue data collection in the field until one gets to a stage where no new information about the categories is being uncovered. In the experience garnered in this project, logistical and contextual constraints make this quite difficult. In the context of third-level students, it would be unfair to ask participants to attend subsequent interviews at a time close to their exams. In this situation all the researcher can do is present the categories that reached saturation point (if this is ever possible given the intrinsic evolution of many theories) and perhaps identify others that are short of saturation but may be looked at in future research. Just as true randomness may be difficult to achieve in quantitative studies, issues like pure theoretical saturation in qualitative studies are also not without difficulty.

4. Impact of Factors on the Study So Far

To date, the factors outlined in this paper have significantly impacted the research conducted so far. Section 3 in this paper discussed in detail how underlying issues in grounded theory impacted the study with particular reference to the variant of grounded theory chosen. The Strauss and Corbin variant advises the researcher to have decided on a research question in advance and this contrasts to the Glaser approach of letting both the research question and its subsequent theory emerge from open-ended interviewing. In this study, the research question was decided upon in advance but its nature was such that whilst it was specific, it was open enough to allow unforced theory to emerge and facilitate any experiences of programming to emerge. This detailed consideration of the nature of the research question paid off for two main reasons. Firstly, the initial aim of the author was to seek out participants who are experiencing significant difficulty with programming. In hindsight, this might have had the effect of categorising participants in advance and to an extent pre-determining the theory in a forced manner. In fact, actual data collection found that all students experienced difficulty, albeit of differing levels and extents and it was the way the dealt with that difficulty emerged as of significance to the study. Secondly, when visiting potential research sites and talking to students, it was important to be able to convey to them the exact nature of what the research was about.

When engaging in data collection and initial analysis methodological clarity was a constant issue where the author constantly referred back to the Strauss and Corbin text in order to ensure that correct procedures were being followed. As mentioned in section 3 the aim of the researcher in data collection is to persist with open-ended questions in order to facilitate unforced, natural theory emergence. At times this was quite

