Learning to do Phenomenography: A reflective discussion

Gerlese Åkerlind, John Bowden and Pam Green

What is it like to learn to do phenomenographic research? What helps and hinders? These are the questions addressed in the preceding stories, from the perspective of five individual researchers. In order to help readers integrate the individual stories and the range of response to these questions within them, all of the authors met to have a discussion of their reactions, as a team. Key highlights of that discussion are presented in this chapter, which also draws liberally on relevant sections of the individual stories.

During the course of the team discussion, the purpose of presenting individual stories was reaffirmed. The stories focus on individuals’ experiences of doing phenomenographic research, especially for the first time. In that sense, the stories are intended to be strongly experiential and practical in nature. They are not intended to provide authoritative accounts of phenomenographic practice or theory (such accounts are available elsewhere in the literature, such as Marton & Booth, 1997; Bowden & Walsh, 2000). Indeed, some of the story authors emphasised that they did not feel expert in phenomenography (inevitable when doing it for the first time) and would feel very uncomfortable if their stories were read as expert accounts. The variation in the language used by the authors to describe phenomenography and phenomenographic practice is an illustration of this, often reflecting their methodological background more than standard phenomenographic terminology.

In line with the practical focus of this monograph, the remainder of this chapter is divided into a discussion of the following issues:

- how different methodological backgrounds can help and hinder learning to do phenomenographic research
- phenomenographic practice at the macro level
– common difficulties experienced by novices
• phenomenographic practice at the micro level
  – key practices in phenomenographic research
  – key questions that phenomenographic researchers need to answer for themselves.

How different methodological backgrounds can help and hinder

Given that the RMIT team members came from quite an array of methodological backgrounds, early meetings of the research team often focused on comparisons between research methods. Although all came from a basically similar research paradigm or theoretical position that could be roughly labelled as ‘interpretivist’, each brought differences with respect to specific approaches and methods used previously.

In his story, John Bowden describes a shift among novices over time, from trying to understand phenomenography externally (by focusing on the ways in which it is similar to, or different from, the research approaches with which they are more familiar) to trying to understand phenomenography internally by focusing on internal consistency within the approach. This shift occurred for the team, but the comparisons with the other methods were significant in terms of being able to take a stand on how (and why) to conduct phenomenography. Further, the comparisons may well be of interest (and perhaps use) to others contemplating phenomenographic research.

Naturalistic inquiry

In her story, Pam Green writes from the perspective of a naturalistic inquirer and describes her adjustment to phenomenographic research as facilitated by the shared interpretive stance that is foundational to both research approaches. This stance is based on the notion of multiple legitimate interpretations or realities, in which ‘the transactions between the researcher and the research participants create understandings that are value-mediated or subjective’. (From a slightly different perspective, a phenomenographer would describe multiple realities in different terms, as existing in the relationship between the perceiver and the phenomenon perceived, between the researcher and the data analysed.)

Other similarities between naturalistic inquiry and phenomenography from Pam’s viewpoint included the common focus on inductive processes, where the research findings arise from a bottom up process grounded in verbatim texts, rather than a top down deductive process of testing a priori hypotheses. The use of purposive sampling (in the case
of phenomenography with the aim of maximising variation), rather than random or broadly representative sampling, was also a familiar notion from her research background. In addition, the common focus between developmental phenomenography and naturalistic inquiry on achieving practical outcomes from the research ensured an overlap for her in the underlying purposes of the research conducted.

Differences in terms of how rigour is established formed the greatest contrast between phenomenography and naturalistic inquiry. While naturalistic inquiry is quite prescriptive with respect to rigour (there is a substantial list of criteria for trustworthiness), particularly in the use of a range of data sources and methods of data collection, phenomenography typically uses one data source and only one round of data collection. This difference arises because in phenomenography the researchers’ data interpretations are made on a collective, not an individual interview basis. The aim is not to capture any particular individual’s understanding, but rather to capture the range of understandings across a particular group. In other words, the analysis goes across and between all of the interview transcripts so that the categories of description that are yielded reflect not individual meanings or conceptions, but rather conceptions from a pool of meanings. The interpretation is, thus, based on the interviews as a holistic group, not as a series of individual interviews.

The rigour of phenomenography, from Pam’s viewpoint, lies more in the iterative process by which interpretations are assiduously checked against the transcript data at each stage of the analysis process. Furthermore, during the processes of team analysis, agreement between researchers was reached through discussion and mutual critique of the data.

**Phenomenology**

For Robyn Barnacle, with her background in philosophical phenomenology, her introduction to phenomenography was facilitated by a shared focus on human meaning as the research object. As with Pam, the overlap in epistemological assumptions between phenomenography and her background research approach substantially eased the transition, in particular, the shared assumptions of meaning as a contested space, where there are multiple legitimate interpretations of a single text or set of transcripts. Any interpretation, however, needs to be justified or supported with evidence. The process of interpretation is viewed as seeking agreement between interpreter(s) and the text/transcripts, and interpretation or understanding is seen as a
process of mutual transformation, with the text as an active player. (In phenomenographic terms, much of this overlap in assumptions would be described in terms of interpretation occurring as a relationship between the transcripts and the researchers).

Robyn voiced a concern with the apparent ‘neatness’ of the categories of description, when individual experience is not neat. Nita Cherry echoed this concern. The apparent neatness of phenomenographic categories of description arises out of the research focus on constituting key aspects of collective experience, rather than the rich detail of individual experience. This leads to categories that are described in terms of critical qualitative similarities and differences, though there is room within these descriptions to include some of the non-critical variation in experience that is perceived.

**Action research**

Nita Cherry describes her action research ‘mindset’ as initially presenting a challenge in trying to come to terms with phenomenographic research. She was, at first, particularly concerned about the lack of involvement of the interviewees in the data analysis process. In her view, this lack of member checking seemed to ‘take the people out of the equation’ – the very antithesis of action research. Nita was concerned initially that this might mean that we ‘think of the categories [of description] as constructs in their own right’. This meant that Nita, in her journey into phenomenography, was particularly concerned about phenomenographic rigour, given that processes for establishing rigour in action research were not appropriate for phenomenography.

Previously in her research work, member checking (or taking the interpretation of an individual transcript back to the relevant research participant for checking) was used as a key aspect of rigour. Normally (that is, in her work as an action researcher), she would ‘have a concern about taking data away from the ‘owners’ of the data in any research method that involved detached ‘coding of other’s material’. However, given that phenomenography seeks meaning across individuals’ stories or examples of their experiences, that is, at a collective level rather than an individual level, member checking is not appropriate. Instead, phenomenographers rely on the iterative process of rigorous checking of interpretations against verbatim transcripts, continuing until the rate of change from each iteration reduces, and the description of categories stabilises.

Nita describes phenomenographic categories of description as ‘elegant constructions’, reflecting the sense of neatness to which Robyn also
referred. While Nita described this ‘elegance’ as intellectually satisfying to create, this elegance for her as an action researcher also carried connotations of an intellectual ‘word-smithing’ exercise divorced from the messiness of real-world experience. It seems clear that the phenomenographic focus on describing key aspects of collective experience, not the richness of individual experience, involves quite a conceptual leap in terms of the move from action research to phenomenography.

The area of greatest overlap between action research and phenomenography for Nita lay in the tracking of the analytic steps involved in arriving at the descriptions of each category. She commented: ‘I would now be a great deal more confident in a rigorous and transparent category system developed through phenomenography than more ambiguous methods that do not track how themes or constructs are constructed.’ This tracking helped address a major concern for her with the rigour of the process, that is, that the team constitution of the categories of description was inevitably a social process, and that the process of social negotiation should be adequately acknowledged.

Furthermore, Nita’s experience of phenomenography as a collaborative, team approach was consistent with her action research experience (even though the collaboration was with other researchers rather than with the interviewees), and aided the transition to phenomenography. However, this also means that if Nita had experienced phenomenography as a lone researcher, she perhaps would have found it even more difficult to come to terms with. In contrast, Robyn, with her background in philosophy, felt quite comfortable with a subsequent experience of individual phenomenographic research, following her team research experience.

Quantitative research

Gerlese Åkerlind, with her background in experimental psychology, describes the converse of Pam and Robyn’s experience of the transition to phenomenography being facilitated by shared epistemological assumptions. The objectivist stance of experimental psychology meant that the transition to a qualitative research paradigm was the greatest barrier to be overcome for her in undertaking phenomenographic research.

Gerlese’s previous interviewing experience formed as much of a barrier as an aid in undertaking phenomenographic interviewing. On the one hand, while others have emphasised the challenge of learning to avoid leading questions in a phenomenographic interview (outside
of the structured questions), this is common practice in psychological research, so Gerlese was already well experienced in this regard. However, when conducting interviews as part of psychological research it is typical to focus on what she terms ‘what’ questions (that is: ‘What did you do?’). This contrasts quite sharply with the focus in phenomenographic interviews on what she terms ‘why’ questions (that is: ‘Why did you do it that way?’). This difference required some practice for her to come to terms with.

Like John and Pam, Gerlese emphasised the importance of pilot interviewing in learning to conduct phenomenographic interviews. She also found that undertaking a phenomenographic analysis of written comments, before undertaking an interview-based analysis, provided a useful learning experience. Part of the transition process for her, in learning to do phenomenographic research, involved coming to terms with the large amount of qualitative data inherent in a set of verbatim interview transcripts. Practising first with a set of written comments enabled her to separate learning about the aims and processes of phenomenographic research from the extra complexities of working with the large data set typically involved.

In phenomenography, small sample sizes with maximum variation sampling, that is, the selection of a research sample with a wide range of variation across key indicators (such as age, gender, experience, discipline areas and so on), is traditional. For the other researchers, such purposeful or purposive sampling (the selection of research participants for a given purpose), and the selective sampling of a relatively small number of participants (commonly 20–30 in phenomenographic research) with the intention of gaining depth of meaning fitted well with the research practices of their methodological backgrounds. However, Gerlese, with her prior background in quantitative approaches, was more accustomed to large samples selected randomly.

This section has highlighted the perceptions of the researchers of the differences between phenomenography and other research methods with which they had been more familiar. The remainder of this chapter deals with specific aspects of phenomenographic research methods.

**Phenomenographic practice at the macro level: Common difficulties experienced by novices**

As would be apparent from the preceding section, researchers experienced in other research methods find some aspects of
phenomenographic research challenging and, at times, difficult to understand and reconcile with their previous research experience. Many of these difficulties are also shared by new researchers and in this section we discuss a number of aspects of phenomenographic research that novices (those new to the method irrespective of other experience) might find puzzling.

**Interviewing**

John highlights the nature of phenomenographic interviewing as a common difficulty for novices. Phenomenographic interviews are semi-structured, with the researcher clearly setting the interview topic(s) through the use of a number of set questions, but then making substantial use of unstructured follow up questions to further investigate interviewees’ responses to the questions. However, during the unstructured part of the interviews, it is important that the interviewer not raise any ideas around the topic(s) that have not previously been introduced by the interviewee. This contrasts with both more structured and less structured approaches to interviewing, as mentioned in Pam’s story, where she describes being used to having greater freedom to exchange comments during interviews.

Gerlese highlights the importance of distinguishing ‘what’ from ‘why’ questions as an interviewer, particularly when following up on participants’ responses to questions. When interview participants provide examples of their experience of the phenomenon being investigated (for instance: examples of going about learning, teaching, researching, or whatever the phenomenon of interest is), it is easy for novices to inappropriately focus their follow-up questions on eliciting more details about what the participant does when learning, teaching, researching, etc. A more useful follow-up question for eliciting participants’ intentional attitude towards the phenomenon (and thus the underlying meaning of the phenomenon for them) is to ask why they go about it in the way they describe. The aim of the latter type of question is to explore the interviewee’s underlying intentions or purposes with regard to the phenomenon. However, exploring the underlying meaning is the key issue and the formal inclusion of the word ‘why’ in the question is less important than actually finding out why the interviewee sees the phenomenon in a particular way.

To help newcomers to phenomenography become familiar with phenomenographic interviewing, John, Pam and Gerlese all strongly recommend pilot interviewing, both to provide an opportunity to
develop the required skills and also to refine the planned questions so that they do in fact catalyse appropriate discussion about the phenomenon.

**Understanding the phenomenographic focus on collective, not individual, experience**

Robyn describes part of this difficulty in her story:

> At times, I, and other members of the team, struggled to come to terms with this aspect of our approach. It is easy to confuse, or equate, the category with the individual. That is, mistakenly believe that a single category should equate fully with the perception of a single individual. While this may occur, if, for example, only one participant in the study expresses a certain perception, this is the exception. Instead, our aim, in using a phenomenographic approach, was to find similarities across the perceptions expressed by a group of individuals.

Phenomenographically, not only is a category of description not equivalent to the meaning expressed in any one transcript, but the meaning expressed in a transcript is not equivalent to the meaning experienced by the interview participant, due to the context sensitive nature of experience. The meaning of the phenomenon for any one participant may vary under different circumstances. Consequently, the interviews attempt to elicit the meaning of the phenomenon for each participant at the particular time and in the situation in which they were being interviewed. Thus, meaning may vary within individuals as well as between individuals, but the range of variation within individuals is likely to be encompassed by the range across individuals. The set of transcripts as a whole represents a snapshot of the ways of experiencing the phenomenon by a particular group of people at a particular time and in response to a particular situation.

The phenomenographic focus on participants as a collective group, rather than as a series of individuals, also makes member checking inappropriate as a validity check for phenomenographic analyses. As described by Robyn, any one transcript is unlikely to correspond precisely with any one category of description, because the categories have been constituted from an analysis of all of the transcripts, as a group. From this perspective, no one transcript can be understood in isolation and can only be interpreted in comparison with the rest of the group of transcripts.
Searching for both variation and commonality in the transcripts

John highlights the danger of novices focusing on variation in meaning more than commonality in meaning as they analyse the interview transcripts. This can lead to focusing on the multitude of surface details on which transcripts inevitably differ, in a way more appropriate to content analysis. By contrast, the phenomenographic aim is to constitute a smaller number of more holistic meanings, with a focus on key aspects of experience which link and separate the different categories of description.

Too strong a focus on details can also lead to going off on tangents during the analysis, where the primary purpose of the research project, in terms of the stated research question, is lost during the course of following interesting sidelines. (Perhaps this is illustrated in Robyn Barnacle’s story when she talks about her desire to explore the role of serendipity in the research phenomenon.)

Distinguishing between critical and non-critical variation in experience

Both Robyn and Nita expressed discomfort with the ‘neatness’ of the categories of description, when individual experience is not neat. The analysis focused on key aspects of the collective experience, expressed as categories of experience, rather than on the detail of individual experience. The descriptions of each category may, however, include some non-critical variation; that is, the category is focused around a particular common theme that distinguishes it from other categories, but with a range of different perspectives about some non-critical aspect.

Thus, a key challenge in the interpretive process is drawing a distinction between critical and non-critical variation in experience. ‘Critical’ variation is that which distinguishes one meaning or way of experiencing a phenomenon as qualitatively different from another. ‘Non-critical’ variation in meaning is that which occurs within a particular way of experiencing, and does not distinguish between different ways.

Phenomenographic practice at the micro level: Key practices in phenomenographic research

In the previous section, the issues discussed are mostly at the macro level. They represent characteristic principles that have to be under-
stood if a researcher is to undertake phenomenographic research. In this section, the discussion is more at the micro level of detailed method. For example, given the purposes of semi-structured interviews highlighted above, how do you actually go about conducting them and how did the different members of the research team experience those processes?

**Interviewing**

As phenomenographic research commonly uses interviews as the sole source of data, it is crucial that they are conducted rigorously. This is essential so that the researchers can feel confident about the integrity of the data with which they are working. It also provides vital external rigour, for those outside of the research process. John argues that being as faithful to the data as possible implies that it is not acceptable during the analysis to infer what else might have been said by participants if it was not actually said during the interview. (Sometimes, this argument is erroneously confused with the idea that one can only work with what is said explicitly and not the implicit implications underlying the words used – as illustrated in Robyn’s story, and discussed further later in this chapter.) Thus, if the analysis is to be comprehensive, it is essential that the interview process be undertaken properly in the first place. Pam also emphasised the importance of getting the interviews right in the first place in her story when she discussed the frustration of being unable to discern the meaning expressed in a particular transcript because the interviewee had not fully articulated their views.

**INTERVIEW PROCEDURES**

Appropriate phenomenographic interviewing procedures are discussed within the individual stories of John, Pam and Gerlese, and these stories should be re-read with this issue in mind. Pam conducted all of the interviews for the team research project, after she and John had jointly prepared the interview scenarios and prompts, and piloted the interviews. As a sole researcher, Gerlese engaged in the same set of processes for her research project. Issues of participant selection and how many to interview are mentioned above and discussed further by John in his story. The nature of the interview scenario, follow-up prompts, and other pointers for interviewing are raised within Pam’s story, while Gerlese emphasises ways of focusing the interview towards eliciting underlying meaning.
John’s story also refers to the need never to abuse the trust of those being interviewed. He describes their particular vulnerability in a process designed to get them to reveal everything possible about how they see a particular phenomenon. He suggests that almost always interviewees reveal something about themselves that they would never have expected they would. It is important that they feel comfortable at all times and that their willingness to cooperate is never abused. One public aspect of that is that no published material should ever contain information that could identify any interviewee. That is a matter of trust.

PILOTING

While pilot interviews are strongly recommended for novices to provide an opportunity to develop the required skills, they also represent an important aspect of any phenomenographic study — in order to check that the scenario or question(s) set actually do yield information on the intended phenomenon in a way that is useful to the study at hand. As Pam noted in her story, the trialing of interview questions was vital for fine-tuning the interview scenario. The consequent addition to the interview questions of the phrase ‘in some way’, for instance, enabled participants to be freed from trying to label their experiences of success in research as being totally ‘successful’ or totally ‘unsuccessful’ to something a little more blurred and thus more realistic.

MAINTAINING FOCUS ON THE CENTRAL RESEARCH QUESTION

As described by Robyn:

The role of the [research] question in determining what is read, or what counts as legitimate reading, was particularly apparent in this study. Our reading of the transcripts was explicitly directed from the beginning by the key question: How do researchers perceive of the conditions for successful research? This meant that much of what participants spoke about was treated as, and remained, peripheral in what we searched for, and focused on, that is, explicit references to success.

In John’s story he argues that the need for constant focus on the research question or phenomenon applies throughout all aspects of the phenomenographic research process. He points out that, in the analysis, you need to read transcripts from a particular perspective.
If the interviewee is saying this about X or Y, what must the phenomenon mean to them? If they are now saying Z as well, does this change my analysis of how they are seeing the phenomenon? It is possible to become distracted and to begin to analyse the various ways of seeing X, for instance. There is a greater danger of losing focus in this way if you cut the interview transcripts into pieces and re-assemble the similar pieces from different transcripts together… Anchoring all analysis to the transcripts alone, reading forward and backwards around key statements and using the devil’s advocate process, all assist in maintaining focus during the analysis.

Any one interview may address a number of phenomena, both deliberately and accidentally. Consequently, the same transcript can be read searching for ways of experiencing different phenomena. However, this would occur across different readings, with only one phenomenon in focus for any one reading. Otherwise, it is possible to accidentally attempt to incorporate descriptions of related phenomena from the transcripts into your interpretation of the primary phenomenon, in effect contaminating the data. As Robyn describes:

During the project, my attention was drawn to a particular aspect of the interpretive process. This was the seduction of the text, or the way in which certain terms or ideas snag, or attach themselves, to certain readers. For me, it was the references in some of the texts to serendipity and luck as factors leading to successful research outcomes… To what extent does success in research [the primary research question] depend on a researcher being able to do this? What is the relationship between successful research and serendipity?

…We debated the issue during the analysis process, but the team felt that there was not enough evidence within the relevant transcripts to suggest that serendipity is a condition for success, rather than just an element of the research process.

**Faithfulness to the transcripts**

John has made the argument that the only evidence that can be used in analysis is what is in the transcripts, and that the categories of description constituted by the researcher(s) should always be justifiable in terms of the transcripts. Pam reinforces that by providing examples of the kinds of questions the research team asked of itself
throughout the analysis, questions such as: Where did this term come from? What is missing? What else might this mean? What does this not mean? What is different about this category?

Pam explains these processes as being based on faithfulness to the transcript data and she gives an illustration in her story of an episode in which recourse to the transcripts prevented an error in analysis. For a time the team misled itself in describing one category in a particular way only to find that a re-reading of the transcripts showed the error and a new description was developed that was more faithful to the transcripts. Pam saw that experience as underlining the rigour of the phenomenographic research process:

We re-read and re-read data so much that our heads swam with complexities. The critical component of such team processes meant that we argued the point, left no stone unturned, grew weary with semantic play, discarded ungrounded descriptions and shared in the joy and satisfaction when forms of clarity finally emerged.

Focus on meaning vs words in the transcripts

In her story, Robyn expressed concern over ‘staying with the transcripts’ because she understood this to mean checking all interpretations against the words in the transcripts. Her concern was that what people say is not necessarily the same as what they mean.

It is clear to me that the directive to ‘stay with the transcript’ is the key strength of the phenomenographic approach. But this directive also has limitations. It is limiting, for example, in that the analysis is restricted to that within the text which is made explicit. Elements of shadow, or aspects of phenomenal experience that are ambiguous and undefined, are not attended to. This is how I felt about the fact that serendipity was not explored within our analysis.

In fact, phenomenographers make a very explicit attempt to distinguish purely linguistic differences in what interviewees have said from differences that represent variation in underlying meaning (such as Bowden, 1994b; 2000b; Prosser, 1994; 2000; Svennson, 1997; Trigwell, 1994; 2000). This includes the aim of looking as much as possible beyond the particular words chosen by the interviewee to
their underlying intentional attitude towards the phenomenon they are describing.

This aim is what underlies the common phenomenographic practice of asking interviewees to provide concrete examples to illustrate their comments. While Robyn was concerned that conceptual interpretations were based on descriptions of examples provided by participants, it is through these examples of actual interaction with the phenomenon being investigated that interviewees’ experiences of the phenomenon can best be interpreted. As Gerlese says:

> In phenomenographic interviews, we are trying to elicit underlying meanings and intentional attitudes towards the phenomenon being investigated. Typically, we do this through exploring concrete examples of the phenomenon provided by the interviewee. However, we are not interested in the details of the example per se, but in using them as a medium for exploring the way in which the interviewee is thinking about or experiencing the phenomenon, that is, those aspects of the phenomenon that they show awareness of. This entails going beyond ‘what’ questions (‘What did you do?’) to ‘why’ questions (‘Why did you do it that way?’).

Robyn draws an important distinction between meaning and what is said, and it is just that distinction that is at the heart of John’s use of the term ‘staying with the transcript’ in phenomenographic analysis. The emphasis when reading transcripts is to discern the fundamental meaning of the phenomenon as expressed in the transcript. The categories of description should be faithful to the meanings in the transcripts. This may mean using in the categories particular words from the transcripts, but that would be a consequence and not a goal of searching for meaning.

**Interpretive rigour in the iterative process of analysis – establishing validity**

Once the interviews have been conducted and then transcribed verbatim, the researcher(s) need to read and re-read the interview transcripts. This leads to a series of iterative cycles between the transcript data, researcher interpretations of the data, and checking of interpretations back against the data. It is in the conduct of this iterative process that phenomenographers most clearly establish their interpretive rigour. Also, if there are any occasions within a transcript in
which it seems that the interviewer has led the interviewee, then the remainder of that section of the transcript should be excluded from the data analysis.

TEAM ANALYSIS

Nita indicates that a characteristic which stands out for her in team phenomenographic research is the rigorousness of the requirement that several people subject the categories they are developing to sustained cycles of scrutiny, debate and testing against the data. Similarly, Pam says:

The power of working within a team of academics who were committed to constructive but critical debate was undoubtedly the greatest strength of this approach. I am convinced that by working as a team we engaged in robust and at times potentially confronting discussion that moved us to develop rigorously considered categories. For instance, Nita Cherry tended to raise questions that ensured that we were explicit about why we were doing what we were doing under the umbrella of phenomenography. While this slowed the process at times, it did mean that we were all clear about the purpose behind our actions as well as the implications for rigour and for other methods. John Bowden frequently played, without orchestration, but more through his own natural inclination, the role of devil’s advocate. This meant that we were frequently challenged to justify the ways in which we framed the categories. We had to go back to the data constantly and check for accuracy.

Robyn saw the iterative process, the use of members of the research team as devil’s advocates for each other, and the constant re-reading of transcripts as category descriptions developed, as forming important sources of the rigour she came to see in phenomenographic analysis. The reliability of phenomenographic analysis is ensured through the researcher engaging in an iterative dialogue with the text and not predicting outcomes in advance by imposing categories of description.’ Robyn was a little surprised by the time and number of iterations required, but felt that slightly changed meanings in the transcripts emerged on each re-reading, especially when different transcripts were compared.

However, Robyn also emphasised that teams can be self-affirming, going down tangential paths together. To help avoid such groupthink, John highlighted the importance of maintaining clear roles within the
team, where someone is explicitly nominated to play a devil's advocacy role on each occasion. Nevertheless, Nita maintained that any team research would be inevitably more socially constructed than grounded in the data, no matter how rigorous the analytical procedures employed.

Both Pam and John refer in their stories to the issue of trust between team members in the team analysis approach. Pam refers to the need for agreement about ways of giving critical feedback so that members feel confident to put forward new ideas without fear of personal attack. John highlights the vulnerability one can feel, especially in the beginning – he makes the distinction between intellectually understanding the devil’s advocacy process and emotionally handling it. Explicit talk about such things at the beginning is essential in Pam's view, but these issues also need to be returned to throughout the analysis. John reiterates that point and emphasises the need for everyone to be sensitive to their colleagues’ feelings, without undermining the process by backing off from being critical. Getting on with the analysis while also taking care of each other’s feelings seems to summarise the position taken by both Pam and John.

INDIVIDUAL ANALYSIS

While individual analysis may seem inevitably less rigorous than team analysis, Gerlese highlighted ways of acting as your own devil's advocate during individual analysis. This included taking substantial breaks during the analysis so that you could return to it with an open mind. Such breaks normally occurred spontaneously, due to the pressures of other work, but should be engineered if they do not happen naturally. Also, when checking tentative outcomes against the transcript data, it is important to consciously and actively look for negative examples as well as supporting examples. Further, there are ways of gathering feedback from others on your preliminary outcomes, even when ostensibly working alone, which she describes in her story.

VALIDITY

Of course all of the above discussion of rigour is directly linked to the issue of establishing validity. John argues that the validity of the outcomes is related to the processes that are used at all stages of phenomenographic research. He emphasises the importance of a common focus through the planning phase (including working out whom to interview), data collection (in this case interviewing), analysis and interpretation. He argues strongly for consistent interview
frameworks with no leading questions except for the planned interventions, team analysis with continual devil’s advocacy, delay in labelling categories and in looking at relations between categories until analysis of the categories is complete, using the transcripts as the only source of evidence in the analysis, and focusing on meaning by taking the ‘whole of transcript’ approach. He has argued (especially in Bowden, 2000a) that you can mount a consistent argument for all of these choices and that if you carry them out, you maximise the validity of the outcomes. He sees the validity in qualitative research lying largely in transparent processes that can be argued for within a coherent framework. Many examples of the rationale behind each of the process aspects are to be found in his story in chapter 2.

Pam contrasts phenomenography with naturalistic inquiry in terms of the different criteria used to establish interpretive rigour. In her story she describes those differences in detail. Naturalistic inquiry emphasises triangulation of data sources, checking data interpretations with original interviewees, and prolonged engagement in the field, with emerging hypotheses leading to further inquiries through follow up interviews. By contrast, phenomenography typically uses one data source and one round of data collection only, and member checking is generally not regarded as an appropriate validity check for this research approach.

Nita suggests in her story that:

Rigour in the process demands complete transparency as to the way the categories are framed and reframed, and recognition and articulation of the logic and social processes that drove the framing. That deconstruction is a significant piece of work in its own right, requiring not only transparency or reporting but critical awareness of the generating processes. Again, it struck me that this sort of critical awareness or critical consciousness fits very well with the action research perspective.

In team analysis, agreement between researchers is reached through discussion and mutual critique of the data and of each researcher’s interpretive hypotheses. This provides a way of balancing the perspective of a single researcher and providing a check on the possibility of prejudiced subjectivity. For a single phenomenographic researcher working alone, rigour lies more in the critical attitude taken towards the interpretations made, that is, how individual presuppositions have been made. Further rigour pertains to the checks and balances that the researcher has employed to help counteract the impact
of one’s particular perspective on the outcomes of the analysis. The documentation of the interpretive steps, by fully detailing the steps and presenting examples that illustrate them (Sandberg, 1996; 1997; Kvale, 1996; Åkerlind, 2002), is part of this process. Seeking feedback on one’s interpretations from other researchers and from representatives of the population interviewed are also common procedures.

**Phenomenographic practice at the micro level: Key questions that each researcher needs to answer for themselves**

As with any research approach, there are key areas of variation in practice as well as commonalities in practice among phenomenographic researchers. That means that, in addition to coming to understand and implement the agreed core research practices of phenomenography, each researcher also needs to make a reasoned decision on areas of variation in practice.

**When should the analysis begin?**

John Bowden argues strongly that the phenomenographic analysis should not begin until all of the interview transcripts are ready for reading as a whole. This view is based on the need to keep the nature of the interview constant throughout all interviews. He suggests that if analysis occurs in parallel with further interviewing there is a danger that the interviews will be altered in some way, either explicitly or unconsciously. This may relate to his focus on developmental phenomenography – he is as interested in what the categories are used for after the analysis as he is in development of the categories themselves. So he is concerned with the credibility of the outcomes in the eyes of non-phenomenographers. Reproducibility is important to him so that he can clearly claim that the categories of description finally developed are related to a particular set of inputs across all of the interviews.

Gerlese argues for the practical value of analysing a preliminary subset of interview transcripts before analysing the whole. This provides a means of making the sheer amount of data collected more manageable. As described in her story, this is a common concern among phenomenographic researchers, often faced with some 20–30 transcripts of 20–30 pages in length.

An alternative and more effective answer might be to combine the two approaches, that is, to conduct all of the interviews before any analysis has begun, but then to do a preliminary analysis on a subset of the transcripts, followed by an analysis of all of the transcripts.
Should the analysis be of the whole transcript or pooled excerpts?

While all of the contributors to this monograph have taken a whole-of-transcript approach to data analysis, Gerlese commented in her story about the varying practices that exist in the amount of each transcript considered at one time during phenomenographic data analysis. Practice varies from considering the whole transcript or large chunks of each transcript related to a particular issue, to the selection of smaller excerpts or quotes seen as representing particular meanings, although these are still initially interpreted within the larger interview context (Marton, 1986a; Svennson and Theman, 1983). In the latter approach, the smaller chunks are separated from the transcript and combined for analysis in one decontextualised ‘pool of meanings’ (Marton, 1986b).

The argument for the ‘whole of transcript’ approach is that the larger interview context is needed to help the researcher(s) faithfully interpret the meaning of particular comments in the transcript. The argument for the ‘pool of meanings’ approach is that greater decontextualisation facilitates the phenomenographic focus on collective meanings, and away from individual meanings.

Should phenomenographic analysis be done individually, or as a team?

The relative confidence one could feel in the outcomes of individual versus team phenomenographic analysis was explicitly discussed by the contributors to this monograph during the team meeting mentioned at the start of this chapter. Although the decision on this issue is often made as much on practical grounds of feasible research resources as on logical argument for the ‘best’ approach.

Both Gerlese, with her extensive experience of individual analysis, and Robyn, who was currently involved in conducting primarily individual analyses arising out of the team project, expressed confidence as individual researchers. One of the advantages they experienced as individual researchers was becoming more deeply familiar with the data than was possible for each member of a team research project. Furthermore, as an individual researcher, one can still seek feedback from others during the analysis, and both had done so. Nevertheless, because such feedback would be more regular and systematic with team research, it seems likely that team research might be more rigorous, in general.

Robyn suggests that:
The process of team analysis highlights the crisis involved in all interpretation: the challenge of realising difference, bringing it to the fore. But within a team situation there is also more because it is not just agreement between text and interpreter that is sought but agreement between text and multiple interpreters. In order to reach consensus one must convince others of the credibility of one’s interpretation. A case needs to be made and defended. …

… I found the process of team analysis provided a rare opportunity to gain insight into, as well as critique, my own interpretive habits. This may have made the analysis process longer than it might otherwise have been, but it was certainly more rewarding.

While team analysis may seem inevitably more rigorous than individual analysis, Gerlese highlights in her story that ‘the large number of existing phenomenographic doctoral theses indicates that high quality phenomenographic research can be accomplished as an individual researcher working largely on one’s own’. She also points out that ‘any outcome space is inevitably partial, with respect to the hypothetically complete range of ways of experiencing a phenomenon’. Consequently, the contrast between different phenomenographic approaches should not be considered in terms of what is most likely to produce right or wrong outcomes, but rather, more or less complete outcomes.

Thus, an individual researcher can, at the least, make a substantial contribution to our understanding of a phenomenon, even if team research might have taken that understanding further.

When have you reached an appropriate endpoint in your analysis?

During the team meeting, all contributors explicitly discussed what gave each of them a sense of confidence in their research outcomes. In fact, there were varying views among different members of the research team as to the confidence they felt in the validity of the final research outcomes that they shared. Some felt that they had been through a rigorous analytic process of iterating to the point of exhaustion, while others suggested that the pressure to complete inevitably meant that not enough time had been put into the iterations. However, all agreed that this stage of the analysis was a very time
consuming one, and that there was no way to predict in advance how much time it would take.

However, researchers must always be wary of closing the analysis too early. As John points out in his story, reading the transcripts again can bring new insights:

Something that I have observed every time I have undertaken a phenomenographic analysis remains a source of joy and a motivation to continue with repeated analyses of interview transcripts. The observation is that as you proceed with the analysis, each new reading of the transcript brings new insights. What you see when you read something is affected as much by what you bring to the task as it is by what is in the transcript you are reading. I am sure literary scholars and others would be able to explain that in more enlightened ways than I can, but the observation leads me to conclude that going through five, ten or fifteen versions of the categories of description is necessary. I do not believe you could read the transcripts once and then write the final categories of description. After the first reading, with the draft categories written, you read the transcripts again with different eyes. The draft categories guide your reading, but you are looking for evidence to undermine that draft representation – to test the coherence of a category description or to question the difference between two different draft categories. Aspects of the transcripts that were not seen as significant before become significant now. You see them afresh. Despite having read the transcripts many times, it seems like you are seeing that aspect for the first time. Every new reading of the transcripts is a new experience. Perhaps when it ceases to be like that is when you’ve reached the last version.

So, how do you manage simultaneously not to exhaust yourself and not to distort the research outcomes? Pam and Nita referred back to their previous research experience as a source of guidance on when enough iteration is enough. In naturalistic inquiry, closure of the analysis is reached when data saturation or redundancy is reached. In phenomenography, Pam felt that the same principle applied. Analytic closure was reached when nothing new emerged from continuing iterations. Nita explained that in action research, closure is reached very quickly, whenever the researcher felt confident to act. But this was a state of preliminary closure, with one cycle of data, analysis and action leading on to another cycle. Furthermore, this confidence and
subsequent action was reached in collaboration with the parties concerned, re-emphasising the significance raised earlier of collaboration in generating a sense of confidence for Nita.

John emphasised a sense of maximising difference, as when one tentative category is confidently separated into two. An illustration of this was the separation between Categories A and B that emerged in the shift from version 3 of the outcome space for the research success project to version 4 (see Appendix A). John described this as ‘seeing a separation’, that is, seeing a separation between two qualitatively different ways of experiencing, when previously they were seen to be the same. Gerlese described an ‘ah ha’ response; the point at which confusion becomes clarity. Pam also emphasised a sense of being able to see a difference, and the feeling that ‘this is a difference that’s real and not simply something that we want to happen’. The example Pam provided in her story of when the team experienced great confidence in generating a particular category description, only to conclude in the next iteration that the description was not in accord with the transcript data, provides a negative illustration of her point. It illustrates a time in the analysis when she felt a particular outcome was reached more because the team wanted it to happen than because it was ‘real’, and how the iterative process subsequently exposed this.

**Structural relationships between categories of description**

Reports of phenomenographic research studies often contain three aspects. The first is a description of the ‘outcome space’, that is, a picture (in either prose or graphic form or both) of the categories and their relation to each other. The second aspect is the detailed elaboration of the categories of description. The third aspect is the detailed analysis of the relationships among categories. In practice, most outcome spaces show some form of hierarchical relationships among categories.

The nature of the hierarchical relationships commonly posited between categories of description in phenomenographic research is frequently misunderstood. The hierarchy is not one based on value judgements of better and worse ways of understanding, but on evidence of some categories being inclusive of others. Thus, the structural relationships searched for in a phenomenographic outcome space are ones of hierarchical inclusiveness. This does not mean that a linear hierarchical structure (such as in Gerlese’s outcome space in Chapter 10) need emerge; forks and branches in the hierarchical structure (such as in the other contributors’ outcome space in Chapter 9) are also common.
Novices might also question whether it is necessary to posit structural relationships at all. However, the very notion of an ‘outcome space’, rather than simply ‘outcomes’, implies structure, and Gerlese describes in her story why searching for structure as well as meaning derives from the theoretical underpinnings of phenomenography:

Phenomenography espouses a non-dualistic ontology: ‘There is not a real world ‘out there‘ and a subjective world ‘in here‘. The world [as experienced] is not constructed by the learner, nor is it imposed upon her; it is constituted as an internal relation between them’ (Marton & Booth, 1997 p.13). The phenomenographic proposition that ways of experiencing represent a relationship between the experiencer and the phenomenon being experienced, leads to the expectation that different ways of experiencing will be logically related through the common phenomenon being experienced. Thus, a core premise of phenomenography is the assumption that different categories of description or ways of experiencing a phenomenon are logically related to one another (Marton & Booth, 1997). These relationships are commonly expected to form a structural hierarchy of inclusiveness, with some ways of experiencing being more complex than others, but including aspects of awareness constituted in less complex ways of experiencing. The inclusive nature of such relationships makes some ways of experiencing not only more complex, but more complete, than others.

Nevertheless, Nita expressed a philosophical objection to searching for structure, as she felt it would inevitably encourage the imposition of researcher views on the data. Pam indicated that, while the team did not search for structure as part of the team research process (due to these different views in the team), she and John constituted structural relationships post hoc. She found this process unexpectedly useful:

The question arose as to how we would think and make informed decisions about the relationality between categories. There was no inherent hierarchy evident here. My inclination then was to assume that there were no relationships between the categories. However, the use of conceptual mapping revealed that there was a two-pronged relationship among the categories. This was taken to the team for discussion and we redrew the concept tree until we felt that we had exhausted
possibilities sufficiently and that the mapping worked for the data. This was a new experience to me. Had I worked alone, I would have omitted this crucial aspect of the work. 

Robyn felt that structuring inevitably occurs, at least implicitly, in the constitution of the categories, otherwise there would be a different category for each transcript. From her perspective, the analysis reached a point where structure became apparent, even though the team had not explicitly constituted the categories with structure in mind.

**The dialectical relationship between meaning and structure**

In phenomenography, it is not enough to simply constitute categories of description; one must also consider the structural relationships between the different categories. This is typically regarded as representing the dialectical relationship between the content, or referential aspect, of meaning (represented by the categories of description) and the structure, or structural aspect, of meaning (represented by a mapping of the inclusive relationships between categories).

While John and Gerlese both see structure as critical in a phenomenographic outcome space, they differed in their approach to constituting these relationships. In her story, Gerlese made clear her view that meaning and structure should be co-constituted contemporaneously, using a mixture of logical and empirical evidence. She argued that this was because structure formed part of the theoretical basis of the phenomenographic approach, part of what made the approach unique, and part of what made the approach useful. John reaffirmed that structure was important, but argued that it should be constituted only after the categories of description are finalised, based on a logical analysis of the empirical outcomes. His argument is that waiting until the end of the analysis to search for structure reduces the chances of researcher bias. Thus, while both emphasised the importance of structure, it seems that from John’s perspective, structure acts to extend the meaning of the categories of description, whereas from Gerlese’s perspective, structure is an integral part of the meaning of the categories of description.

**HOW BEST TO INTEGRATE STRUCTURE AND MEANING IN THE OUTCOME SPACE**

As Gerlese noted in her story, there is considerable variation in practice in terms of the extent to which structural relationships are constituted on
logical versus empirical grounds. John argues strongly for the final structure of a phenomenographic outcome space being justified entirely empirically, solely in terms of the content of transcripts. On this basis, any suggestion of inclusiveness must be evident in the transcripts. Gerlese emphasises an iterative process of alternating between logic and data in constituting structure. She felt that active empirical support for the proposed structure was the ideal to aim for, but allowed in principle for the possibility of some researchers proposing a structure that could not be demonstrated in the data, as long as it was not contradicted by the data. (It would also be essential for the researcher to make the absence of active empirical support clear when describing the proposed structure.)

At one level, however, the variation in practice is primarily one of timing. All researchers start their analysis with the data. Neither categories of description nor structural relationships are anticipated in advance of the data. So, a secondary question is, how quickly do you move from considering only the transcript data in forming categories of description to also considering their logical structure as part of the constitution of categories? The ‘success’ project, following John’s lead, took the extreme position of completing the categories of description before looking for the relationships among them. Others include a search for structural relationships from the beginning of the analysis. John argues that the latter approach has the danger of categories of description being derived as much from the researcher’s ideas as from the data.

Gerlese describes in her story how she has taken a position that is different from both of the extremes. She discusses this issue at length and readers should go to her story to examine her arguments.

What did we all get out of the experience?

What makes phenomenography worth using in one’s research?

Different researchers highlight different benefits from using phenomenography. Nita valued the focus on balancing commonality and difference in experience of the same phenomenon. She also valued the elegant descriptions of experience that emerged. With more detailed descriptions, such elegance is impossible – indeed, thick descriptions deliberately challenge the notion of simplicity and elegance. Gerlese valued the ability phenomenography provides to look at variation in experience holistically, through the simultaneous focus on constituting the range of ways of experiencing a phenomenon and structural relationships between the different ways of experiencing. John
emphasised phenomenography’s distinctive focus on the collective experience, rather than individual experience, although individual experience can subsequently be mapped against and explicated by collective experience. Pam highlighted phenomenography’s basis in variation theory, while Robyn highlighted the method of analysis, which ensured that:

The transcript remains an active partner in the interpretive process and that the phenomenal descriptions that are elicited emerge out of this two-way, or dialogical, exchange. This is both challenging and stimulating for the researcher, in that one at times struggles to stay with the data but also benefits from gaining genuine insights.

All saw potential for using phenomenography again in the future, with Nita’s comment representing the general response:

I quickly found that I really liked the ‘polishing’ of the categories which finally emerged – this involved constant iteration between text and categories, considering different perspectives, and finding ways to both focus and enrich the central ideas if done in a highly transparent and iterative way, in the company of others and with rigorous testing against the data. The careful and methodical working and re-working of the data that this method invites is very powerful. I particularly valued these dimensions of rigour, given my disquiet with processes which invite individual researchers to simply immerse themselves in their data and see what themes emerge intuitively, without a subsequent process of critique and review by others. As a result, I would encourage anyone working with complex data, from a large number of people, to seriously consider using some of the protocols of phenomenography.

**Looking back**

Learning ‘to do’ phenomenography is not a journey with a distinctive beginning and a clear end. While this section has presented five different perspectives of experience in working with phenomenography, it does not purport nor even support the idea that any of us might be at the end of our individual (and/or collective) learnings about phenomenography. However, we have learnt much from our collaborations. We have argued the point and tried to reconcile our differences. Where we
have not agreed, we have presented our various positions and the accompanying reasoning. In other words, these accounts represent just a snapshot or reflection upon our uniquely positioned experiences. We have deliberately tried to present not a white-washed, consensus-based view of phenomenography, but rather a multiplicity of complex views that have been individually influenced by various methodological experiences and preferences. In so doing we have attempted to allow the reader ‘in’ with respect to phenomenographic matters at both macro and micro levels. Thus, what remains now is for readers to consider whether or not to embark on a phenomenographic journey and, if so, to raise their own methodological questions, make appropriate and informed choices, and reflect upon their own phenomenographic journey(s). The following chapters contribute further to this consideration and the decision-making that might flow. The next chapter outlines comprehensively the methodological details pertaining to a phenomenographic study on understandings of growth and development as an academic, as well as understandings of being an academic. As such it provides further food for thought for the reader who might then move to the outcome chapters relating to both of the studies noted in this chapter and the preceding stories.