

Reflections on trustworthiness in phenomenographic research: Recognising purpose, context and change in the process of research

Brandon I. Collier-Reed

University of Cape Town, South Africa

Åke Ingerman,

University of Gothenburg, Sweden

Anders Berglund

Uppsala University, Sweden

Abstract

In interpretive research, trustworthiness has developed to become an important alternative for measuring the value of research and its effects, as well as leading the way of providing for rigour in the research process. The article develops the argument that trustworthiness plays an important role in not only effecting change in a research project's original setting, but also that trustworthy research contributes toward building a body of knowledge that can play an important role in societal change. An essential aspect in the development of this trustworthiness is its relationship to context. To deal with the multiplicity of meanings of context, we distinguish between contexts at different levels of the research project: the domains of the researcher, the collective, and the individual participant. Furthermore, we argue that depending on the primary purpose associated with the collective learning potential, critical potential, or performative potential of phenomenographic research, developing trustworthiness may take different forms and is related to aspects of pedagogical legitimacy, social legitimacy, and epistemological legitimacy. Trustworthiness in phenomenographic research is further analysed by distinguishing between the internal horizon – the constitution of trustworthiness as it takes place within the research project – and the external horizon, which points to the impact of the phenomenographic project in the world mediated by trustworthiness.

Key words: Phenomenography, context, trustworthiness, credibility, transferability

Introduction

It can be argued that one of the primary values inherent in research is a search for knowledge – descriptive or predictive – that is powerful in human culture and enterprise. Furthermore, the character of scientific knowledge as open and contestable makes it an important contributor to the democratic form of life. We would argue that educational research should be located within this premise. In the context of South Africa today, this implies contributing to transformation through, of, and within higher education as well as the educational system in general; supporting the establishment of proto-democratic practices (see for example Booth, 2008). In this article, we consider what it would entail to contribute towards the emergence of such practices through the enactment of higher education research projects; with a particular focus on the qualitative research approach known as phenomenography (Marton, 1981; Marton & Booth, 1997). In support of such a contribution, we will argue for

the alignment of the phenomenographic research project with its underpinning purposes and potentials, through the explication of trustworthiness.

Phenomenography as a research tradition is located broadly within an interpretive epistemological orientation and focuses on the variation in how a phenomenon is experienced by a group of individuals. Phenomenography is underpinned by, amongst others, a focus on the relational nature of human experience, a non-dualistic ontological perspective, an explicit focus on the experience of phenomena, and the adoption of a second-order perspective. The research outcome is a set of categories that describe the qualitatively different ways of experiencing that phenomenon, and are logically related in structure and meaning. The categories are not descriptive of how *individuals* perceive the phenomenon – rather they describe the phenomenon at the collective level. Learning, in the phenomenographic tradition, points to coming to discern phenomena in new and more powerful ways.

Conventional notions of validity and reliability in research typically rely on being able to measure in an absolute sense; treating the research object, and the project as a whole, as a closed system. Qualitative researchers, particularly those adopting an interpretive epistemology, typically rather operate in an “*open system*” (Lincoln & Guba, 1985, p.329) and often draw on the work of Lincoln and Guba (1985) for understanding and discussing the trustworthiness of their research outcomes. An important aspect of Lincoln and Guba’s argument is that “...trustworthiness is a matter of concern to the *consumer* of inquirer reports” (ibid, p. 328). As such, the research outcome is defined (and restricted) by structured reporting, irrespective of the format in which it is published. In this article, we will suggest how to make use of trustworthiness as something that is more than simply applicable for the benefit of “consumers”. Rather, we explore trustworthiness as a means for a holistic *strengthening* of the research outcome and its impact. We will consider this with respect to its purpose and inherent potential in situations where it may be of value.

The traditional potential of phenomenographic research is that of *collective learning*. It entails building knowledge, potentially accumulative, which can be a factor contributing to desirable societal change. Embracing a broader perspective of what the research outcome entails – in addition to “inquirer reports” – we emphasise a troika of learning: the learning of the researcher, the learning of the college of teachers, and the learning of students. As such, research outcomes are an integral part of articulating and shaping change. Recognising these additional research outcomes as important to take into consideration, and as legitimate parts of research endeavour, two new *potentials* of phenomenographic research become apparent: the critical potential and the performative potential (Booth, 2008). The *critical* potential of phenomenography concerns “bringing problematic features of the teaching and learning nexus to the fore” (ibid). It includes the “emancipation from a distorted set of practices that take learning for granted” (ibid) and aims to bring about change in the specific context from which the object of research originates. The *performative* potential of phenomenography “stresses an action research framework for phenomenographic research for effectuating change” (ibid) and thus aims to align and integrate the process of change into the research process itself. Research transforms from building knowledge about the world to being transformative in itself. A simple illustration of this is the change that is brought about in the world by the very formulation of the descriptors that are core to the research process. We argue that these three potentials of phenomenographic research (critical, performative, and collective learning) should be emphasised as relevant to the transformation of higher education.

This article developed from the experience of three phenomenographic studies in higher education science and technology contexts (Berglund, 2005; Collier-Reed, 2006; Ingerman,

2002). The purpose of our article is not to provide a full description of how phenomenographic research can or should be conducted. Rather, we address a recurrent issue in phenomenographic research where results are often communicated to non-educationalists who typically base their understanding of rigour on how well a study reports on issues of validity and reliability. We show how trustworthiness is a vital part of establishing rigour in the research process and that it applies to the entire process of the phenomenographic research endeavour. Furthermore, we suggest that it is important to consider trustworthiness not only in relation to the methodology and framework required but also in relation to the *context* within which the research is conducted, analysed and disseminated. Finally, the article develops the argument that trustworthiness plays an important role in not only effecting change in a research project's original setting, but also that trustworthy research contributes toward building a body of knowledge that can play an important role in societal change.

Towards trustworthiness in phenomenographic research

In many traditional research approaches, validity, reliability, and generalisability are taken as the philosophical constructs that demarcates good research from non-science and have reached the status of a 'scientific holy trinity' (Kvale, 1996, p.229). However, in some research approaches (in particular those with an interpretive epistemology) trustworthiness has developed to become an important alternative for measuring the value of research and its effects, as well as leading the way by providing for rigour in the research process (see for example Morse, Barrett, Mayan, Olson, & Spiers, 2002; Rolfe, 2006). Making use of trustworthiness as an alternative construct rather than adapting the meanings of validity and reliability makes explicit some basic differences in the assumptions of how the world is constituted.

Lincoln and Guba (1985), in their seminal work *Naturalistic Inquiry*, argue that rigour can be appropriately reported on in interpretive studies. Their concern at the time was that interpretive studies were seen as 'undisciplined ... [with] "sloppy" research, [and researchers] engaging in "merely subjective" observations' (p.289). This book was written at a time when the positivist approach to research was dominant in the social sciences and the merits of interpretive research was judged based on characteristics that Lincoln and Guba felt were only appropriate to the former approach. Even though their argument may appear dated, Morse et al. (2002) argue that "their work on 'trustworthiness' is still regarded as seminal and pertinent" (p.16). The core of Lincoln and Guba's argument was that it is not appropriate to argue for the positivist standards of validity, reliability and objectivity in measuring the value of interpretive research, but rather that the concept of the trustworthiness of the investigation should be employed. To argue for trustworthiness, they brought in the notions of 'credibility', 'transferability', and 'dependability' as equivalent to the traditional research notions of internal validity, external validity and reliability respectively. As part of arguing for the 'trustworthiness of results' (or in a later refinement (Lincoln & Guba, 1986, p.78), the 'authenticity of results'), rather than a focus on validity and reliability of results as the appropriate measure of establishing rigour in research, it has been possible to judge the value of interpretive studies by a different set of appropriate criteria.

Credibility can be considered the "truth value" (Lincoln & Guba, 1985, p.294) of an investigation. It considers if there is "correspondence between the way the respondents actually perceive social constructs and the way the researcher portrays their viewpoints" (Mertens & McLaughlin, 2004, p.106). Transferability draws on the notion of "applicability" (Lincoln & Guba, 1985, p.298) of research outcomes. Lincoln and Guba argue that "the original enquirer cannot know the sites to which transferability might be sought, but the

appliers can and do” (ibid). It is thus up to the researcher to “provide sufficient detail to enable the reader to make ... a judgement ... to determine the degree of similarity between the study site and the receiving context” (Mertens & McLaughlin, 2004, p.107). Dependability is underpinned by the idea of “consistency” (Lincoln & Guba, 1985, p.298) of research findings and is employed to “attest to the quality and appropriateness of the inquiry process” (Mertens & McLaughlin, 2004, p.107). These constructs will be developed below.

Validity in a study is the extent to which it is possible to show that a study measures what it sets out to measure. By drawing on the argument made by Lincoln and Guba (1985), it can be shown that there are other approaches that may be more appropriate as measures of this aspect of rigour in interpretive research. Cohen et al. (2000) agree with this position, and argue that it is important to discuss issues of validity in the context of the research perspective employed in a study. Thus, for interpretive research to be considered sound, it should not be a requirement to strive to meet the traditional criteria for validity, i.e. the need to ‘demonstrate concurrent, predictive, convergent, criterion-related, internal and external validity’ (Cohen et al., 2000, p.106). This notwithstanding, many phenomenographic studies have argued for ‘validity’ in their results in terms of the sorting of the descriptions into *already* constituted categories (see for example Johansson, Marton, & Svensson, 1985). This can be seen as a form of generating trustworthiness. Even so, there are still those who would argue for a more traditionally orientated approach to reporting validity in phenomenographic studies. One of the stronger proponents of this post-positivist type approach is Cope (2000). His attempt to introduce ‘rigour’ through application of the traditional approach to validity does not, however, adequately address the fact that a phenomenographic result is an interpretive (and thus by definition not objective) process describing individuals’ experience of a phenomenon and not the phenomenon itself.

With respect to the notion of credibility, Booth (1992) argues that it is understandable for the research community to scrutinise the results and conclusions of an interpretive study with specific regard to their ‘credibility and trustworthiness’ (p.65). She highlights three approaches to ensuring that this measure of rigour occurs and although she frames these using the term ‘validity’, it is clear from the description of each that they bear little resemblance to the traditional notion of validity described earlier. The approaches she highlights are termed ‘content-related’ (see also Cohen et al., 2000, p.109), ‘methodological’, and ‘communicative’ (see also Kvale, 1996, p.244) validity. Content-related validity concerns the researcher’s familiarity with the subject matter under investigation, methodological validity looks at how the goals of the study match its design and execution, and communicative validity involves the researcher’s ability to argue their interpretation of the data. In applying these constructs to the present argument we have chosen to use the term ‘credibility’ rather than ‘validity’ as in, for example, *content-related credibility* because of the associated legacy that terminology such as ‘validity’ brings to interpretive studies. We would argue that it is more appropriate to re-characterise the term ‘credibility’, as first proposed by Lincoln and Guba, instead of trying to take what is essentially a positivist concept and refashion it to work in an interpretive context.

The need for ‘reliability’ in an interpretive study has been questioned by some scholars (see for example Lincoln & Guba, 1985; Stenbacka, 2001). Lincoln & Guba claim that there can be no validity without reliability (see also Cohen et al., 2000) and thus, in the context of using their terminology to describe these constructs, no there can be no credibility without dependability. The consequence of this claim is that in a study, if credibility has been established, the need to establish dependability to show that a study is trustworthy becomes unnecessary. However, they do acknowledge that in practice there is still the need to ‘deal with dependability directly’ (Lincoln & Guba, 1985, p.317).

Dependability in a phenomenographic study has been likened by Booth (1992) to a journey of exploration. She equates the planning of a study to the planning of an expedition and the execution of a study similar to 'following an outline chart while simultaneously ensuring the potentially interesting ways are followed and striking features are noted' (p.67). The results of the study are analogous to a description of the terrain in terms of what the explorer saw and experienced along the way. A different explorer may have the same goals for their expedition, but would most likely follow a different path to achieve these goals and would thus give a different description of what they saw and experienced along the way. However, should the second researcher be given the former's maps and charts, log-books, diaries, and itinerary, it is likely that the resultant description of the terrain covered during the exploration would be similar to that presented by the first explorer. Booth (1992) does add a rider to this expectation of similarity of description. She maintains that this does 'presuppose that both explorers (or researchers, now), have similarly thorough experiences of what it is to explore foreign lands (or research other people's conceptions)' (p.68). Similarly, in terms of a phenomenographic study, researchers should have a similar understanding of the phenomena focussed on as the object of study. It is clear from this analogy that dependability relates to 'consistency in data [interpretation]' (Åkerlind, 2005, p.331) and thus consistency in research findings and should as far as possible be actively pursued in a phenomenographic study. Kvale (1996) identifies three areas in an interpretive study that can influence the dependability of research findings. The first area involves the interviewer during the data collection exercise, the second is relevant during the transcription of the interviews and finally, during analysis, dependability issues are again raised.

The relationship between context and trustworthiness

An essential aspect in the development of the trustworthiness of phenomenographic research is its relationship to context(s). This is important in the collection of empirical data, how the analysis process is conducted, and the implementation of the research outcomes. To deal with the multiplicity of meanings of context, we see the need to distinguish between context in different levels, or rather domains, of the research project.

The domain of the researcher

The domain of the researcher includes the context of the research study as such, the purpose of the research study, and the motivation for the researcher's endeavour. A central element in the phenomenographic researcher's context is their own awareness of the role they are playing in the study and the preparations that were made for the study's setup. This entails the delimitation of the object of research, the empirical data collection schema, and the use of the primary research outcome – the set of categories – in relationship to the context of the study and the potential for change with respect to the triad of learning as defined earlier.

The domain of the collective

The domain of the collective is a tool within a phenomenographic analysis. The primary outcome of a phenomenographic analysis is a set of categories that reflect, on a collective level, the possible ways interviewees have of experiencing a phenomenon. The individual categories are given meaning in the context of the full set of categories and the full extent of the extracts from the interviews. The meanings of individual extracts are seen within the collective of other extracts, each conveying part of the possible ways of experiencing the phenomenon.

The domain of the individual participant

When a participant experiences something or talks about a phenomenon, for example during a phenomenographic interview, some aspects of the phenomenon come into focus, while

others remain in the background (see for example Booth, 1992; Gurwitsch, 1964). In this instance, the individual context is then the theme of awareness, lending meaning to what is in focal awareness. Important from a phenomenographic perspective is that this individual context is *experienced*. This contrasts with what we would call the *prepared* context of the interview. This prepared context is how a researcher attempts to frame the situation in such a way that there is a shared experience amongst the participants; rich in the areas of the researchers' interest. For example, this could be as simple as phrasing questions in a form that is recognisable to a certain kind of situation for the participants, or including photographs of objects of interest in the interview (Collier-Reed, 2006).

Other research approaches (e.g. phenomenology: cf. Ashworth & Lucas, 2000, on entering the lifeworld of individual students) are often more concerned than phenomenography with context as a part of the research outcome, for example, by taking the full history of individuals into account. This yields quite different research outcomes, given different research purposes.

Trustworthiness serving the purpose of research

It is important to reflect on *why* it is important to consider trustworthiness, *who* trustworthiness is directed at, and *what* the essential characteristics of trustworthiness are. We can begin to address these questions by recognising that developing trustworthiness is essential in building relationships between the object of study (and the situation it is in), the context of the researcher, and research purpose and outcome (including its tentative impact). These relationships are at the heart of the development of trustworthiness and can partially be broken down into developing links associated with particular sub-parts of the research project as a whole: e.g. the way in which interviews are conducted, transcriptions are made, and categories constituted on the basis of transcriptions.

Consideration of the *why* question is what makes trustworthiness serve the purpose of the research. It is essential to develop trustworthiness to be able to use the research as a lever for ensuring any of the three kinds of potentials we identified in the introduction. Depending on what *primary purpose* is tied to these potentials, developing trustworthiness may take different forms:

Critical potential: The research outcome in this sense acts as a basis for effecting change in the original setting. It is thus important that the research builds a 'true' reflection of the original setting. This would emphasise aspects such as the relationship between study aims, the data collection, and the delimitation of the phenomenon in such a way that it is central in the original setting.

Performative potential: The act of researching something is transformative in itself and the development of trustworthiness thus also forms part of this transformation. Several participants are involved in the process of transformation – researchers, colleagues, and students – each of whom may perceive the research as trustworthy for different reasons. As an example, simply the fact that Berglund (2005) undertook the research that he did created an awareness of the variation in the students' understanding both within the Department of Computer Science and within the Nordic computer science education community. It is therefore important to recognise the impact of the research and the actions of the researcher on the situation itself, with due regard given to the legitimacy of these acts in the eyes of *all* the participants.

Collective learning potential: This involves contributing to a base for change together with learning in a general sense. Trustworthiness is thus important when it comes to the appropriateness of how the collective learning from those contributions has come about. It

concerns measures of credibility, dependability, and transferability, to ensure rigour in the research.

These three potentials may be related to aspects of pedagogical legitimacy, social legitimacy and epistemological legitimacy. The critical potential warrants pedagogical legitimacy, the performative potential warrants pedagogical and social legitimacy, and the collective learning potential warrants epistemological legitimacy (but in its implications also pedagogical legitimacy).

We will further our analysis of trustworthiness by distinguishing between the internal horizon and the external horizon of trustworthiness. By *internal horizon*, we mean the constitution of trustworthiness as it takes place within the phenomenographic research project – in essence, the trustworthiness of the process of reaching the outcome space. It concerns how the parts of the project are constituted in terms of trustworthiness, and how these parts are related to each other, giving shape to the trustworthiness of the phenomenographic project as a whole. The *external horizon* of trustworthiness points to the impact of the phenomenographic project in the world mediated by trustworthiness. Here trustworthiness is constituted internally through the parts, and externally in the context of the purposes of the project as a whole. The external horizon draws the boundary that delimits the extent and relevance of the project and its outcomes.

The internal horizon of trustworthiness

Within the internal horizon, the research project in itself is taken for granted. It is an environment in which it is possible to define ‘good practice’, making use of the analytical and theoretical tools of the phenomenographic approach. Of concern is the alignment of research question to empirical setup, methods and forms of data collection, stability of and transparency in analysis procedures and structure and documentation of research results.

Credibility

The notion of credibility in a phenomenographic study is something that must be considered throughout the study. Contrary to the view of those arguing for the return of the traditional condition of validity, the idea of credibility is not something that is only left to those interrogating the research findings. Rather it is a way of designing the study that begins with the definition of the object of research and follows through each aspect of the study to its conclusion. Furthermore, it includes a tacit relationship with the community outside the study through an interaction around the research findings as they are constituted. The measures of credibility presented below, namely content-related credibility, credibility of method, and communicative credibility, ensure that the outcome of phenomenographic research can be taken seriously.

Content-related credibility

Content-related credibility relates to a researcher having a comprehensive grasp, or understanding, of topics related to the phenomenon under investigation. The research community could justifiably question the rigour of the results should a researcher not be completely familiar with the subject matter. An interesting example where this was particularly important in the choice of interviewees is Ingerman (2002), which undertook a study of how physicists expounded on their research. This relied on interviewing fellow researchers in the subfield mesoscopic physics, where he had been involved in research himself. Booth (1992) maintains that it is important that this knowledge of the subject matter in question be an *open* understanding as the ‘researcher has to be open for ways of understanding it [the subject matter] which differ from those generally accepted’ (p.66). We would argue that the need for an ‘open understanding’ of the topic of investigation is more

important than being a leading expert in this topic because, in a phenomenographic analysis, the focus is on others' ways of experiencing a phenomenon, which may turn out not to be related at all to the scientific (or other) ways of understanding it. It is important that during analysis, a researcher is able to 'bracket', or carefully examine and then suspend one's own understanding of the topic, so that the categories of description constituted from the data are not influenced by the researcher's bias in understanding the phenomenon. Ashworth and Lucas (2000) argue that the act of bracketing allows a researcher to set aside, as far as is possible, their own assumptions about a phenomenon 'in order to register the student's own point of view' (p.297).

Credibility of method

Booth (1992) has argued that credibility of method (or what she terms methodological validity) 'lies in the match between the goals of the study and its design and execution' (p.66). To this end, issues relating to sample composition, the context in which the interview takes place, the structure and content of the interview, and the analysis of the data are centrally important.

In a phenomenographic study, it is crucial that the sample selected for the study be appropriate and relevant to the central research question under investigation. The result of a phenomenographic investigation is the variation in the ways a phenomenon is experienced and without careful consideration given to the selection of the sample, a justifiable critique from the research community could be the possibility that the full extent of the variation in ways of experiencing the phenomenon is not assured.

The second construct ensuring credibility of method lies in the context of the interview. It has been argued that the 'mutual definition of the situation determines the value of the discourse' (Booth, 1992, p.66). We argued earlier that context is a necessary element of a phenomenographic study that should be kept in mind during implementation of the research approach. By ensuring a shared experience of a phenomenon, there is a greater likelihood of 'establishing a joint definition of what is being talked about' in the interview situation (Säljö, 1996, p.23-24). This will directly influence the credibility of the results of the subsequent analysis. Berglund's (2005) research is an example of how a shared context was constituted. In his interviews, he discussed certain computer science concepts against the background of a project that the students were in the process of completing.

In a phenomenographic interview, questions are not necessarily as structured as often found in other forms of interpretive studies. The typical phenomenographic interview is of a semi-structured nature with only a few key questions predetermined. This introduces the importance of the third construct ensuring credibility of method and that is to ensure that the structure and content of the interviews are richly reported in a study so that the research community can determine the completeness of the data collected. During the interview conversation itself, the researcher must ensure that the possible power relationship that could exist between themselves and the interviewee does not limit the 'open and deep discourse' (Booth, 1992, p.66) required to obtain rich, meaningful data from which credible categories of description can be constituted during analysis.

It is important to ensure credibility of method through the adoption of a particular attitude towards analysis. Ensuring rigour through analysis 'lies in the researcher's open and thorough attitude, eschewing preconceived ideas and being receptive for the meaning that interviews themselves reveal' (Booth, 1992, p.66). The actual process of phenomenographic analysis varies from researcher to researcher. However, what is common to all approaches to analysis is that credibility of this aspect of a study rests on the 'search for meaning' in the data

collected and during this process ‘retaining a sense of the whole while pursuing the particular’ (p.66).

Communicative credibility

Communicative credibility relates to the requirement of the researcher to be able to ‘argue persuasively for the particular interpretation that they have proposed’ (Åkerlind, 2005, p.330). It is also the ability to present the results and conclusions of a study to the research community in an open way that allows the study as a whole to be scrutinised. This allows the research community to recognise and judge for themselves the credibility and legitimacy of the researcher’s interpretation of the results (Booth, 1992; Kvale, 1996). Unlike in a positivist-oriented study, here researchers do not argue for results as being the ultimate truth. In a phenomenographic study, the aim of a researcher is to be able to defend their interpretation of the data to the ‘outside world’. This ‘outside world’ has a number of distinctive participative ‘actors’ that between them define an external and an internal aspect of communicative credibility (Booth, 1992).

External communicative credibility relates to other researchers with similar interests being able to recognise the legitimacy of the interpretation made of the data. There are a number of avenues for this to take place, both during the course of the analysis of a study as well as after a study is complete. These include conference presentations and seminars where interim constitutions of categories of description can be presented for discussion, as well as peer-reviewed journal articles (Åkerlind, 2005). Internal communicative credibility relates to the participants of the study, and others indirectly involved in the study (Booth, 1992; Kvale, 1996). Even though Francis (1996) suggests that it is important to ensure the ‘accuracy [of the interpretation] through the eyes of the individual concerned’ (p.41) – and thus make use of internal communicative credibility – the focus on having an *individual* recognise the legitimacy of the constituted categories of description is inherently problematic in a phenomenographic study and runs counter to its methodological underpinning. In a phenomenographic analysis, individuals contribute fragments of an interview conversation to a ‘pool of meaning’ after which categories of description are constituted. Individuals will generally not be able to recognise ‘their’ contribution to the outcome space. The categories of description do not capture their ways of experiencing the phenomenon, but rather the experience of the phenomenon by all those in the study. A further issue raised in this regard by Åkerlind (2005) is that one of the fundamental underlying assumptions of the phenomenographic approach is that an individual’s experience of phenomenon is ‘context sensitive’ (p.331) and can change, even during the course of an interview, depending what aspect of the phenomenon is in focus at that particular moment. Thus, at the time the participants are asked about their interpretation of the final categories of description as constituted, there is no guarantee that they would still be experiencing the phenomenon in the way they were during the interview.

Dependability

Earlier in this article we argued that *dependability* is a more appropriate construct to use for the positivist term *reliability*; and is an important part of a phenomenographic study. Ensuring the dependability of a study is important as it allows for consistency of data interpretation and thus consistency in the research findings of an investigation. To ensure dependability, care must be taken during the interview conversation, during transcription of the data, and most importantly, during constitution of the categories of description.

Dependability as a function of the interview conversation

During data collection, the structure of the interview conversation and the issue of prompting the interviewee with leading questions is of critical importance (Francis, 1996, p.38; Kvale, 1996, p.157). The researcher must be conscious at all times of ensuring that the interviewees are expressing how *they* have experienced the phenomenon in question. Leading questions can influence the interviewee to attempt to 'see their experience through the eyes of the interviewer rather than through their own' (Francis, 1996, p.38). In a phenomenographic interview, only a few key questions are predetermined. This places an extra burden on the researcher to ensure that non-leading questions are used during the interview conversation. Questioning strategies that develop based on what is brought into focus by the interviewee need to be carefully considered while the interview itself is still in progress. Francis (1996) argues that 'prompt trails' (p.39), based on a whole range of possible interviewee responses, could be set up before the interview to help ensure that the interview is well grounded in the interviewee's experience of the phenomenon in question.

Dependability as a function of accuracy of transcription

One of the first decisions required of a researcher after an interview is complete is to decide the degree of accuracy of the transcript. A decision of this nature is important because, as Kvale (1996, p.160) points out, the act of transcription is an 'interpretive process' in itself. Since a phenomenographic analysis does not have the same focus on linguistic elements as a method such as discourse analysis, it is not necessary to record every tonal inflection or pause in speech. What is important, however, is that the 'spoken word' is transcribed as accurately as possible.

Dependability as a function of analysis

Within the process of analysis of phenomenographic data, intersubjective agreement forms the basis for assuring dependability of results. Åkerlind (2005) has identified two possible types of intersubjective confirmation of categories of description in the context of analysis of phenomenographic data. One she labels the *dialogic reliability check*, where 'agreement between researchers is reached through discussion and mutual critique of the data and of each researcher's interpretive hypothesis' (p.331). The second she labels the *coder reliability check*, where two or more researchers 'independently code all or a sample of interview transcripts and compare categorisations' (p.331). In keeping with the use of the construct 'dependability' in place of 'reliability', Åkerlind's two types of intersubjective confirmation of categories of description are termed here a *dialogic dependability check* and a *coder dependability check*.

Bowden (2000), Prosser (2000) and Trigwell (2000) are proponents of various forms of the *dialogic dependability check*. As an example, Bowden's approach to analysis is to assign one member of his research team the responsibility of constituting draft categories of description, then having the other researchers re-read the transcripts, and independently make tentative allocations of each transcript to the categories. The allocations are compared and 'where there were disagreements about categories of description or allocation of transcripts, they were resolved with reference to the transcripts as the only evidence of students' understandings' (p.52). Bowden claims that by making use of a group iterative approach to the analysis of phenomenographic data, new insights to the constitution of the categories of description can be achieved. He is also of the opinion that by carrying out a solitary analysis of the data, he would not have been able to constitute as accurate a set of categories of description as was possible by the group (p.59).

A strict implementation of the *coder dependability check* during the development of the categories of description (where researchers independently code the interview transcripts and compare categorisations) has been argued by Marton (1986) to be unreasonable to expect in phenomenography as ‘the original finding of the categories of description is a form of discovery, and discoveries do not have to be replicable’ (p.35 – see also Säljö, 1988). However, once the categories have been found, Marton (1986) argues that ‘it must be possible to reach a high degree of intersubjective agreement concerning their presence or absence if other researchers are to be able to use them’ (p.35). This idea of categories being recognised by others forms the basis for a practice widely practised in phenomenography (Cope, Horan, & Garner, 1997; Johansson et al., 1985; Marton, 1996; Säljö, 1988), often referred to by the traditionally positivist constructs of inter-rater reliability or interjudge reliability.

There are varying levels to which this latter check can be implemented. Some researchers (Cope, 2004; Johansson et al., 1985; Säljö, 1988) have argued for a statistic to be generated (note the positivist associations!) illustrating the agreement between how co-researchers have classified data in the categories of description identified by the original researcher. The appropriate level of agreement, after consultation between researchers, has been variously argued to be between 75% and 100% (Johansson et al., 1985, p.251) or between 80% and 90% (Säljö, 1988, p.45). Some researchers (for example Booth, 1992) have been critical of the use of intersubjective agreement of this kind, primarily because there is a difficulty in finding co-judges who are as well versed in the subject matter under investigation.

Sandberg (1997) has criticised the use of the *coder dependability check* on a number of fronts. Firstly, he argues that because a statement made by an individual could appear to a co-judge to encompass more than one conception, there is potential conflict as to where the statement should be located. In reality, it is possible that this statement (or more accurately, aspects of this statement) has helped constitute more than a single category of description. Another consideration is that a statement may not in itself be easily identified with a category of description, but it may have contributed a critical aspect that has helped constitute a particular category. He has argued that this ‘integrated interpretation can be difficult for co-judges to see, since they do not have the same familiarity with the data as the original researcher’ (p.206). Sandberg’s (1997) second criticism relates to the extreme form of coder dependability where a particular percentage correlation is demanded between co-judges. He argues that searching for such correlation is not appropriate in a research approach that does not have an objectivist epistemology. He claims that a ‘serious consequence of such theoretical and methodological inconsistency in the research process’ (p.208) is that the scientific community is able to call into question the results of a phenomenographic study. In order to overcome these concerns, he has argued for interpretive awareness to be the means whereby trustworthiness is ensured in phenomenographic studies. To ‘maintain’ this interpretive awareness means to ‘acknowledge and explicitly deal with our subjectivity throughout the research process instead of overlooking it’ (p.209).

The external horizon of trustworthiness and the grounds for transferability

There comes a point where you reach the *boundary of applicability* of the internal horizon of trustworthiness. The categories of description that have emerged from that empirical data may have been rigorously developed and thus trustworthy, but at the boundary, what others see of the results are a set of idealised and – from their standpoint – a not necessarily ‘real’ set of categories. Direct application to a broader context becomes limited and at this point, the strength of the argument supporting the original research work often gets lost. That does not mean that it is impossible, nor potentially fruitful, to cross such boundaries. However, the

strength of the original outcome is reduced to that of the insight of others who need to take the results forward in application. Results may in such cases be reformulated through a process of open empirical and analytical reconstitution. The implication may be that the research must be repeated in this new context, but there is an overall epistemological gain during this process in terms of a growing body of knowledge and understanding and the constitution of outcomes on a different level. The empirical and theoretical boundaries of such collated research have to be critically considered. For example, the kinds of generalisations posited in *Learning and Awareness* (Marton & Booth, 1997) were based on a collation of 25 years of teaching and learning research; making use of a evolving phenomenographic framework.

Our claim is that part of the trustworthiness of phenomenographic research is developed through consideration of where the boundaries are and in what respect it is *appropriate* to cross them while still retaining the relevance of the research outcome and staying true to the methodological underpinning of the phenomenographic approach. In particular, we would critique phenomenographic work that has not been sufficiently analysed beyond the phenomenological notion of a ‘rich description’. We would argue that it is valuable, where possible, to analyse the essential aspects of the origin, context, and structure of a research situation so that it is possible for similarities and differences to be seen in relation to other situations where the results are potentially of relevance. For example, this may include providing analytical tools with which another context can be analysed to be more thoroughly compared with the original setting. In so doing, the framework for the transferability of results is explicitly articulated. This is not typically considered part of the phenomenographic research endeavour.

Directing focus on the external horizon of trustworthiness implies making clear what space could, and should, be made for the undertaking of a phenomenographic study. This clearly has associated implications for the internal horizon where there may be certain elements that are required to deem it to be good phenomenographic practice in terms of the level of importance of certain parts, as part of a bigger whole. The meaning of the phenomenographic project, and its parts, are thus not only judged as part of phenomenographic research practice, but also as part of a larger academic enterprise.

Concluding remarks

By returning to the three kinds of potentials for phenomenographic research we identified earlier – critical, performative, and collective learning – it is clear that developing trustworthiness plays an important role for not only effecting change in the research’s original setting, but also for contributing toward building a body of knowledge that can play an important role in societal change. As a consequence, we suggest that embarking on a research project simply for the sake of doing research is seldom justifiable – particularly in the South African context given the situation in higher education, the racial legacies brought about by apartheid, and the current pedagogical framings. We believe that a researcher is morally obliged to act on his/her results and in so doing keep true to what it means to be a researcher in South Africa. By implication, this may mean taking results further, both within and external to the phenomenographic tradition. In other situations, where the social context is not as pertinent, other elements may compel a researcher to act – for example, a passion to support student learning. The primary outcome to consider is the researcher’s learning, which encompasses the formal results, the experience of the original settings, and the imperative to undertake the research. This will inevitably take the results further than the formal endpoint of phenomenography.

In this article, we have argued that trustworthiness in phenomenographic research is what makes the research have impact in terms of being able to effect change in the original research setting, the transformation of those participating in the research, as well as having the potential to contribute to a broader knowledge base that can help shape change in society. In the ideal world, the strength of this trustworthiness is scrutinised in both its parts and its whole, internally and externally, regulating its impact on the broader world. Trustworthiness is the amorphous fluid that forms the shape giving relevance to the structure of an outcome, and guides the ways in which it may interact with the broader pedagogical, social and epistemological contexts of the world.

The way to ensure rigour in phenomenographic research is to adopt the principles of trustworthiness described in this article. However, it is not enough to only focus on the internal horizon of trustworthiness. It is also important to be mindful of the pedagogical, social, and epistemological boundaries when undertaking research – and thus focus on the external horizon – to ensure that any results obtained can meaningfully contribute toward impacting on societal change.

Acknowledgements

We wish to acknowledge funding from Swedish Research Links (VR-SIDA-NRF) as part of the Phenomenographic Horizons collaboration as well as the Swedish Research Council (VR-UVK). We are particularly grateful to Shirley Booth for the comments she made on an early draft of this article. Some of the work drawn on for this article was developed as part of Brandon Collier-Reed's PhD project, where the contributions by Jenni Case and Cedric Linder are acknowledged. The reading group of the Centre for Research in Engineering Education (CREE) has been the location for many stimulating engagements around methodological issues such as those discussed in this article.

References

- Åkerlind, G. S. (2005). Variation and commonality in phenomenographic research methods. *Higher Education Research and Development*, 24(4), 321-334.
- Ashworth, P., & Lucas, U. (2000). Achieving empathy and engagement: A practical approach to the design, conduct and reporting of phenomenographic research. *Studies in Higher Education*, 25(3), 295-308.
- Berglund, A. (2005). *Learning computer systems in a distributed project course: The what, why, how and where* (Vol. 62): Acta Universitatis Upsaliensis.
- Booth, S. (1992). *Learning to program: A phenomenographic perspective*. (Vol. 89): University of Gothenburg.
- Booth, S. (2008). *Transformation and phenomenography: Strange bed-fellows? Reflections on the potential significance of phenomenography and variation theory on research on transformation and higher education in South Africa*. Keynote presented at the European Association for Research on Learning and Instruction SIG 9: Implications of Phenomenography and Variation Theory in Practice, Kristianstad, Sweden
- Bowden, J. A. (2000). Experience of phenomenographic research: A personal account. In J. A. Bowden & E. Walsh (Eds.), *Phenomenography* (pp. 47-61). Melbourne: RMIT University Press.
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). London ; New York: RoutledgeFalmer.
- Collier-Reed, B. I. (2006). Pupils' experiences of technology: Exploring dimensions of technological literacy. Unpublished PhD thesis. University of Cape Town.

- Cope, C. (2000). *Educationally critical aspects of the experience of learning about the concept of an information system*. Unpublished PhD thesis, La Trobe University, Bundoora, Victoria.
- Cope, C. (2004). Ensuring validity and reliability in phenomenographic research using the analytical framework of a structure of awareness. *Qualitative Research Journal*, 4(2), 5-18.
- Cope, C., Horan, P., & Garner, M. (1997). Conceptions of an information system and their use in teaching about IS [Electronic Version]. *Informing Science Journal*, 9-22.
- Francis, H. (1996). Advancing phenomenography - Questions of method. In G. Dall'Alba & B. Hasselgren (Eds.), *Reflections on phenomenography - Toward a methodology?* (Vol. 109, pp. 35-47): University of Gothenburg.
- Gurwitsch, A. (1964). *Field of consciousness*. Pittsburgh,: Duquesne University Press.
- Ingerman, Å. (2002). *Exploring two facets of physics. Coherent current transport in superconducting structures. Phenomenographic studies of sense-making in physics.*, Chalmers University of Technology, Gothenburg.
- Johansson, B., Marton, F., & Svensson, L. (1985). An approach to describing learning as a change between qualitatively different conceptions. In L. Pines & T. West (Eds.), *Cognitive structure and conceptual change* (pp. 233-257). New York: Academic Press.
- Kvale, S. (1996). *Interviews: An introduction to qualitative research interviewing*: Sage Publications.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*: Sage Publications.
- Lincoln, Y., & Guba, E. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. In D. D. E. Williams (Ed.), *Naturalistic evaluation* (pp. 73-84). San Francisco: Jossey-Bass.
- Marton, F. (1981). Phenomenography - Describing conceptions of the world around us. *Instructional Science*, 10(2), 177-200.
- Marton, F. (1986). Phenomenography - A research approach to investigating different understandings of reality. *Journal of Thought*, 21, 28-49.
- Marton, F. (1996). Cognosco ergo sum - Reflections on reflections. In G. Dall'Alba & B. Hasselgren (Eds.), *Reflections on phenomenography - Toward a methodology?* (Vol. 109, pp. 163-187): University of Gothenburg.
- Marton, F., & Booth, S. (1997). *Learning and awareness*. New Jersey: Lawrence Erlbaum Associates.
- Mertens, D. M., & McLaughlin, J. A. (2004). *Research and evaluation methods in special education*. Thousand Oaks, Calif.: Corwin Press.
- Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International Journal of Qualitative Methods*, 1(2), 1-19.
- Prosser, M. (2000). Using phenomenographic research methodology in the context of research in teaching and learning. In J. A. Bowden & E. Walsh (Eds.), *Phenomenography* (pp. 34-46). Melbourne: RMIT University Press.
- Rolfe, G. (2006). Validity, trustworthiness and rigour: quality and the idea of qualitative research. *Journal of Advanced Nursing*, 53(3), 304-310.
- Säljö, R. (1988). Learning in educational settings: Methods of inquiry. In P. Ramsden (Ed.), *Improving learning: New perspectives* (pp. 33-48). London: Kogan Page.
- Säljö, R. (1996). Minding action - Conceiving of the world versus participating in cultural practices. In G. Dall'alba & B. Hasselgren (Eds.), *Reflections on phenomenography* (pp. 19-33). Gothenburg, Sweden.

- Sandberg, J. (1997). Are phenomenographic results reliable? *Higher Education Research & Development*, 16(2), 203-212.
- Stenbacka, C. (2001). Qualitative research requires quality concepts of its own. *Management decision*, 39(7), 551-555.
- Trigwell, K. (2000). A phenomenographic interview on phenomenography. In J. A. Bowden & E. Walsh (Eds.), *Phenomenography* (pp. 62-82). Melbourne: RMIT University Press.