THE PROBLEM OF DETECTING GENUINE PHENOMENA AMID A SEA OF NOISY DATA

Christine Knipping, Eva Müller-Hill

University of Bremen, University of Cologne

In our paper we consider two principles that are essential for research, and in particular for qualitative comparative work in mathematics education. First, we consider the explanation of phenomena rather than data as fundamentally important in educational research. Second, we discuss a holistic conception of human knowledge, suggested by Sellars, emphasising the role of common discourse for any theoretical scientific discourse. We relate our discussion to a number of comparative studies in mathematics education, as well as to conceptual modelling, a trans-disciplinary methodological framework that follows and reflects an interactive process in the transition from data to phenomena.

Keywords: Comparative Studies, Philosophy, Methodology, Data, Phenomena

INTRODUCTION

Comparative research in mathematics education, especially qualitative comparative research, is relatively new (see ZDM 2002 Vol. 34,6). One reason for this probably is the origin of mathematics education in psychology. The first comparative studies were quantitative, as for example the First International Mathematics Study (FIMS) in 1964 (see Husén 1967) and compared results rather than processes. As the field of mathematics education has embraced more sociological and anthropological approaches, more qualitative comparative research has been done, but compared to psychological methods in mathematics education it is often less accepted and understood. Kaiser (2002) for example states in the introduction to the special ZDM issue on *Comparative Studies in Mathematics Education*: “it seems that the same questions have repeatedly been asked in large-scale studies, and that qualitative strategies are still not well considered …” (p. 240).

In this paper we seek to provide a clearer foundation for qualitative comparative research methods in order to increase their acceptance in the wider mathematics education community. We focus on two principles of research: the emphasis on phenomena rather than data, and incorporating stakeholder views into research. Woodward (1989) discusses the relationship between data and phenomena in quantitative work in the natural sciences, and while we acknowledge that qualitative work in the social sciences is normally more closely focused on phenomena, a more conscious attention to phenomena could increase the recognition of qualitative comparative research methods. The second aspect is the recognition of the perspectives of the stakeholders, e.g., students, future teachers, in-service teachers, and administrators. Qualitative research is more likely than quantitative research to take such perspectives into account, but the relationship between these perspectives and results arrived at by scientific analyses is rarely described clearly. By drawing on

Sellars' idea of a "stereoscopic" view we provide a way to describe this relationship, which we feel would make qualitative comparative research methods more compelling.

Compared to qualitative comparative research methods in social science and anthropology, comparative research is less theorised in mathematics education. Anthropology and the social sciences developed a strong theoretical basis for their methodology through their longer history. The challenge is to explain the value of qualitative comparative research methods to mathematics education without having this long history to draw on. The methods can be copied, but the theoretical basis cannot be, because its origins are different. Instead of recapitulating the hundred year history of comparative methods in anthropology, we suggest in this paper to go more directly from principles derived from the philosophy of science. We consider especially emphasis on phenomena rather than data, and incorporating stakeholder views into research.

Throughout we will draw on comparative research studies in mathematics education to provide examples of the two principles in practice. These include Knipping (2003), a qualitative small scale comparison of proof teaching in France and Germany; the Learner's Perspective Study (Clarke, Keitel & Shimizu, 2006; Clarke, Emanuelsson, Jablonka, and Mok, 2006); and the TIMSS Video Study (Stigler & Hiebert, 1999).

THE DATA-PHENOMENA DISTINCTION

The data-phenomena distinction was posed and discussed for the general case of empirical science by Woodward and Bogen in the 1980s. Since then, it has shaped debates in general theory of science.

According to Woodward and Bogen, scientific theories are not about data. In particular, it is not data that is explained by mature scientific theories, but phenomena (Woodward & Bogen, 1988, p. 305).

> Underlying the distinction between data and phenomenon is the idea that the sophisticated investigator does not proceed by attempting to explain his data, which typically will reflect the presence of a great deal of noise. Rather, the sophisticated investigator first subjects his data to a great deal of analysis and processing, or alters his experimental design or detection technique, all in an effort to separate out the phenomenon of interest from extraneous background factors. [...] Figuring out what one should even try to explain—what the phenomena are in a given domain of inquiry—and what is mere noise is, as we shall see, an important aspect of scientific investigation, especially in relatively immature areas of inquiry like the social sciences. (Woodward, 1989, p. 397)

Data are characterized as the results of idiosyncratic, local measurement processes:

> As a rough approximation, data are what registers on a measurement or recording device in a form which is accessible to the human perceptual system, and to public inspection.

> […] They typically are of no theoretical interest except insofar as they constitute evidence for the existence of phenomena (Woodward, 1989, p. 393 f.).

In contrast, phenomena are:

> relatively stable and general features of the world which are potential objects of explanation and prediction by general theory (Woodward, 1989, p. 393 f.).

Though the examination of data serves both as a source of discovery and as evidence of the existence of phenomena, phenomena themselves are "not observable in any interesting sense of that term." (Woodward & Bogen, 1988, p. 305 f.)

Examples for data in physics and empirical psychology are patterns of discharge in electronic particle detectors and records of reaction times and error rates in various psychological experiments.

In Knipping (2003) the data includes photographs and transcripts of audio tapes made in French and German 8th grade classrooms during lessons on the proof of the Pythagorean theorem.

Examples for phenomena in physics, chemistry, empirical psychology and sociology are "gravitational radiation, Brownian motion, capacity limitations and recency effects in short term memory, and the proportionately higher rate of technical innovation among middle-sized firms in moderately concentrated industries" (Woodward, 1989, p. 393 f.).

An example of a phenomenon in Knipping (2003) is that "proofs" in the German classrooms observed follow a leitmotif that can be characterized as "anschauendes Deuten" (a visual contemplative approach), whereas the French classrooms observed follow a discursive leitmotif, where explicit verbal descriptions are essential. These phenomena only became transparent through comparison of the proving processes in the different classrooms and particularly by rigorous argumentation analyses of the proving processes (see Knipping, 2008, for a description of this process). The fine argumentation analyses of the talk in the class and the written texts on the blackboards were necessary to reveal these phenomena, which were at first, without comparisons, not directly observable in the data.

As a practical and methodological consequence of a clear conceptual distinction between data and phenomena, a large amount of research effort has to be spent to face the resulting problem of detecting a genuine phenomenon rather than some artefact of the experimental setting [1]:

> The problem of detecting a phenomenon is the problem of detecting a signal in this sea of noise [that is: in data], of identifying a relatively stable and invariant pattern of some simplicity and generality with recurrent features—a pattern which is not just an artifact of the particular detection techniques we employ or the local environment in which we operate. (Woodward, 1989, p. 397)

For example, many attempts have been made to explain the phenomenon of the shift in performance of Finnish students on international assessments from average performance prior to 1999 to leading the European countries in recent assessments. However, this phenomenon may simply be an artifact of the methods used to compare mathematics achievement in different countries. Finland's recent scores are on PISA surveys, while the older scores are on TIMSS. "It may be the case that Finnish educators chose to participate after 1999 in a test oriented to the kind of mathematics curriculum they had been training new teachers to implement" (Stotsky, 2012, p. 299). In other words, the phenomenon of increasing scores is an artifact of the change in testing.

Criteria for evaluating comparative studies in mathematics education are and should be diverse, related to the research questions and aims of these studies, and the theoretical perspectives and paradigms they are committed to. Nevertheless, we suggest that the philosophical principle presented here may contribute to the selection of appropriate criteria scientific adequacy in qualitative comparative research methods. To put it in a nutshell, the data-phenomena distinction may form a basis for finding appropriate qualitative counterparts to the criteria of objectivity and reliability used in quantitative approaches. We now turn to Sellar's stereoscopic view, which has potential to serve as a foundation for an adequate specification of the concept of validity in qualitative comparative research methods.

THEORETICAL SCIENTIFIC DISCOURSE AND SOPHISTICATED COMMON SENSE

So far, the data-phenomena-distinction may have helped to highlight that "detecting a signal in this sea of noise [that is: in data], of identifying a relatively stable and invariant pattern" (ibid., p. 397) is essential. But how can this be done? We suggest that a holistic view, as outlined in the following, is particularly promising for comparative approaches in mathematics education.

In the philosophy of language and general epistemology, so-called holistic conceptions of human knowledge (prominently argued by outstanding philosophers such as, W.V.O Quine, Wilfrid Sellars, or Donald Davidson, and taken up and developed further by, e.g., Robert Brandom) have strongly influenced current views of the development of and relation between common, every-day discourse and knowledge on the one hand and theoretical scientific discourse and knowledge on the other. We consider such an approach as important as it can help to identify phenomena by including the perspectives of the involved participants, teachers and students in mathematics education.

As a first step, Sellars calls for scientific theories to connect to pre-scientific and everyday constructions and interpretations of the observable physical world and to take into consideration the complicated inner logic of everyday circumstances.

> I suggested that the most fruitful way of approaching the problem of integrating theoretical science with the framework of sophisticated common sense into one

> comprehensive synoptic vision is to view it not as a piecemeal task—e.g. first a fitting together of the common sense conception of physical objects with that of theoretical physics, and then, as a separate venture, a fitting together of the common sense conception of man with that of theoretical psychology—but rather a matter of articulating two whole ways of seeing the sum of things, two images of man-in-the-world and attempting to bring them together in a "stereoscopic" view. (Sellars, 1963b, p. 19)

However, the stereoscopic view is itself still part of the scientific and not the common image of the world. A stereoscope allows a viewer to see a three dimensional scene by looking at two dimensional images taken from different perspectives through a suitable frame. Sellars' intent is for research to be done in such a way that the differing perspectives offered by different theoretical ways of seeing be presented in such a way that a researcher sees a single true-to-lived-experience representation of the world.

As a second step, Sellars addresses the question when such a scientific image is "complete". This is not the case when the "common man" has understood and taken over the whole scientific image itself. The Sellarian concept of completing the scientific image is rather functional with regard to the "common man": He shall be able to relate the circumstances and purposes of his actions to the general statements of the scientific theories.

> Thus, to complete the scientific image, we need to enrich it not with more ways of saying what is the case, but with the language of community and individual intentions, so that by construing the actions we intend to do and the circumstances in which we intend to do them in scientific terms, we directly relate the world as conceived by scientific theory to our purposes, and make it our world and no longer an alien appendage to the world in which we are living. (Sellars, 1963b, p. 40)

The Learner's Perspective Study attempts to provide such a stereoscopic view through "complementarity" (Clarke, Keitel, & Shimizu, 2006).

> Complementarity is fundamental to the approach adopted in the Learner's Perspective Study. This applies to complementarity of participants' accounts, where both the students and the teacher are offered the opportunity to provide retrospective reconstructive accounts of classroom events, through video- stimulated post-lesson interviews. It also applies to the complementarity of the accounts provided by members of the research team, where different researchers analyse a common body of data using different theoretical frameworks. (pp. 4-5)

A second element in the Learner's Perspective Study that supports a stereoscopic view is the bringing together of insiders' perspectives with the more typical outsider perspectives. This is especially the case in Clarke, Keitel & Shimizu (2006) where the authors describe their own school systems and cultures, supported by data including the voices of teachers and students in the classrooms studied in their countries. This is done in the framework provided by the overall study, however, so the authors are presenting their perspective with awareness of the differences that exist internationally.

Contrasting approaches to phenomena are evident when the TIMSS Video Study (Stigler & Hiebert,, 1999) is compared with the Learner's Perspective Study. Both sought to compare characteristics of teaching in different countries. In the TIMSS Video study the phenomenon of national teacher scripts for mathematics education was assumed and data was collected in order to study this phenomenon more closely. Because the phenomenon was assumed to occur everywhere, it was not necessary to collect data in multiple locations in each country. An outsider perspective was taken, in order to reveal culturally ingrained characteristics of teaching that are invisible (due to their familiarity) to insiders.

In contrast, in the Learner's Perspective Study the Hong Kong team (for example) started off by asking whether a Chinese teacher script of mathematics education actually exists (Mok, 2006). Multiple lessons were recorded, in two Chinese cities and in the classrooms of six teachers. Both outsider and insider perspectives were sought and both the teacher and the students' actions in the classroom were recorded (Clarke, Emanuelsson, Jablonka, and Mok, 2006). Data was collected not to study a presupposed phenomenon, but rather to understand teaching "based on the perspective of relevant stakeholders such as teachers and students" (Mok, 2006, p. 133).

In the Learner's Perspective Study participants were asked explicitly for their views to bring in the common sense perspective. In contrast, Knipping (2003) observed the participants in their everyday activity of teaching and learning, and their comments on their own activity were later juxtaposed with the results of the scientific analysis to provide a stereoscopic view.

The focus of Knipping's research was classroom proving processes. In order to reconstruct the rationale of mathematical proving practice of teachers and students in the classroom she needed both a view grounded in classroom practices as well as a rigorous analysis of classroom argumentations. Formal mathematical logic cannot capture the rationale of proving practice, as arguments that are produced during proving processes in the mathematics classroom follow their own peculiar rationale.

As a method Knipping (2008) proposes a three stage process: reconstructing the sequencing and meaning of classroom talk; analyzing arguments and argumentation structures; and finally comparing these argumentation structures and revealing their rationale through an interplay between the structures and the reconstructed classroom talk.

As a first stage the reconstruction of the common sense meaning of proof in classroom talk is essential. Interpretative methods are used to reveal what teachers and students say and mean, when they produce arguments and proofs in conversation.

Second, the arguments and argumentation structures are analyzed to provide the scientific view. This involves two moves, first analyzing local arguments on the basis of Toulmin's functional model of argumentation, and second analyzing the global argumentation structure of the proving process.

Comparing argumentation analyses of classroom proving processes in different contexts provided a stereoscopic view. It allowed not only the reconstruction of different leitmotifs of proving processes in classrooms, but also made the perspectives of the participating teachers on proving visible. These two-dimensional images come together into a three dimensional view of the proving processes in the classrooms. In this three dimensional view phenomena stand out and become more accessible to further study.

For example, teachers had communicated to their students during the proving process what they considered to be important, but this was at first not noticed in the data analyses. Only when the phenomena were described in form of leitmotifs, were these statements recognized and their significance acknowledged. One German teacher tried to encourage her students during the proving process as follows:

Teacher: Mmh. We don't know yet what exactly to write in the middle. But, you know, what I really like about your answer is that you looked for squares, you could somehow find the area of. But we don't know exactly the lengths of the sides of the inner square. b squared would be a square, that is here somewhere.

Maren: Mmh.

Teacher: … that does not work so well. Maybe you can find something else. Sarah, don't write, don't write, just think, just look. We can write this down later. Jan.

In a French classroom, the teacher guided her students' proving in this way:

Thierry : DCH and BCG.

Teacher: # are ..

Thierry : DCH are complementary.

Teacher: Yes. (9 sec.) And then, are complementary, did you write that? Yes? So?

Stephanie: You write that C, so the angle C is equal to 180 minus

Teacher: You have to say first that HC, why 180?

Stephanie: 180, because it's straight.

Teacher: Well, you have to say it at least, eh? We have not said it yet. We have said it, but we have not written it down. In a proof we have to write everything that we said, so, next line, so HCG equals 180 degrees.

The statements "Sarah, don't write, don't write, just think, just look. We can write this down later" and "We have said it, but we have not written it down. In a proof we have to write everything that we said" were part of the data all along, but once the phenomena of the leitmotifs was recognized in the scientific analysis, such comments could be recognized as also describing the phenomena, providing the "sophisticated common sense" framework for them. Comparing both the argumentation structures and the actual classroom talk provided "two whole ways of seeing the sum of things."

In the Learner's Perspective Study and in Knipping's research we have examples of comparative mathematics education research in which an attempt was made to investigate both phenomena that were identified by insiders in the community, and also phenomena identified by outside researchers, but which were described in ways that made them visible to the wider community as a whole.

There are other approaches represented at the CERME 8 conference that may address this challenge in comparative empirical research in mathematics education. We conclude with some further suggestions in this direction.

THOUGHTS TO THE FUTURE

In this paper we have discussed two principles we consider as essential for research in general, and for comparative work in mathematics education in particular. First, we emphasised the essential distinction between data and phenomena. Second, we discussed the issue and relation of common discourse and scientific discourse and why a 'stereoscopic' view is necessary to overcome unintended shortcomings of educational comparative research. We have referred to ideas from the philosophy of science and discussed how these ideas make a valuable and significant contribution to comparative research in mathematics education. More can be learnt from the works cited and generally from philosophy, and particularly from philosophically motivated socio-empirical studies in the context of mathematics.

Müller-Hill (2011), for example, has studied formalisability in proving practices in mathematical research. She finds that proving practices vary in a significant way in the different subfields of mathematical research, e.g. mathematical logic, algebra, geometry, applied mathematics. In a comparative approach, interviewing mathematicians as experts in these different fields, she investigates to what extent formalisability really plays a fundamental epistemic role, as is often presumed by general epistemology, in these subfields. Her socio-empirical study was based on an approach, called 'conceptual modelling', developed by Löwe & Müller (Löwe & Müller, 2011; see also Löwe, Müller & Müller-Hill, 2010). Conceptual modelling is a trans-disciplinary methodological framework that follows and reflects an interactive process in the transition from data to phenomena. The sophistication in this approach cannot be discussed in detail here. We only want to point out that looking more closely at, and reflecting more carefully on, the conceptualisation and application of the transition from data to phenomena, is vital. It is of particular interest for research in general and for comparative research in mathematics education in particular. This is especially important in respect to differences in quantitative approaches compared to qualitative approaches, and a possible mix of these methods. There is much more potential in the work of philosophy of science and particularly the works we have cited than we can discuss here.

Also the distinction of common discourse and scientific discourse is reflected in Müller-Hill's research and the work of her colleagues. She makes a distinction of 'armchair' epistemology of mathematics, related to normative philosophers' discourse about epistemic aspects of mathematical practice, and the discourse of mathematical practice, related to mathematicians' discourse about their work. Conceptual modelling proved to be particularly helpful to investigate the issue of formalisation not just as a philosophical or logical issue, by not only interviewing practicing mathematicians, but also combining and reflecting on these perspectives to provide a 'stereoscopic' view of the issue.

NOTES

1. In (Woodward, 1989, p. 453), the following example is given:

The authors of a recent study (Kamien and Schwartz 1982) devote approximately half of their book to arguing that the relationship [a relatively higher rate of technical innovation in moderately concentrated industries] is indeed real—that it has the characteristics of a phenomenon and is not an artifact of various statistical and measurement assumptions they employ. They investigate the relationship by regressing various measures of technical innovation on various measures of firm size and market concentration, the underlying assumptions about functional form being supplied by economic theory. They show that the relationship is relatively robust under different assumptions about how to measure these quantities and that it is fairly constant and stable across different industries. It is only in the second half of their book that the authors turn their attention to what they call 'theoretical explanation'.

REFERENCES

Biggs, J. B., & Watkin, D. A. (1996). *The Chinese learner: Cultural, psychological, and contextual influences.* Hong Kong: CERC & ACER.

Clarke, D., Keitel, C. and Shimizu, Y. (Eds.) (2006). *Mathematics classrooms in 12 countries: The insiders' perspective*. Rotterdam: Sense Publishers.

Clarke, D., Emanuelsson, J., Jablonka, E., & Mok, I.A.C. (Eds.) (2006). *Making connections: Comparing mathematics classrooms around the world.* Rotterdam: Sense Publishers.

Husén, T. (1967) (Ed.). *International Study of Achievement in Mathematics.* Stockholm: Almqvist & Wiksell.

Kaiser, G. (2002) (Ed.). Comparative Studies in Mathematics Education. *Zentralblatt für Didaktik der Mathematik, 34*(6).

Knipping, C. (2003). *Beweisprozesse in der Unterrichtspraxis: Vergleichende Analysen von Mathematikunterricht in Deutschland und Frankreich*. [Proving processes in teaching practices – Comparative analysis of mathematics teaching in France and Germany]. Hildesheim: Franzbecker Verlag.

Knipping, C. (2008). A method for revealing structures of argumentations in classroom proving processes. *ZDM The International Journal on Mathematics Education 40*(3), 427-441.

Löwe, B., & Müller, T. (2011). Data and phenomena in conceptual modelling. *Synthese, 182*, 131-148.

Löwe, B., Müller, T., & Müller-Hill, E. (2010). Mathematical knowledge: A case study in philosophy of mathematics. In B. Van Kerkhove, J. De Vuyst, J.P. Van Bendegem (Eds.), *Philosophical Perspectives on Mathematical Practice* (185-203). London: College Publications.

Mok, I.A.C. (2006). Shedding light on the East Asian learner paradox: Reconstructing student-centredness in a Shanghai classroom. *Asia Pacific Journal of Education 26*(2) 131-142.

Müller-Hill, E. (2011). *Die epistemische Rolle formalisierbarer mathematischer Beweise* [The epistemic role of formalizable mathematical proof] (Doctoral dissertation, University of Bonn, Germany). Retrieved from http://hss.ulb.uni-bonn.de/2011/2526/2526.htm

Science, Perception and Reality, Routledge & Kegan Paul Ltd; London, and The Humanities Press: New York, 1963; reissued in 1991 by Ridgeview Publishing Co., Atascadero, CA. (Stanford encyclopedia of philosophy).

Sellars, W. (1963a). Empiricism and the philosophy of mind. *Science, Perception and Reality* (pp. 127-196). London: Routledge & Kegan Paul.

Sellars, W. (1963b). Philosophy and the scientific image of man. In *Science, Perception and Reality* (pp. 1-40). Atascadero: Ridgeview.

Stigler, J., & Hiebert, J. (1999). *The teaching gap*. New York: Simon & Schuster.

Stotsky, S. (2012). Finnish lessons: What can the world learn from educational change in Finland by Pasi Sahlberg. *Journal of School Choice, 6*(2), 295-310.

Watkins, D. A., & Biggs, J. B. (Eds.) (2001). *Teaching the Chinese learner: Psychological and pedagogical perspectives.* Hong Kong: Comparative Education Research Centre.

Woodward, J. (1989). Data and phenomena. *Synthese, 79*, 393-472.

Woodward, J., & Bogen, J. (1988). Saving the phenomena. *The Philosophical Review, 97*, 303-352.