



HAL
open science

Why look into mathematics classrooms? Rationales for comparative classroom studies in mathematics education

Eva Jablonka

► To cite this version:

Eva Jablonka. Why look into mathematics classrooms? Rationales for comparative classroom studies in mathematics education. CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Charles University in Prague, Faculty of Education; ERME, Feb 2015, Prague, Czech Republic. pp.1724-1730. hal-01288008

HAL Id: hal-01288008

<https://hal.science/hal-01288008>

Submitted on 14 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Why look into mathematics classrooms? Rationales for comparative classroom studies in mathematics education

Eva Jablonka

King's College London, London, UK, eva.jablonka@kcl.ac.uk

Larger comparative studies of mathematics classrooms became most prominent as an appendix of international achievement studies. With the advance of video technology and the potentials it offers for researchers, however, comparative classroom observation studies became attractive and feasible also for smaller scale and low budget projects. This paper intends to provide a basis for discussing rationales for comparative studies of mathematics classrooms. It is suggested that the affiliation of classroom observation with school inspection on the one hand, and ethnographic research on the other hand, lingers on in comparative classroom research. The paper provides a narrative, illustrated by examples, that exposes this tension between evaluation and documentation inherent in the field.

Keywords: Classroom research, mathematics education, international and comparative education, teacher evaluation.

INTRODUCTION

There is no shared set of specific goals or methodologies that would justify characterising 'comparative studies of mathematics classrooms' as a sub-field, other than reference to the classroom as a focus of interest and 'comparative' as a research strategy. One might contemplate what constitutes a classroom as a unit; for example, the fact that all inside happen to gather in the same 'room' [1] in the presence of a teacher, or that there is a common practice in which all engage? The latter is not the case for individualised instruction or when students work on-line and access different sites, even if they are in the same room. Further, one can look at classrooms as micro-cultures, as reflecting school culture and schemes of work, as sites for teachers enacting broader curriculum traditions,

as sites for transmitting norms and values of larger communities, or as sites for differential distribution of curricular knowledge, amongst others. The 'unit of comparison' [2] will be different in each case. Not all of these interpretations are featured as comparative studies in the mathematics education literature. These are usually comparisons across (groups of) countries, regions or districts. While comparative studies of mathematics classroom practice are driven by diverse interests and methodologies, all assume that classrooms are an obvious site to look at.

The range of comparative studies of mathematics classrooms presented in this working group at previous CERME conferences highlights well the diversity of goals pursued [3]. The intention of this paper is neither to provide a comprehensive review of comparative classroom research nor to present its history in terms of research traditions, but to offer a narrative that exposes the tension between documentation and evaluation inherent in the field.

SCRUTINISING THE QUALITY OF EDUCATION

There is a long tradition of carrying out comparative classroom observations that seek to identify how pedagogy and curriculum relate to students' learning, with the aim of scrutinising the quality of pedagogic practice. In 1891–92, Joseph Mayer Rice took stenographic notes from lesson observations in classrooms in primary and grammar schools in 36 United States cities (East Coast and Mid West). In addition to lesson observations, he talked to teachers, parents and staff in education authorities, visited teacher education institutions, collected student productions and also tested year-3 pupils in arithmetic, amongst others. He classified schools into 'classes' of excellence, ranging from a mechanical 'antiquated' drill-and-practice to a

'scientific' approach. Further, he argued that the differences between the two higher 'classes' of practices can be brought out more effectively by a method of comparison. Accordingly, he summarised commonalities and differences that accounted for his ranking: In the schools of the two 'higher classes', teachers aimed at children's development in all faculties, took into account their 'developmental stage' and made the work interesting. In addition, he observed an 'excellent spirit' amongst the teachers, who treated the children kindly, were enthusiastic and constantly strived to increase their professional strength. The schools in the highest class differed in their practice of integrating school subjects in an 'attempt being made to teach the subjects in their natural relation to each other' (Rice, 1893, p. 222). Further, he notes a strong focus on having children express their ideas in written form in all subjects as a distinguishing feature of the highest class. He also reports that the supervision in these districts takes the form of guidance, instruction and inspiration rather than inspection.

Notably, Rice took the school as his unit of analysis and occasionally included the local policy context as well as the economic conditions of the schools' district in his report. For example, he observed the 'poorest' teaching as regards 'methods and tone' (Rice, 1893, p. 131) in a primary school in one of the poorest neighbourhoods in Boston. [4] This points to another feature of his report, namely the assemblage of impressions on curriculum and authority relations from classroom observations across a range of school subjects.

The report is explicitly evaluative and critical, with clear preferences for a 'progressive' curriculum and pedagogy as outlined in the introduction (notably the labelling of the preferred pedagogic practice as 'scientific'). However, the principles for compiling his notes and the collected student work remain largely implicit. The writing style is journalistic and often scathing, which might be partly motivated by Rice's trip being financed by the magazine *The Forum* that first published part of his report as a series of articles and later financed a second study trip. His overall approach to data generation could be classified as ethnography, while his goal clearly is critical evaluation. Hence he engaged in two different activities.

Around 120 years later, comparative classroom studies produce large data sets due to an increased level

of methodological differentiation, such as refined multi-camera video observation, complex scaling and other statistical techniques for constructing reliability and validity of a range of measurements, inclusion of a variety of context variables, systematic procedures for 'coding' segments of classroom video footage, often embedded in research processes characterised by a division of labour in academic work (e.g., coders/report writers). These large studies are affiliated with educational psychology and its tradition of measurement. On the other hand, there are many smaller projects, mostly within a tradition of ethnographic classroom research. The distinction between the activities of evaluation and documentation, however, does not necessarily specialise studies into these two forms. The picture looks more complex.

COMPARING AND MEASURING THE QUALITY OF TEACHING

Comparing and measuring the quality of teaching can be seen as an elaboration of strategies that have been used in many places for assessing teachers during classroom visits, including judgment of the level of teachers' subject-related knowledge and appropriateness of pedagogic strategies, in addition to curriculum coverage and classroom management. Instead of a holistic evaluation based on largely implicit performance criteria (practiced in many places by school inspectors), coding of lessons is based on developing a list of relevant aspects and identifying specific performances that relate to these aspects. Presence/ absence or proportion of lesson time spent on different activities, are then used as a basis for numerical measures, but for lack of theory also expert ratings for each aspect are conducted ('high-inference' coding). The latter leaves the criteria for the 'marks' on each aspect implicit. Scales based on ratings of a range of aspects of teacher performance are increasingly employed in the USA for formative teacher assessment, evaluation of curriculum policy and professional development (Hill et al., 2012). Sapiro and Sorto (2012) used adaptations of such a scale (the Mathematical Quality of Instruction - MQI score) for comparing teaching quality in Botswana and South Africa, where teaching appeared much less complex in terms of pedagogical techniques and use of resources than in the USA, which renders the application of the measure questionable. They complemented the measure with curriculum coverage and other codes. In this context, with reference to (Knight & Sabot, 1990), Carnoy (2012)

uses the term ‘natural experiment’ for the situation where national social conditions are similar but policies and outcomes differ. Even though measures of instructional quality originate in the idea that students’ scores on mathematics tests are not an appropriate measure of the quality of teaching and hence classroom teaching practice needs to be looked at, correlations with some measures of student achievement are still often incorporated in studies that use such measures and are sometimes used as an argument for their validity. Comparison and evaluation become intertwined.

Studies that aim at comparing instructional effectiveness of different teaching approaches by means of randomised controlled trials have not (yet) been associated with comparative studies in mathematics education. These studies reflect a comeback of experimentalism, asserting it a position of scientific superiority for identifying ‘what works’. They occasionally include classroom observation in order to check the fidelity of the teachers’ dispensing of the intervention (treatment), or to complement measurement of gain scores with scores from classroom observations (e.g., Clements, Sarama, Spitler, Lange, & Wolfe, 2011). Classroom observation in these studies is asserted an auxiliary role, as the classroom is only relevant in relation to the statistical regularity the black box produces as its achievement outcomes. Hence these studies do not qualify as comparative studies of mathematics classrooms.

RELATING INTERNATIONAL ACHIEVEMENT RANKINGS TO TEACHING APPROACHES

The international comparative measurement of average outcomes, which may include affective co-productions (e.g. attitudes) in addition to academic achievement, preceded international studies that included mathematics classroom observations. As a result, in many classroom observation studies, the nominal unit of comparison is countries and administrative units, or larger units defined by shared cultural traditions (such as the ‘Confucian Heritage Culture’ or CHC), with the aim of identifying characteristics of representative mathematics teacher practice based on random samples of lessons. This is motivated by an interest in high achieving countries (e.g., Hiebert et al., 2003). Comparisons of teacher practices in countries at the bottom end of the achievement rankings rarely attract wider than local interest.

In reflecting upon experiences with the video surveys from the Third International Mathematics and Science Study (TIMSS) and its follow-up study (TIMSS-R), Stigler, Gallimore and Hiebert (2000, p. 87) noted that “the most obvious reason to study classrooms across cultures is that the effectiveness of schooling, as measured by academic achievement, differs across cultures (e.g., Peak, 1996)”. A similar observation by the Under Secretary of State for Education recently prompted the invitation of 60 Shanghai teachers to give workshops for teachers in England. Stigler and colleagues (2000), however, also mention the illuminating effect of making the familiar look unfamiliar when confronted with other cultural practices as a rationale for the TIMSS video studies, which does not suggest replanting of teaching practice. They describe the TIMSS video surveys as ‘integrating’ the tradition of ethnographic classroom research with a survey tradition of schools and classrooms that aims at allowing generalization to a wider population.

As ‘comparing’ only rests on generating analytical categories to find commonalities and specialising these for describing differences or variations, there is no restriction to the level of detail to which this strategy can be applied; neither is there any restriction to what could be usefully looked at in mathematics classrooms. Lesson structure, based on low-inference codes for time spent on a range of activities, has been a focus of many studies. In doing so, lesson structure (‘script’, ‘pattern’) is taken as both an unconscious routine (e.g., Hiebert et al., 2003) or as an outcome of a more conscious act of teachers’ planning (e.g., Leung, 1995; see also the discussion in Clarke et al., 2007). Generally, it is not much looked at how instruction is constrained by traditions of classroom management other than through interpreting differences in student behaviour in terms of culturally typical inclinations, e.g., of Chinese students being obedient, a form of analysis, which Jablonka (2013) sees affiliated with cultural essentialism. Wong (2004) is critical towards attempts of interpreting differences in terms of culturally typical behaviour of Chinese students.

As the achievement measures are reported in the form country/region averages, relating classroom teacher practice directly to these measures can only be suggestive. In the first TIMSS study, the classroom samples for the achievement test and the video samples did not overlap in Japan and US, but only in Germany. Klieme and Bos (2000) found a differential item function on

the Japanese test results in the achievement test for the type of mathematical tasks they observed in the lessons. For the German videos, relations between teaching practice and achievement were explored through high-inference ratings of some aspects of teaching quality with not much significant outcome. Clarke and colleagues (2007) observed more variability in lesson structure than reported in the TIMSS video study, and see the location of the lesson in a topic sequence as a key influence on a lesson's structure. The TIMSS video studies have not only been criticized for taking single lessons as analytical units, but also for not exploiting the potential of their rich qualitative data sets (Stigler et al., 2000; Andrews, 2007; Clarke et al., 2007). Other studies reported more consistent patterns in lesson structure in some places, as for example in China (Beijing) and also included a wider range of analytical categories than the TIMSS (Leung, 1995). In a random sample of lessons from Finland and Iceland (20 each), Savola (2010) found that the Finnish lessons followed a 'conventional' review-lesson-practice script, whereas more than half of the Icelandic lessons exhibited versions of individualized learning. When looking at the range of codes employed for characterising lesson structure, a major challenge appears to be the apparent embeddedness of instructional and regulative discourse [5]; hence categories for coding tend to include 'mixes' of privileged teacher-student relations and mathematical knowledge structures.

Relating characteristics of teaching practices to national achievement levels did not work for the TIMSS-R video studies either. The official executive TIMSS report (National Center for Education Statistics, 2003) is not explicitly appreciative of any practice, but the mentioning of the achievement levels of the countries implicitly suggests a ranking of some aspects identified in high achieving regions: "Results from the 1999 study of eighth-grade mathematics teaching among seven countries revealed that, among the relatively high-achieving countries, a variety of methods were employed rather than a single, shared approach to the teaching of mathematics." (p. 11). The Pythagoras Study (Klieme, Pauli, & Reusser, 2009) set out to further investigate relations between characteristics of teaching practice and student achievement as well as motivation in 20 classrooms from Germany and Switzerland each, covering a broad range of achievement levels. They focussed on particular topics (the Pythagoras Theorem and algebra 'word problems'),

including some briefing for the teachers about how to approach these. This study, then, departs from the ethnographic tradition mentioned by Stigler and colleagues (2000) and moves towards an intervention and evaluation study. It included a broad range of methodological questions and produced a range of publications with detailed analyses. These studies are affiliated with educational psychology usually associated with measurement and with a conception of the curriculum as socially and culturally neutral content delivered in different ways by teachers.

As to the PISA, there is a limit in identifying the schools and classrooms that were chosen for further research. Hence, the use of the available contextual school data for comparative classroom studies is restricted. There is anecdotal evidence, however, that some complementing video studies are planned by the OECD. All attempts at relating regularities of teacher practice to a country's average achievement are further hampered by severe methodological problems in the production of achievement rankings, including sampling problems. Also, different international achievement measures privilege different forms of mathematical knowledge (Wu, 2010), which are balanced differently in national curricula (Cai & Howson, 2013).

This does not derogate the value of looking at larger samples of classroom videos. Most of the reports, however, are ignorant of the socio-political contexts, in which these classrooms are situated. Not much is said, for example, about practices of streaming and selection based on educational credentials, school fees, accountability regimes or working conditions of teachers. The highly competitive nature of the college entrance test in China or the amount of private tutoring in some places (cf. Bray & Kwo, 2014) are rarely mentioned in reports. In addition, classrooms are featured as culturally and socially homogeneous conglomerates, and one is left in the dark about the student intake and neighbourhood of the schools.

EXPLORING REGULARITIES IN MATHEMATICS CLASSROOMS WITHIN AND ACROSS DIFFERENT CURRICULUM TRADITIONS

A range of comparative observation studies are more in line with an ethnographic tradition or interpretive sociology (e.g., Kaiser, 2002; Knipping, 2003). Rather than analysing random samples of classrooms, typicality is achieved through selecting classrooms of

experienced teachers judged locally as competent (Clarke, 2006). The practices of these teachers and students can then be seen as representing a range of pedagogic discourses in different curriculum traditions. Identification of regularities and exposure of similarities and differences facilitates interrogation of assumptions and so opens up possibilities for theoretical differentiation and practical innovation that might not otherwise be recognised. Exploring regularities of classroom practice can also be aided by use of statistical techniques (e.g., Andrews, 2009).

The Learner's Perspective Study (LPS) (e.g., Clarke, Keitel, & Shimizu, 2006a) views classroom practice as co-constructed by teachers and students and so included student video-stimulated reconstructive interviews and a camera focussing on changing groups of students (during ten lessons), in addition to a teacher camera and a whole-class view, which affords considering multiple points of view. This is a reminder of the insight that different techniques for producing video data reflect different analytical gazes (cf. Hall, 2000). The variety of analyses conducted by different research groups and teams composed of insiders and outsiders is reported in a series of volumes (Clarke et al., 2006a; Clarke, Emanuelsson, Jablonka, & Mok, 2006b; Shimizu, Kaur, Huang, & Clarke, 2010; Kaur, Anthony, Ohtani, & Clarke, 2013) and other publications. This demonstrates the productivity of a departure from the search for 'best practice' and shows how a substantial data set can be used for iteratively developing languages of descriptions ('theories') as well as for complementary analyses. The advantage of publishing reports that contain transcripts, student work and examples of teaching material is that the audience can engage in their own interpretations.

DEPARTURES AND POSSIBILITIES

While classroom practice also produces valued learning outcomes not captured by the examination system, curricular choices might appear most aligned with external examination specifications or with criteria derived from other policies (e.g., school inspection regimes). In systems with strong regulation there is very limited space for teachers to act upon curriculum. Weaker regulation affords more diversity in classroom practice and more adaption to apparent needs of different categories of students. Hence it would seem reasonable to analyse classroom practice with reference to a level where teachers and schools

make deliberate decisions about curriculum, which clearly differs in different socio-political contexts. This strategy departs from attempting to uncover taken-for-granted lesson scripts, patterns or rituals.

Cross-subject studies with the same students or the same teacher would allow to create differentiated accounts of regulative principles of instruction and behaviour in relation to teacher authority and mathematical knowledge structures. Comparison is then a methodological strategy for creating substantial variation in the empirical data as a starting point for developing a language of description (see Gellert & Jablonka, 2009; Knipping, Reid, Gellert, & Jablonka, 2008). Studies with a broader conception of curriculum that illuminate issues of power, identity and subjectivity are still almost absent amongst comparative studies of mathematics classrooms.

Irrespective of the substantial amount of work produced by comparative mathematics classroom studies, those that employ quantifications, in particular for establishing descriptive causality ('findings') between characteristics of teacher practice and some 'student outcomes', still (or again) earn more scientific respectability, despite their often antiquated mechanistic conception of curriculum.

REFERENCES

- Andrews, P. (2007). Negotiating meaning in cross-national studies of mathematics teaching: kissing frogs to find princes. *Comparative Education*, 43(4), 489–509.
- Andrews, P. (2009). Comparative studies of mathematics teachers' observable learning objectives: validating low inference codes. *Educational Studies in Mathematics*, 71(2), 97–122.
- Bernstein, B. (2000). *Pedagogy, symbolic control and identity. Theory, research and critique. Revised edition*. Oxford, UK: Rowman & Littlefield.
- Bray, M., & Kwo, O. (2014). *Regulating private tutoring for public good: Policy options for supplementary education in Asia*. Hong Kong, China: Comparative Education Research Center (CERC): The University of Hong Kong.
- Cai, J., & Howson, A. G. (2013). Toward an international mathematics curriculum. In M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick, & K. S. F. Leung (Eds.), *Third international handbook of mathematics education* (pp. 949–974). Berlin, Germany: Springer.

- Carnoy, M. (2012). Introduction. Comparing learner performance in southern Africa: A natural experiment. *Prospects*, 42, 363–369.
- Clarke, D. J. (2006). The LPS research design. In D. J. Clarke, C. Keitel, & Y. Shimizu (Eds.), *Mathematics classrooms in twelve countries* (pp. 15–29). Rotterdam, The Netherlands: Sense Publishers.
- Clarke, D. J., Keitel, C., & Shimizu, Y. (Eds.) (2006a). *Mathematics classrooms in twelve countries: The insider's perspective*. Rotterdam, The Netherlands: Sense Publishers.
- Clarke, D. J., Emanuelsson, J., Jablonka, E., & Mok, I. A. C. (Eds.) (2006b). *Making connections: Comparing mathematics classrooms around the world*. Rotterdam, The Netherlands: Sense Publishers.
- Clarke, D. J., Mesiti, C., O'Keefe, C., Jablonka, E., Mok, I., Xu, L. H., & Shimizu, Y. (2007). Addressing the challenge of legitimate international comparisons of classroom practice. *International Journal of Educational Research*, 46(5), 280–293.
- Clements, D. H., Sarama, J., Spitler, M. E., Lange, A. A., & Wolfe, C. B. (2011). Mathematics learned by young children in an intervention based on learning trajectories: A large-scale cluster randomized trial. *Journal for Research in Mathematics Education*, 42(2), 127–166.
- Gellert, U., & Jablonka, E. (2009). "I am not talking about reality": Word problems and the intricacies of producing legitimate text. In L. Verschaffel, B. Greer, W. van Dooren, & S. Mukhopadhyay (Eds.), *Words and worlds: Modelling verbal descriptions of situations* (pp. 39–53). Rotterdam, The Netherlands: Sense Publishers.
- Hall, R. (2000). Video-recording as theory. In R. Lesh & A. Kelly (Eds.), *Handbook of research design in mathematics & science education* (pp. 457–486). Mahwah, NJ: Lawrence Erlbaum.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Hill, H. C., Charalambous, C. Y., Blazar, D., McGinn, D., Kraft, M. A., Beisiegel, M., et al. (2012). Validating arguments for observational instruments: Attending to multiple sources of variation. *Educational Assessment*, 17, 1–19. doi: 10.3102/0002831210387916
- Jablonka, E. (2013). Boredom in mathematics classrooms from Germany, Hong Kong and the United States. In B. Ubuz, Ç. Hazar, & M. A. Mariotti (Eds.), *Proceedings of the 8th Congress of European Research in Mathematics Education* (pp. 1885–1894). Ankara, Turkey: Middle East Technical University, ERME.
- Kaur, B., Anthony, G., Ohtani, M., & Clarke, D. J. (Eds.) (2013). *Student voice in mathematics classrooms around the world*. Rotterdam: Sense Publishers.
- Kaiser, G. (2002). Educational philosophies and their influence on mathematics education: an ethnographic study in English and German mathematics classrooms. *ZDM – The International Journal on Mathematics Education*, 34(6), 241–57.
- Klieme, E., & Bos, W. (2000). Mathematikleistung und mathematischer Unterricht in Deutschland und Japan: Triangulation qualitativer und quantitativer Analysen am Beispiel der TIMSS-Studie. *Zeitschrift für Erziehungswissenschaft*, 3(3), 359–379.
- Klieme, E., Pauli, C., & Reusser, K. (2009). The Pythagoras Study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. In T. Janik & T. Seidel (Eds.), *The power of video studies in investigating teaching and learning in the classroom* (pp. 137–160). Münster, Germany: Waxman Verlag.
- Knight, J., & Sabot, R. (1990). *Education, productivity, and inequality: The East African natural experiment*. Washington, DC: Oxford University Press, for the World Bank.
- Knipping, C. (2003). Learning from comparing: a review and reflection on qualitative oriented comparisons of teaching and learning mathematics in different countries. *ZDM. Zentralblatt für Didaktik der Mathematik*, 35(6), 282–293. doi:10.1007/BF02656692
- Knipping, C., Reid, D. A., Gellert, U., & Jablonka, E. (2008). The emergence of disparity in performance in mathematics classrooms. In J. F. Matos, P. Valero, & K. Yasukawa (Eds.), *Proceedings of the Fifth International Mathematics Education and Society Conference, Part 2* (pp. 320–329). Lisboa & Aalborg: Universidade de Lisboa & Aalborg University.
- Leung, F. K. S. (1995). The mathematics classroom in Beijing, Hong Kong and London. *Educational Studies in Mathematics*, 29(4), 297–325.
- National Center for Education Statistics (2003). *Highlights from the TIMSS 1999 Video Study of eighth-grade mathematics teaching*. Washington, DC: U.S. Department of Education, National Center for Education Statistics. Retrieved from the NCES website: <http://nces.ed.gov/timss>.
- Peak, L. (1996). *Pursuing excellence: Initial findings from the Third International Mathematics and Science Study, NCES 97–198*. Washington, DC: U.S. Government Printing Office.
- Rice, J. M. (1893). *The public-school system of the United States*. New York, NY: Century.
- Sapire, I., & Sorto, M. A. (2012). Analyzing teaching quality in Botswana and South Africa. *Prospects*, 42, 433–451. doi: 10.1007/s11125-012-9250-x

- Savola, L. (2010). Structures of Finnish and Icelandic Mathematics Lessons. In B. Sriraman, C. Bergsten, S. Goodchild, C. Michelsen, G. Palsdottir, O. Steinhorsdottir, & L. Haapasalo, (Eds.), *The first sourcebook of Nordic Research in Mathematics Education* (pp. 519–538). Charlotte, NC: Information Age Publishing.
- Shimizu, Y., Kaur, B., Huang, R., & Clarke, D. J. (Eds.) (2010). *Mathematical tasks in classrooms around the world*. Rotterdam, The Netherlands: Sense Publishers.
- Stigler, J., Gallimore, R., & Hiebert, J. (2000) Using video surveys to compare classrooms and teaching across cultures: examples and lessons from the TIMSS video studies, *Educational Psychologist*, 35(2), 87–100.
- Wong, N. Y. (2004). The CHC learner's phenomenon: its implications on mathematics education. In L. Fan, N. Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 503–534). Singapore: World Scientific.
- Wu, M. (2010). Comparing the similarities and differences of PISA 2003 and TIMSS. *OECD Education Working Papers*, No. 32. OECD Publishing. doi: 10.1787/5km4psnm13nx-en

ENDNOTES

1. Including, for example, open-air arrangements found in rural areas in periphery countries or in refugee camps.
2. The distinction between unit of analysis and unit of comparison has been discussed in the working group on comparative studies in mathematics at CERME 7.
3. These are published in the CERME proceedings. For lack of space for an extensive bibliography, the examples mentioned in this paper will remain limited.
4. It needs to be pointed out that Rice did not comment on the racist nature of some student writing assignment, which he collected from a grammar school.
5. The notion of 'embeddedness' of these two discourses is taken from Bernstein (e.g. 2000).