

## 16 Comparative Studies in Mathematics Education

Eva Jablonka (DE), Paul Andrews (SE), David Clarke (AU) Constantinos Xenofontos (CY)

### 1. Introduction

#### 1.1 History of the group

The group *Comparative Studies in Mathematics Education* was established in 2007 at CERME 5), reflecting a growing interest of researchers and regional school authorities in comparative and international dimensions of education, not least in response to cross-national achievement studies carried out by supranational organisations. The group has been truly multi-linguistic since its inception; but more recently it has attracted participants from locations far beyond geographical Europe or ERME affiliated regions, as for example Ghana, South Africa, China, Japan, Australia, Vietnam, Mexico, Chile and USA.

Immediate access to colleagues from contexts that participants had chosen for their cross-cultural comparisons proved invaluable to better mutual understanding of cultural, social and political contexts of mathematics education. At each CERME, participants included international postgraduate research students who worked with supervisors at universities in Europe and individual researchers who had moved between contexts. Hence the meetings provided an excellent forum for productive interaction of insiders’ and outsiders’ perspectives and exchange of experiences in support of the development of culturally sensitive research designs and analyses. The number of presenting participants remained relatively small throughout the group’s history, with many opportunities for sustained and focussed discussion in an inclusive atmosphere, greatly aided by the reservoir of languages available for communication and the help provided by many participants who were fluent in a couple of these.

#### 1.2 Developing aims and frameworks for comparative studies

One of the group’s initial aims was to overcome simplistic identifications of culture with country and associated misinterpretations of comparative studies. In acknowledgement of cultural complexities within many countries or administrative units, the scope of the work has not been restricted to comparative studies across national educational systems. Consequently, the group adopted an eclectic perspective in its interpretation of comparison as referring to any study that documents, analyses, contrasts or juxtaposes cross-cultural or cross-contextual similarities and differences across all aspects and levels of mathematics education.

Besides sharing outcomes of empirical studies, the group aimed at developing and defining research methodologies specific to comparative studies and developing better understanding of how various theoretical approaches and conceptual frameworks shape the goals and the design of comparative research in mathematics education.

While understanding the role of culture in the construction of mathematics teaching and learning remained a shared interest of the groups' participants, studies accepted for presentation in the working group often included much smaller units of 'culture'. Consequently, rather than assuming or seeking coherence in the field of comparative studies in mathematics education, the working group invited contributions based on a broad range of methodologies and analytical frameworks that reflected the diversity of aims and concerns of researchers working in different contexts.

Yet, comparisons of aspects of curriculum and pedagogy across (national) education systems remained the focus of many studies presented and discussed in the group. While agreement was reached that comparison cannot itself constitute the goal of such studies, comparison was seen as being always of interest in providing a new 'lens' and making the familiar look unfamiliar. The often very vivid discussions revealed how comparison across systems can afford insights beyond the reach of research confined to one context. Further, cross-system comparison of 'distant' cultures was discussed as a research strategy because contrasting cases may have substantial theoretical and methodological bearing.

## **2. Foci of Comparison**

### 2.1 Aspects of mathematics education addressed and compared in empirical studies

A majority of cross-system studies addressed aspects of curriculum and pedagogy, with some also including assessment practices and policy. Curriculum studies included analyses of curriculum goals and 'standards' in official documents (i.e. An, Mintos, & Yigit, 2013; Bjarnadóttir, 2007; Gosztonyi, 2015), representation in textbooks and teacher guides (i.e. Bofah & Hannula, 2011; Cabassut & Ferrando, 2013; Ponte & Marques, 2007; Xenofontos & Papadopoulos, 2015), as well as curriculum enactment in classroom practice (Clarke & Xu, 2010). Many of these studies focussed on the teaching of selected mathematical topics, some on broader competences, such as number sense (Andrews, Sayers, & Marschall, 2015), problem solving and modelling (Saeki, Matszaki, Kawakami, & Lamb, 2015), or on specific aspects of

pedagogy, such as language use and modes of questioning (Hommel & Clarke, 2015) and assessment methods (Brown, 2007).

In-service teachers were the focus of some cross-system studies examining participants' professional motives (Andrews, 2010), reactions to classroom interventions and professional development activities (Cabassut & Villette, 2011), and views on 'good' examination tasks (Peng, Sollervall, Stadler, Shang, & Ma, 2015). Other studies were concerned with prospective teachers, looking at issues like mentor-prospective teacher relationships (Knutsson, Hemmi, Bergwall, & Ryve, 2013), the use of metaphors for describing mathematics (Kiliç, 2011), and problem solving in pre-service teacher training (Kuzniak, Parzysz, Santos-Trigo, & Vivie, 2011).

Students also featured commonly in cross-system studies. Examples comprise investigations of students' mathematical understanding and meaning construction (Nguyen & Gregoire, 2013; Vollstedt, 2007), examinations of factors that influence learners' mathematical literacy (Törnroos, 2007), explorations of emotional aspects and affective co-productions of mathematics learning (Jablonka, 2013; Tuohilampi, Hannula, Giaconi, Laine, & Näveri, 2013), and identities of 'elite' students (Saari, 2010).

The group recognised the difficulties with generalisations at the level of 'culture', particularly in contexts where cultural homogeneity of school students or teachers cannot be assumed, and agreed that a more nuanced articulation of 'culture' is necessary. Some participants foregrounded social relations in their conceptualisation of 'culture', while others focussed more on shared traditions and systems of values.

## 2.2 Systemic contexts chosen for comparison:

A variety of reasons were provided by the authors of the cross-system papers, explicitly or in some cases implicitly, in order to justify their preferences for the comparisons of specific systemic contexts. It should be noted, however, that the reasons presented below are not mutually exclusive, and that for every study more than one reason may apply. The majority of studies located their choices within an argument of “cultural proximity versus cultural distance”. In regards to the former, we find examples like those of Ponte and Marques (2007), comparing data from Portugal, Brazil, and Spain, of Hannula, Lepik, Pipere, and Tuohilampi (2013), comparing data from Estonia, Latvia, and Finland, and of Vula, Kingji-Kastrati, and Podvorica (2015), from Albania and Kosovo. Other studies emphasise the different educational traditions of the systems under scrutiny, such as that of Saari (2010) who compares data from Finland and the USA, of

Kuzniak, Parzys, Santos-Trigo, and Vivier (2011) and their study in France and Mexico, as well as of Nguyen & Gregoire (2013), highlighting the cultural and linguistic differences between the French Belgium and Vietnam. A few studies comment on the geographical proximity of the compared contexts, like Bjarnadottir (2007), examining issues regarding Iceland, Denmark, Norway, and England.

Co-authors' locations appear to be another feature that, implicitly or explicitly, enables cross-system comparisons. For instance, Xenofontos and Papadopoulos (2015) talk explicitly about their respective locations in Cyprus and Greece, while in the cases of Cabassut and Ferrando (2013), comparing issues in France and Spain, and of Saeki, Matszaki, Kawakami, and Lamb (2015) and their study concerning Australia and Japan, this is implied by the authors' affiliating institutions. In a few studies, typically single-authored, the choice of a comparative approach is justified by claims of the author's familiarity with the systems under examination (see, for instance, Cabassut, 2007, Gosztonyi, 2015, and Kiliç, 2011). A relatively large number of papers draw on data from larger cross-system projects, and compare findings from typically two to four of the participating countries. Such examples constitute the studies of An, Mintos, and Yagit (2013), Andrews (2011), Clarke and Xu (2010), and Jablonka (2013). Finally, a couple of studies explain how the choice of a comparative cross-national approach was adopted for testing the cross-cultural affordances of frameworks or theoretical models under development (see Andrews, Sayers, & Marschall, 2015, and Bofah, 2015).

### 2.3 Comparisons beyond 'culture'

In addition to studies that used the location in a particular school system for defining their unit of comparison, a couple of presentations reported investigations with other comparative foci. These included comparisons across school sectors and between students at different levels of attainment or in different streams, or belonging to an ethnic minority. In regards to comparisons between students at different attainment levels, we find studies like those of Maréchal (2011), who talks about 'praxeologies' in different classes for 'ordinary' or other students in Geneva, Schäfer and Winkler (2011), investigating metacognitive actions used by high-achieving and low-achieving German students chosen from different streams, and Saika and Rosiek (2015), who report an experimental study of eye tracking of high and low-achieving students and their problem solving competences in Poland. Another focus of comparison concerned school sectors within a single educational system. For example, Nilsen (2010) compares Norwegian lower- and upper-secondary teachers' beliefs, while Larson and Bergten (2013) explore 'praxeologies' in lower and

upper secondary classrooms in Sweden. Finally, we see studies comparing settings with demographic differences in school intake, as, for example, Eisenmann and Even (2007) who follow the same teacher, applying different pedagogies in different schools; studies comparing differently qualified teachers, like Vantourout (2007); and longitudinal comparisons within the same context, like Branchetti and her colleagues (2015), who follow the same cohort of Italian students and compare their test performance in grade 6 and grade 8.

These studies were comparative in so far as differences at the individual level were taken as contributing to the re-presentation of the group in comparison with the other group. A methodological question discussed in this context was whether these groups exist independently of the comparative research; if not, then the notion of culture or micro-culture that is being compared gets lost. These questions initiated a continuous methodological discussion. Consequently, in 2015 (CERME 9) the call for papers included an invitation to contribute to a discussion that outlines the delineation of comparative and non-comparative research.

### 3. Developing Methodology and Theory

#### 3.1 Methodological approaches

The methodological stances adopted by the working group’s members have shown considerable variation, although there have been many more qualitatively focused than quantitative. In broad terms, they have fallen into five forms, summarised in table 1.

	2007	2009	2011	2013	2015	Total
Case study	3	5	4	3	5	20
Documentary analysis	3	1	2	3	3	12
Small survey		1	4	3	1	9
Position paper	1	1		2	2	6
Large survey	1				2	3

**Table 1: Summary data of the forms of study presented to the working group**

#### *Case studies*

The most frequently occurring methodological tradition was case study in various forms, although the majority of studies identified below were not reported as case studies by their authors. Typically, the case element of the study was left implicit with authors attending to details pertaining to other methodological aspects such as theoretical perspectives or the articulation of an analytical procedure. In some respects, acknowledging the brevity of a CERME paper, this is probably not surprising. In categorising case study choices have to be made between, say,

whether to categorise by approaches to data collection and analysis or by the nature of the cases under scrutiny. In this instance we have chosen to categorise first by the nature of case as it better represents the diversity of the working group’s papers.

In broad terms, case studies as identified in this manner fell into four categories in which the cases were countries, classrooms, teachers or students. The most common of these, accounting for 12 of these 20 studies, involved the country as the case. These included analyses of lessons focused on particular topics and students of various ages (Andrews, 2011; Andrews, Sayers & Marschall, 2015; Back, Sayers & Andrews, 2013; Navarra, Malara & Ambrus, 2010; Saeki, Matszaki, Kawakami & Lamb, 2015), comparisons of particular teacher education interventions (Knutsson, Hemmia, Bergwall & Ryve, 2013; Kuzniak, Parzysz, Santos-Trigo & Vivier, 2011), analyses of classroom interactions and their meaning for participants (Clarke & Xu, 2010; Hommel & Clarke, 2015; Vollstedt, 2007), teachers’ perceptions of the nature of good problems (Peng, Sollervall, Stadler, Shang & Ma, 2015) and a study of elite students experiences of mathematics learning (Saari, 2010).

The remaining eight studies were undertaken in single cultural contexts in which different groups within a particular setting were compared. Case studies focused on the student included differently conceptualised comparisons of the mathematical behaviour of high- and low-achieving students (Sajka & Rosiek, 2015; Schäfer & Winkler, 2011) examinations of minority students’ mathematical resilience (Mulat & Arcavi, 2010) and a comparison of upper and lower secondary students’ perceptions of school mathematics (Larson & Bergsten, 2013). Two studies exploited students’ socio-economic affiliation as the case, as in Eisenmann and Even’s (2007) comparison of grade 7 algebra lessons taught by the same teacher in different schools and Maréchal’s (2011) investigation of the teaching of addition to grade one students in three structurally different classrooms. Finally, two cases involved the teacher. One of these compared preservice teachers’ approaches to the assessment of fictitious student solutions (Vantourout, 2007), while the other compared lower secondary and upper secondary teachers’ beliefs about the teaching of mathematics (Nilsen, 2010).

Typically, case studies drew on qualitative data, whether derived from interview or observation transcripts, although the means of analysis varied considerably. A number exploited the constant comparison process of the grounded theorists. In some papers, this was explicit (Andrews, 2011; Vollstedt, 2007; Peng et al., 2015) and others implicit (Knutsson et al., 2013; Mulat & Arcavi, 2010; Saari, 2010). In other studies, established theoretical frameworks were applied to the

researchers’ qualitative data (Eisenmann & Even, 2007; Larson and Bergsten, 2013; Maréchal, 2010; Schäfer & Winkler, 2011), while others applied frameworks derived as part of their, typically, larger project of which the reported paper was part (Andrews, et al., 2015; Back et al., 2013; Saeki et al., 2015). Finally, in this section, (Nilsen, 2010) offered an unspecified analysis of qualitative data derived from lesson observations and interviews.

A smaller number of studies exploited quantitative analyses of their different data. Two, due to the ways in which data were focused on particular tasks, undertook analyses of time on various aspects of the tasks, as with Sajka and Rosiek’s (2015) eye-tracking study and Vantourout’s (2007) examination of teacher education students’ approaches to the assessment of fictitious students problem solutions. Two studies offered more conventional analyses of teacher education student or school student responses to mathematical tasks respectively (Kuzniak, Parzysz, Santos-Trigo & Vivier, 2011; Navarra, Malara & Ambrus, 2010). Finally, two studies, both drawing on LPS data, offered quantitative summaries of differently focused data derived from codes applied to lesson transcripts (Clarke & Xu, 2010; Hommel & Clarke, 2015). Of course, alternative interpretations may not have presented many of the studies above as case studies, but our view is that the privileging of, typically, two groups for comparison created particular cases.

#### *Documentary analyses*

The second most common type of study presented to the working group exploited some form of documentary analysis. In broad terms these 12 studies drawing fell into three categories. The largest of these three groups, comprising six studies, focused on comparisons of different classroom focused texts. Five of these focused on comparisons of how the school textbooks of different countries facilitate students’ learning of different topics. These included an analysis of the opportunities to learn calculus in Finland and Ghana (Bofah & Hannula, 2011), proportion in middle school mathematics in of Portugal, Brazil, Spain, and USA (Ponte & Marques, 2007), fractions in Albania and Kosovo (Vula, Kingji–Kastrati & Podvorica, 2015), the history of mathematics in Cyprus and Greece (Xenofontos & Papadopoulos, 2015) and a summary of various studies involving England, France and Germany (Pepin, 2010).

The second group of studies comprised comparisons of curricular documents in various ways. These included the introduction of quadratic equations in the Caribbean, China, Turkey, and the USA (An, Mintos & Yigit, 2013), mathematical modeling in France and Spain (Cabassut & Ferrando, 2013) and the graphical calculator in the examinations of Denmark, Australia and the International Baccalaureate (Brown, 2007). The third group comprised two papers summarising

the history of the ‘New Math’ movement in Iceland, Denmark, Norway and Sweden (Bjarnadóttir, 2007) and France and Hungary (Gosztonyi, 2015). Finally, there was one paper focused on document-based comparison of two mathematics teaching and learning traditions – the metaphor and the Grundvorstellungen (Soto-Andrade & Reyes-Santander, 2011).

From an analytical perspective studies typically invoked a variety of approaches. Two exploited a project-derived framework for supporting quantitative analyses (An, Mintos & Yigit, 2013, Vula, Kingji–Kastrati & Podvorica, 2015). Four exploited external frameworks in different ways. For example, Cabassut and Ferrando (2013) presented an ATD-based qualitative analysis, while Ponte and Marques (2007) offered a categorical analysis based on PISA’s assessment framework. Two studies offered, essentially, categorical analyses based on frameworks drawn from elsewhere in the literature (Hemmi, Koljonen, Hoelgaard, Ahl & Ryve, 2013; Bofah & Hannula, 2011). Three studies offered qualitative analyses based on frameworks derived from the vocabulary of the analysed documents themselves (Bjarnadóttir, 2007; Brown, 2007; Gosztonyi, 2015), while just one invoked the traditions of constant comparison (Xenofontos & Papadopoulos, 2015).

### *Small surveys*

Nine studies were construed as small surveys. In this respect it is important to explain how we distinguished, on the one hand, between small surveys and large surveys and, on the other hand, small surveys and case study. In the former the distinction was based less on sample size, although this was not ignored, than on the source of the data analysed. Studies classed as large surveys drew their data from large scale assessments of achievement, whether national or international, while small surveys exploited instruments developed by the project team or similar. In the latter case, small surveys were construed as requiring sufficient data units to achieve thematic saturation (O’Reilly & Parker, 2013), which case studies typically do not seek to achieve. Three forms of small survey were identified, one reflecting the former case above and two the latter.

Of the nine studies in this broad categorisation, five were construed as small questionnaire surveys, of which three focused on teachers. Two of these examined different aspects of mathematics teachers’ beliefs. These included a survey of teachers’ beliefs about mathematical modelling undertaken in France, Germany, Hungary and Spain (Cabassut & Villette, 2011) and a study of teachers’ beliefs in Estonia, Latvia and Finland (Hannula, Lepik, Pipere & Tuohilampi, 2013). A third study adapted an American instrument to examine Norwegian teachers’

mathematical knowledge for teaching (Jakobsen, Fauskanger, Mosvold and Bjuland (2011). Of the remaining two small surveys, one study evaluated grade three students’ mathematics-related affect structures (Tuohilampi, Hannula, Giaconi, Laine & Näveri, 2013) and one exploited an atypical social media distributed survey to examine interested parties’ perspectives on mathematics education in South Africa and England (Joubert, 2015). Analyses were conducted differently according to both project goals and the nature of the instruments used.

Two studies were construed as interview surveys. On the one hand, Andrews (2010) exploited semi-structured interviews, which were subjected to a constant comparison analysis, to examine English and Hungarian teachers’ professional motivations. On the other hand, Jablonka (2013) analysed a representative sample of LPS interviews to examine student perspectives on boredom in the mathematics lessons of Germany, Hong Kong and the United States. Finally, in this section, one study was construed as a written small survey in that Turkish and Belgian teacher education students were invited to describe any metaphors they had for describing mathematics (Kiliç, 2011), while one study exploited a battery of project-developed test items to examine the impact of the number-name system on third grade French-speaking Belgian and Vietnamese students’ understanding of the role of zero in their understanding of number in various forms (Nguyen & Grégoire, 2013).

#### *Large surveys*

Three studies presented new analyses of large extant data sets. One of these, Branchetti, Ferretti, Lemmo, Maffia, Martignone, Matteucci and Mignani (2015), compared the same cohort of students on two iterations of the Italian national test with the aim of understanding their understanding of fractions. The other two studies, Törnroos (2007) and Bofah (2015) drew on data from PISA 2003 and TIMSS 2011 respectively to analyse the impact of non-mathematical factors on students’ mathematics performance in, on the one hand, Finland and Sweden and, on the other hand, five diverse African countries, Botswana, Ghana, South Africa and Tunisia. In all cases, data were analysed by means of appropriate statistics, the latter two studies in particular exploiting confirmatory factor analysis to identify the structural properties of the entities under scrutiny.

#### *Position papers*

The six studies categorised as position papers, unsurprisingly, followed similar formats in their drawing on literature to support the authors’ proposed arguments. However, their foci typically

fell into two distinct categories. On the one hand were discerned papers with a methodological emphasis for comparative mathematics education research, as with Cabassut (2007), Clarke (2013, 2015), Jablonka (2015), Knipping and Müller-Hill (2013). All such papers offered significant insights into the conduct of comparative mathematics education research. On the other hand, a single study (Xenofontos, 2010) drew on literature to propose a particular agenda for future research on mathematical problem solving.

### 3.2 Theories and analytical framings

As comparative studies are dependent on various theoretical interests and changing methodological standards established in research in general, delineation will always be in flux. Yet, the view was shared that in studies that include comparison across educational systems there are two branches, one deriving from documentation, description and analysis of classroom and school practices, and another stemming from large-scale international achievement surveys. As is apparent from the overview of methodologies, researchers affiliated with the first tradition constituted a clear majority in the working group. However, it was noted that within this group studies with a broader conception of the enacted mathematics curriculum, which aim at illuminating issues of power, identity and subjectivity, were still rare (Jablonka, 2015).

Members of the group employed and developed a wide range of theories and analytical framings. These included literature based frameworks for the teaching/ learning of particular mathematical topics (An, Mintos, & Yigit, 2013; Andrews, 2011; Andrews, Sayers, & Marschall, 2015), mathematical knowledge for teaching (Jakobsen, Fauskanger, Mosvold & Bjuland, 2011), conceptual metaphors and Grundvorstellungen (Kiliç, 2011; Soto-Andrade, Reyes-Santander, 2011), cognitive psychology and theories of mathematical abstraction (Schäfer & Winkler, 2011), social cognitive theory (Mulat & Arcavi, 2010), particular theories from educational psychology (Adu-tutu Bofah, 2015; Andrews, 2010; Tuohilampi, Hannula, Giaconi, Laine, & Näveri, 2013; Vollstedt, 2007; Sarri, 2010), Anthropological Theory of Didactics (Cabassut & Ferrando, 2013; Gosztonyi, 2015; Larson & Bergsten, 2013; Maréchal, 2011), Theory of Didactic Situations (Gosztonyi, 2015; Kuzniak, Parzysz, Santos-Trigo, & Vivier, 2011), Neo-Vygotskian theories (Hemmi, Koljonen, Hoelgaard, Ahl, & Ryve, 2013; Pepin, 2010), theory of speech acts (Schäfer & Winkler, 2011), Bernstein's theory of pedagogic discourse (Larson & Bergsten, 2013), and semiotics (Branchetti et al., 2015). Some of the studies aimed at theory-networking or complementary analysis (Jablonka, 2013; Larson & Bergsten, 2013; Xenofontos, 2010).

The diversity of approaches and theoretical lenses aroused interest in the role of cultural origins of didactical principles and related issues pertaining to conceptual equivalence of categories used for comparative analyses. For instance, a couple of studies (e.g. Kiliç, 2011; Soto-Andrade & Reyes-Santander, 2011) triggered a discussion of the role of metaphors in a culture. Group members suggested to ask whether there are structural differences in the metaphors used in different contexts. Given the unavoidable cultural authorship of many theoretical frameworks, awareness of blind spots as well as of the danger of imposing categories and taxonomies that arise in one curriculum and teaching tradition onto another was recognised as important in order to overcome unquestioned assumptions of a shared view on the nature of the subject being compared. A more general issue examined in this context was whether we can successfully use a ‘home-grown’ theory from one culture to ask questions about another. The use of the French Anthropological Theory of Didactics in settings with different institutional layers and curriculum traditions or the appropriateness of tools developed in the West for use in the Global South served as examples.

Most group members agreed that comparisons are undertaken between constructed representations, whose structure and attributes are reflective as much of the value system of the researcher as of the objects being represented for the purposes of comparison. Any cross-cultural comparative analysis faces the challenge of honouring the separate cultural contexts, while employing an analytical frame that affords reasonable comparison. In particular, in an international comparative study, any evaluative aspect is reflective of the cultural authorship of the study. More specifically, seven “dilemmas” were identified by Clarke (2013) to reveal some of the contingencies under which international comparative research might be undertaken. Each dilemma can also serve as an interrogatory instrument: a tool directing the researcher’s attention to salient characteristics that, while presenting impediments to comparison, simultaneously provide insight into nuances of meaning and practice.

#### **4. Trends and Directions**

At each CERME, the group identified a number of practical challenges, methodological issues and complexities that comparative studies need to address, and shared experiences how this can be done. An increase in reflective awareness of methodological and theoretical issues is evidenced in the number of position papers with methodological and theoretical contributions submitted to the group. It was concluded that comparison can assist in basic, fundamental research that opens up new issues and that the role and potential of comparative approaches in

theory building needs to be explored further. It proved useful to consider comparative research from the perspective of boundary crossing. Recognition of the significance of acts of comparison in both boundary crossing and boundary construction foregrounds comparison as a key tool in the essential act of boundary deconstruction.

Notions of culture were examined at all meetings with the initial aim of developing a more differentiated view of culture than evidenced in large-scale surveys. Group members contended that there are different levels and dimensions of culture, and limitations of ‘zooming in’ or ‘zooming out’ were highlighted. In the case of employing psychological constructs, a potential ‘ecological fallacy’ by taking aggregated scores of individuals to represent a culture’s score was acknowledged. Further, it was concluded that cultural homogeneity cannot be assumed, even if there are shared political institutions. Hence it was suggested to employ notions of identity rather than talking about cultural affiliation, in recognition of individuals with multiple identities.

A major concern at all CERMEs was the development of culturally sensitive research that provides valid comparison. Rather than developing a set of fixed criteria, conceptualising methodological decisions in terms of balancing tensions was seen as useful. Some unavoidable ‘dilemmas’ in comparative studies were discussed in more depth, such as the advantages or disadvantages of exploiting too inclusive or too distinctive categories in the pursuit of valid comparisons. A study may risk to sacrifice validity in the interest of comparability (equivalence of constructs assumed) or limit comparability by the use of culturally-specific categories for cross-cultural description or evaluation. It was also noted that when imported instruments fail to be reliable in new contexts, this is an important research outcome. Openness towards the empirical field was seen as a condition for allowing the voice in the new context to be heard.

Further, it was acknowledged that all data can be seen as constructed rather than given, and that in cross-cultural research criteria for what constitutes data must hold legitimately across settings. Although matching an appropriate question to an appropriate unit of analysis is a major decision in all research studies, the group concluded that the choices are more visible in comparative studies. Consequently, methodological questions relating to the unit of analysis in comparative studies deserve more discussion, in particular with respect to appropriate theoretical framings for different levels or units of comparison. Another direction to be pursued arises from the need to better understand power relations between participants during data collection, and to acknowledge culturally and micro-culturally constructed interpersonal relationships, as for

example in characterising and comparing classroom instruction from the vantage points of apparently ‘neutral’ observer, teacher practice and learners’ experience.

A question resulting from the diversity of systemic contexts chosen for comparisons pertained to the tension between ‘cultural proximity’ and ‘cultural distance’ of the researchers involved. Generally, the group acknowledged the value of both, the insider’s and outsider’s perspectives, and noted that insiders share blind spots due to cultural blindness. Also, a distinction between form and function proved productive; while insiders can easily recognise the function of events in a culture, outsiders first of all attend to the form and might be more open to alternative interpretations. Similarly, the group also considered examples in which the use of theoretical frameworks linked to the researcher’s cultural heritage is a strength and examples where it is a limitation that could be overcome by cooperation and exchange between researchers, such as through the group meetings at the CERME.

## References

- An, T, Mintos, A., & Yagit, M. (2013). A cross-national standards analysis: quadratic equations and functions. *CERME8* (pp. 1825-1834).
- Andrews, P. (2010). Comparing Hungarian and English mathematics teachers’ professional motivations. *CERME6* (pp. 2452-2462).
- Andrews, P. (2011). The teaching of linear equations: Comparing effective teachers from three high achieving European countries. *CERME7* (pp. 1555-1564).
- Andrews, P., Sayers, J. & Marschall, G. (2015). Developing foundational number sense: Number line examples from Poland and Russia. *CERME9* (pp. 1681-1687)
- Back, J., Sayers, J., & Andrews, P. (2013). The development of foundational number sense in England and Hungary: a case study comparison. *CERME8* (pp. 1735-1844).
- Bjarnadottir, K. (2007). Development of the mathematics education system in Iceland in the 1960s in comparison to three neighbouring countries. *CERME6* (pp. 2403-2412).
- Bofah, E. A. (2015). Reciprocal determinism between students’ mathematics self-concept and achievement in an African context. *CERME9* (pp. 1688-1694).
- Bofah, E. A., & Hannula, M. (2011). The case of calculus: Comparative look at task representation in textbooks. *CERME7* (pp. 1545-1554).

- Branchetti, L., Ferretti, F., Lemmo, A., Maffia, A., Martignone, F. Matteucci, M., & Mignani, S. (2015). A longitudinal analysis of the Italian national standardized mathematics tests. *CERME9* (pp. 1695-1701).
- Brown, R. G. (2007). Policy change, graphing calculators and ‘High stakes examinations’: A view across three examination systems. *CERME6* (pp. 2413-2422).
- Cabassut, R. (2007). Examples of comparative methods in the teaching of mathematics in France and in Germany. *CERME6* (pp. 2423-2432).
- Cabassut, R., & Ferrando, I. (2013). Modelling in French and Spanish syllabus of secondary education. *CERME8* (pp. 1845-1854).
- Cabassut, R., & Villette, J-P. (2011). Exploratory data analysis of a european teacher training course on modelling. *CERME7* (pp. 1565-1574).
- Clarke, D. (2013). The validity-comparability compromise in crosscultural studies in mathematics education. *CERME8* (pp. 1855-1864).
- Clarke, D., & Xu, L. H. (2010). Spoken mathematics as a distinguishing characteristic of mathematics classrooms in different countries. *CERME6* (pp. 2463-2472).
- Clarke, D. (2015). The role of comparison in the construction and deconstruction of boundaries. *CERME9* (pp. 1702-1708).
- Ponte, J. P. & Marques, S. (2007). Proportion in school mathematics textbooks: A comparative study. *CERME6* (pp. 2443-2452).
- Eisenmann, T., & Even, R. (2007). Types of algebraic activities in two classes taught by the same teacher. *CERME6* (pp. 2433-2442).
- Gosztonyi, K. (2015). The ‘New Math’ reform and pedagogical flows in Hungarian and French mathematics education. *CERME9* (pp. 1709-1716).
- Hannula, M., Lepik, M., Pipere, A., & Tuohilampi, L. (2013). Mathematics teachers’ beliefs in Estonia, Latvia and Finland. *CERME8* (pp. 1865-1874).
- Maréchal, C. (2011). What kind of teaching in different types of classes? *CERME 7* (pp.1605-1613)
- Mulat, T., & Arcavi, A. (2010). Mathematical behaviors of successful students from a challenged ethnic minority. *CERME 6* (pp. 2473-2482).

- Navarra, G., Malara, N. A., & Ambrus, A. (2010). A problem posed by J. Mason as a starting point for a Hungarian-Italian teaching experiment within a European project. *CERME 6* (pp. 2483-2494).
- Nguyen, H. T. T., & Gregoire, J. (2013). Re-examining the language supports for children's mathematical understanding: A comparative study between French and Vietnamese languages. *CERME 8* (pp. 1925-1934).
- Nilsen, H. K. (2010). A comparison of teachers' beliefs and practices in mathematics teaching at lower secondary and upper secondary school. *CERME 6* (pp. 2494-2503).
- O'Reilly, M., & Parker, N. (2013). 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research, 13*(2), 190-197. doi: 10.1177/1468794112446106
- Peng, A., Sollervall, H., Stadler, E., Shang, Y., & Ma, L. (2015). Swedish and Chinese teachers' views on what constitutes a good mathematical test task: a pilot study. *CERME 9* (pp. 1738-1744).
- Pepin, B. (2010). Mathematical tasks and learner dispositions: a comparative perspective. *CERME 6* (pp. 2504-2512).
- Saari, J. v. R. (2010). Elite mathematics students in Finland and Washington: access, collaboration, and hierarchy. *CERME 6* (pp. 2513-2522).
- Saeki, A., Matszaki, A., Kawakami, T., & Lamb, J. (2015). Examining the heart of the dual modelling cycle: Japanese and Australian students advance this approach. *CERME 9* (pp. 1745-1751).
- Saika, M., & Rosiek, R. (2015). Solving a problem by students with different mathematical abilities: A comparative study using eye-tracking. *CERME 9* (pp. 1752-1758).
- Schäfer, I., & Winkler, A. (2011). Comparing the construction of mathematical knowledge between low-achieving and high-achieving students – a case study. *CERME 7* (pp. 1614-1624).
- Soto-Andrade, J., & Reyes-Santander, P. (2011). Conceptual metaphors and Grundvorstellungen: a case of convergence? *CERME 7* (pp. 1625-1635).
- Törnroos, J. (2007). Factors related to students' mathematical literacy in Finland and Sweden. *CERME 5* (pp. 2453-2462).

- Tuohilampi, L., Hannula, M., Giaconi, V., Laine, A., & Näveri, L. (2013). Comparing the structures of 3rd graders' mathematics-related affect in Finland and Chile. *CERME 8* (pp. 1935-1944).
- Vantourout, M. (2007). A comparative study of assessment activity involving 8 pre-service teachers: What referent for the assessor? *CERME 5* (pp. 2463-2472).
- Vollstedt, M. (2007). The construction of personal meaning – A comparative case study in Hong Kong and Germany. *CERME 5* (pp. 2473-2482).
- Vula, E., Kingji-Kastrati, J., & Podvorica, F. (2015). A comparative analysis of mathematics textbooks from Kosovo and Albania based on the topic of fractions. *CERME 9* (pp. 1759-1765).
- Xenofontos, C., & Papadopoulos, C. (2015). The history of mathematics in the lower secondary textbook of Cyprus and Greece: Developing a common analytical framework *CERME 9* (pp. 1766-1773).
- Xenofontos, C. (2010). International comparative research on mathematical problem solving: Suggestions for new research directions. *CERME 6* (pp. 2523-2532).