

Comparative studies of mathematics teaching: does the means of analysis determine the outcome?

Paul Andrews · Judy Sayers

Accepted: 29 November 2012 / Published online: 11 December 2012
© FIZ Karlsruhe 2012

Abstract This paper addresses four questions concerning the influence of culture on mathematics teachers' professional practice. Firstly, drawing on categorical data yielded by the application of low inference coding schedule to video recordings of sequences of lessons taught by case study teachers on four common topics in England, Flanders, Hungary and Spain, we undertook an exploratory factor analysis to examine the ways in which such coded variables interact. This process yielded five factors, each of which was interpretable against the literature and highlighted the extent to which dichotomisations of mathematics teaching as reform or traditional are not necessarily helpful, not least because all project teachers exhibited characteristics of both. Secondly, factors scores were analysed by nationality to reveal culturally located practices resonant with the available literature. Thirdly, cluster analyses yielded four well-defined cross-cultural clusters of episodes, each indicative of particular didactical perspectives that appeared to challenge the exclusivity of these culturally located practices. Finally, the key methodological finding was that the manner in which data are analysed influences greatly the outcomes of comparative mathematics research.

Keywords Comparative mathematics education · Cultural patterns of mathematics teaching · England · Flanders · Hungary · Spain

P. Andrews (✉)
Faculty of Education, University of Cambridge,
184 Hills Road, Cambridge CB2 8PQ, UK
e-mail: pra23@cam.ac.uk

J. Sayers
School of Education, University of Northampton,
Boughton Green Road, Northampton NN2 7AL, UK

1 Introduction

For many years researchers have been investigating the ways in which teachers present mathematics in different cultural contexts. Much of this research, which has highlighted both similarities and differences in the learning opportunities teachers offer their students, has attended to the national mathematics teaching script; that is, the culturally determined patterns of belief and behaviour, frequently beneath articulation, that distinguish one set of teachers from their culturally different colleagues. Typically, such projects have exploited coding schedules to highlight, through the presentation of frequencies, systemic emphases with regard to particular didactic strategies or mathematical outcomes. However, they rarely examine how such coded behaviours interact. Moreover, researchers have tended to use the nation as the framing construct and not consider how alternative analytical approaches may identify culture-independent typologies of teaching.

In this paper, construing a teacher's cultural context as reflecting the curricular traditions and expectations that govern his or her actions, we explore these issues by means of two analyses of categorical data derived from a video study of European mathematics teaching. In so doing we hope to make a significant methodological and substantive contribution to the field. Based on low-inference codes or "objective counts of discrete behaviors" (Evertson et al., 1980, p. 44), derived from sequences of lessons taught by four teachers in each of England, Flanders, Hungary and Spain, data were subjected to factor analyses to examine their interactions. The five factors that emerged, reflecting well defined and coherent perspectives on mathematics teaching, were then analysed with nationality as the framing variable to determine whether teachers in each cultural context enacted their roles in ways that might have

been predicted from the extant literature. Secondly, cluster analyses, also based on the five factors, were undertaken to determine whether there existed patterns of teacher behaviour that transcended such cultural boundaries or borders.

In undertaking our analyses we were conscious of the limitations of our data, particularly from the perspective of representativeness and generalisability. Consequently, we do not locate our analyses within notions of national teaching script but introduce two concepts to facilitate discussion of them. A *within culture practice* (WCP) is a pattern of practice common to a group of teachers working within a particular set of cultural norms. While not excluding the possibility that all teachers working within a particular cultural norm behave similarly, therefore adhering to what might otherwise be described as a national teaching script, WCP is not, of itself, a proxy for a national teaching script. Initially we construed WCP as *within system practice* because all teachers operate within particular sets of systemic expectations. However, *system* was rejected in favour of *culture* because a privileging of the former denied the impact of teacher agency, whereas an emphasis on culture reflected both systemic and individual constructions of practice. *Cross culture practice* (CCP) is construed as a pattern of practice found across cultural groups. Such a conception does not exclude the possibility of a CCP being an amalgam of two or more WCPs. That being said, we use the word *script* in our framing of the paper as that is the term frequently used in the literature.

2 Background

Several studies of the 1990s, particularly the Trends in International Mathematics and Science Study (TIMSS) video studies (Stigler et al. 1999; Hiebert et al. 2003) and the Survey of Mathematics and Science Opportunities (Schmidt et al. 1996), concluded that mathematics teaching, drawing on a subconscious routine and consistent re-enactment of particular pedagogies (Cogan and Schmidt 1999; Kawanaka et al. 1999), is culturally normative. That is, teachers of mathematics adhere, consciously or otherwise, to a culturally determined script. Motivations for investigating the script frequently, but not always, stem from the desire to understand the causes of the consistently high performance of East Asian students, when compared with their Western counterparts, on various international tests of achievement like the TIMSS, the Programme of International Student Assessment (PISA) and their repeats.¹ In this respect it is not surprising to find research showing how mathematics teaching in countries described as culturally east or Confucian differs from that of the culturally west or Socratic. Indeed, one study found that

every examined variable discriminated between the teaching of mathematics in Japan, Taiwan and the US (Stigler and Perry 1988), while Leung (2001) identified six features of mathematics classrooms that dichotomise the two regions in significant ways. However, differences in the mathematics didactics of such culturally diverse countries is not the focus of this paper, not least because researchers are increasingly aware that such broad categorisations mask significant differences between groups hitherto considered similar. For example, within the Chinese context, there is mathematics teaching variation across urban and rural regions (Ma et al. 2004). Also, Confucianism has been so mediated in China by Mohism, Daoism and Buddhism (Wong 2008) and in Taiwan by Buddhism (Leu 2005) that the existence of a national script has been questioned (Wong 2009). Similarly, within the apparently similar Socratic traditions of the West, developments in the Protestant and Catholic churches influenced educational developments in England and France (Sharpe 1997) and societal emphases on the individual, community and nation, respectively, informed educational practices and expectations in England, Denmark and France (Osborn 2004). In short, dichotomisations of the Confucian/Socratic form may fail to account for cultural variation within these broad categorisations.

Many cross-national studies of mathematics education have exploited coding schedules developed to facilitate an understanding of the similarities and differences with respect to the ways in which teachers conceptualise and present mathematics to their students. However, developing such tools is not straightforward, not least because decisions concerning examined variables influence greatly a project's outcomes; if they are too specific then every examined dimension differentiates between cultures (Stigler and Perry 1988), while if they are too broad then studies tend to show little cross-national variation, as with the studies of Anderson et al. (1989) and LeTendre et al. (2001). This problem was exemplified in our study of three teachers', one from each of Finland, Flanders and Hungary, presentations of linear equations (Andrews and Sayers, 2012). We found, at a macro level, each teacher's sequence of lessons passing through the same four phases. However, at a micro level, deep-seated differences emerged with respect to how each teacher construed and presented mathematics. Such matters lead us to suggest, in addition to the issue of examined variables, that decisions with respect to the mode of analysis may significantly alter the conclusions derived. That being said, even when selected variables allow for both similarities and differences to

¹ For TIMSS, the Trends in International Mathematics and Science Study, see <http://timss.bc.edu/> and for PISA, the Programme of International Student Assessment, see <http://www.oecd.org/pisa/>.

emerge, studies typically report frequencies and fail to examine the ways in which variables interact, as with, for example, the TIMSS video studies (Givvin et al. 2005; Hiebert et al. 2003, 2005; Stigler et al. 1999). Such interactions, we conjecture, are likely to be significant as students' mathematical competence, drawing on secure juxtaposition of conceptual knowledge, procedural skills and various dispositions (Kilpatrick et al. 2001), is unlikely to be achieved through simple didactical strategies exploited independently of other didactic strategies. Furthermore, researchers, particularly with respect to quantitative studies, have typically assumed the sufficiency of analyses by nationality and not asked questions concerning the uniqueness of the national scripts they have identified. For example, the TIMSS video studies presented many tables and charts, structured by nationality, showing similarities and differences in the mathematics privileged and the didactical strategies employed by project teachers. However, they did not ask, to what extent are the scripts unique to the cultures under scrutiny or, are there typologies of mathematics teaching scripts that transcend national boundaries?

Thus, acknowledging such ambiguity, our purpose in this paper, acknowledging the unrepresentative nature of our samples, is to dig deeper into these issues by addressing the following questions:

1. How do observed teacher behaviours, categorised by a low inference coding schedule, interact?
2. When categorised by the cultural context in which teachers work are there patterns of interaction indicative of within culture practices (WCP)?
3. When not categorised by the cultural context in which teachers work are there patterns of interaction indicative of cross culture practices (CCP)?

In seeking a resolution to these questions we hope, also, to address a fourth, essentially methodological, question; to what extent does the means of analysis determine the outcomes of comparative studies of mathematics teaching?

3 Methods

The European Union-funded Mathematics Education Traditions of Europe (METE) project set out to examine how four case study teachers in each of Flanders, England, Finland, Hungary and Spain, each of whom was construed against local criteria as competent, conceptualised and presented mathematics to students in the age range 10–14. Participating teachers were videotaped over four or five successive lessons taught on topics agreed by project colleagues as representative of their curricula. The topics were, for students in grades 5 or 6, percentages and

polygons, and for students in grades 7 or 8, polygons again and linear equations.

These video-recorded lessons were coded against an analytical framework developed in the year prior to the main study's data collection. The development of this framework entailed a week of live observations in each project country. Each morning one lesson was observed by at least one member from each project team and simultaneously video-recorded. These teachers, who were known to project members as being locally respected and amenable to a group of up to eight observers sitting at the back of their classrooms, were not involved in the subsequent video study. During the afternoons, facilitated by home colleagues' linguistic and contextual support, discussions focused, in the first instance, on constructing an accurate account of the lesson and, in the second, discussing descriptive categories of classroom activity to be tested against both subsequent and previous lessons. This constant comparison process led to a set of generic categories—seven learning objectives and ten didactic strategies—against which all subsequent video-recorded lessons were coded (Andrews 2007). Working definitions of the seven generic learning objectives and the ten generic didactic strategies can be seen in Tables 1 and 2, respectively.

A key element of this development was the identification of the unit of analysis. Unlike the TIMSS video studies, which exploited time-determined codes (Stigler et al. 2000; Givvin et al. 2005) the METE team adopted the episode. Here an episode, defined as that period of a lesson where the teacher's observed didactical intent remained constant, retained the structure, and therefore the unique characteristics, of an individual lesson. Each episode was coded 1 for each category observed and 0 for each category missing. Thus, most episodes were multiply coded.

Videographers were instructed to capture all teachers' utterances and as much board-work as possible. Teachers wore radio microphones while telescopic microphones were unobtrusively placed to capture student talk. Video files were compressed for later sharing and coding. Importantly, the first two lessons in each sequence were transcribed and translated into English, allowing colleagues to code other countries' lessons to evaluate inter-coder reliability. Cohen Kappa coefficients, which account for agreements due to chance, showed acceptable inter-coder reliability between the coders of England and Flanders ($\kappa = 0.877$), England and Hungary ($\kappa = 0.875$) and England and Spain ($\kappa = 0.793$). For reasons beyond our control, the Finnish data were never coded.

Importantly, when such small numbers of teachers are involved in a study it would be inappropriate to make any claims with respect to "national typification of practice", although "any regularities of practice...demand some consideration of possible causes" (Clarke et al. 2006, p. 3).

Table 1 Working definitions of the seven inferable generic learning objectives

Conceptual knowledge	The teacher is seen to encourage the conceptual development of his or her students
Derived knowledge	The teacher is seen to encourage the development of new mathematical entities from existing knowledge
Structural knowledge	The teacher is seen to encourage links or connections between different mathematical entities; concepts, properties and so on
Procedural knowledge	The teacher is seen to encourage the acquisition of skills, procedures, techniques or algorithms
Mathematical efficiency	The teacher is seen to encourage learners' understanding or acquisition of processes or techniques that develop flexibility, elegance or critical comparison of working
Problem-solving	The teacher is seen to encourage learners' engagement with the solution processes of non-trivial or non-routine tasks
Mathematical reasoning	The teacher is seen to encourage learners' development and articulation of justification and argumentation

Table 2 Working definitions of the ten generic didactic strategies

Activating prior knowledge	The teacher explicitly focuses learners' attention on mathematical content covered earlier in their careers in the form of a period of revision as preparation for activities to follow
Exercising prior knowledge	The teacher explicitly focuses learners' attention on mathematical content covered earlier in their careers in the form of a period of revision unrelated to any activities that follow
Explaining	The teacher explicitly explains an idea or solution. This may include demonstration, explicit telling or the pedagogic modelling of higher level thinking. In such instances the teacher is the informer with little or no student input
Sharing	The teacher explicitly engages learners in a process of public sharing of ideas, solutions or answers. This may include whole-class discussions in which the teacher's role is one of manager rather than explicit informer
Exploring	The teacher explicitly engages learners in an activity, which is not teacher directed, from which a new mathematical idea is explicitly intended to emerge. Typically this activity could be an investigation or a sequence of structured problems, but in all cases learners are expected to articulate their findings
Coaching	The teacher explicitly offers hints, prompts or feedback to facilitate their understanding of or abilities to undertake tasks or to correct errors or misunderstandings
Assessing	The teacher explicitly assesses or evaluates learners' responses to determine how the lesson proceeds
Motivating	The teacher, through actions beyond those of mere personality, explicitly addresses learners' attitudes, beliefs or emotional responses towards mathematics
Questioning	The teacher explicitly uses a sequence of questions, perhaps Socratic, which lead pupils to build up new mathematical ideas or clarify or refine existing ones
Differentiation	The teacher explicitly attempts to treat students differently in terms of the kind of tasks or activities, the kind of materials provided, and-or the kind of expected outcome in order to make instruction optimally adapted to the learners' characteristics and needs

In the case of the METE project, all teachers were selected, in the manner of the Learner's Perspective Study (Clarke et al. 2006), on the basis of local criteria of competence. For example, all were involved in teacher education activities with project universities. Two of the Flemings had been videotaped as models of good practice for use in a government-funded teacher education initiative, while two of the Hungarians had extensive experience of contributing to that country's well known tradition of textbook writing. Thus, should any patterns of interaction emerge from analyses of teachers working within the same cultural context then it would not only be pertinent to seek an explanation, in this case by examining the literature with respect to mathematics teaching in that cultural context, but also reasonable to assume the existence of some form of

WCP pertaining to those teachers working within that cultural context. That being said, this study is less concerned with confirming the existence of WCPs than it is with highlighting the consequences of different approaches to analysis.

4 Results

4.1 How do observed teacher behaviours, categorised by a low inference coding schedule interact?

In the following we exploit an exploratory factor analysis (EFA) to examine the interactive patterns in our lesson observation data. While not widely used, and depending on

whether or not researchers aim to test theoretically derived models, both confirmatory and exploratory approaches have been undertaken on data derived from observation schedules. For example, with respect to the former, several studies have performed confirmatory factor analyses on data derived from adaptations of the Classroom Assessment Scoring System (La Paro et al. 2004), including UK secondary trainee teachers (Malmberg et al. 2010) and Finnish kindergarten teachers (Pakarinen et al. 2010). These have shown it to be an appropriate tool for confirming a priori theorisations of classroom quality as measured by rating scales.

However, with respect to the METE data, there were no theoretical expectations with regards to the interactions of the learning outcomes and didactic strategies. In such circumstances, where researchers are seeking to uncover relationships between variables, exploratory factor analyses (EFA) in general and principal components analyses in particular are appropriate (Fabrigar et al. 1999; Henson and Roberts 2006). With respect to classroom observation rating scales, EFA have been exploited in a number of kindergarten-related studies, including an examination of children's play with particular toys (Trawick-Smith et al. 2011) and the quality of their classrooms (La Paro et al. 2004; Lambert et al. 2008; Jeon et al. 2010). Of significance to this study is the growing evidence that data derived from observational rating scales, even in dichotomous forms, can be subjected to EFA (Jeon et al. 2010).

Before factor analyses were undertaken, due to their infrequency across all 16 lesson sequences, 4 items were removed from the analysis. These were, from the generic learning outcomes, *derived knowledge* and, from the didactic strategies, *exercising prior knowledge*, *exploring* and *differentiation*. After this a principal components, with varimax rotation, EFA was performed on the data set derived from all coded episodes. The Kaiser Criterion and scree tests, procedures used to determine the appropriate number of factors to be extracted, indicated a five factor solution, details of which can be seen in Table 3. Importantly, a decision was made to accept factor loadings, or the correlation between the individual item and the underlying construct, in excess of 0.447 as this would ensure that the factor would account for at least 20 % of the item's variance. Each factor, as we discuss below, is construable as an orientation towards mathematics teaching that facilitates cross-cultural analysis.

4.1.1 Interpreting the factors

Firstly, it interesting to note that both learning outcomes and didactic strategies combined in three of the five factors, indicating that teaching is typically a juxtaposition of what is to be taught and how it is to be taught. Secondly, all five

factors were amenable to straightforward interpretation, although the latter three, due to their comprising two items in opposition, required some further explication by way of making them operational for subsequent analysis. In the following we discuss each factor in turn.

The first factor draws on didactic strategies, sharing and questioning, which encourage high level student participation in that sharing involves students in presenting publicly their solution processes and questioning encourages students to engage in a public development of mathematical ideas. Thus, embedded in the factor are expectations of high levels of participation. The three learning objectives implicated in the factor, reasoning, efficiency and structural knowledge, represent learning outcomes beyond the mere acquisition of concepts and procedures. The invocation of students to reason resonates well with the participative acts of questioning and sharing, as does the encouragement of students to engage with mathematical efficiency. Also, structural knowledge, reflecting relationships between concepts and topics is a higher order learning outcome resonant with notions of relational understanding as described by Skemp (1976). Finally, the characteristics of this factor reflect well current perspectives on reform mathematics teaching, whereby learners actively construct rather than passively receive knowledge (Cady et al. 2007; Peterson et al. 1989) and, "through interaction, are able to challenge one another's constructions in ways that facilitate the construction of increasingly shared and powerful knowledge" (Beswick 2005, p. 43) in classrooms that emphasise the making of mathematical connections (Cady et al. 2007; Peterson et al. 1989; Spillane and Zeuli 1999). In sum, we construe this factor as a relational participation. A factor score of one would represent an episode in which all five elements were present, while a score of zero would reflect an episode in which none were present.

The second factor, drawing on two didactic strategies, coaching and motivating, alongside procedural knowledge, is a publicly managed encouragement of, in Skemp's (1976) terms, instrumental learning. Coaching, which is a public activity focused on students completing tasks correctly, also resonates well with the acquisition of procedural competence. The exploitation of these two didactic strategies in the development of students' procedural knowledge matches well-known description of traditional approaches to mathematics teaching, which emphasise "the teaching of procedures...with little attention paid to the development of concepts or the connections between their learnt procedures and the concepts that show why they work" (Hiebert 1999, p. 11), where "the teacher is in complete control and the students' only goal is to learn operations to get the right answer" (Stipek et al. 2001, p. 214). Thus, acknowledging the instrumental focus and

Table 3 Factor loadings

	Factor				
	1 Relational participation	2 Instrumental participation	3 Activity progression	4 Exploited knowledge	5 Privileged Knowledge
Sharing	0.675				
Reasoning	0.670				
Questioning	0.658				
Efficiency	0.604				
Structural knowledge	0.533				
Coaching		0.762			
Motivating		0.554			
Procedural knowledge		0.551			-0.455
Explaining			0.847		
Assessing			-0.667		
Activating				0.782	
Problem solving				-0.553	
Conceptual knowledge					0.783

the public nature of the didactic strategies, we construe this factor as an instrumental participation. A factor score of one represents an episode in which all three elements are present, while a score of zero represents an episode in which none are present.

The remaining three factors, each of which comprised only two items in opposition, require a little more by way of explication. Factors three and four are structured similarly around, essentially, exclusive activities. That is, when one was observed the other was essentially absent. In particular, the third factor, which drew on two didactics, explaining and assessing, was construed as facilitating an activity's progression. For example, when teachers explain, they typically have a clear objective in relation to moving a lesson forward. In similar vein, assessing, as construed by the METE team, is a strategy focused on evaluating a class' understanding with a view to making in-the-moment decisions with regard to how an activity should progress. In this instance, a score of one would represent an episode in which explaining was the means of activity progression, while a score of zero was indicative of assessing as the means of activity progression. Importantly, a score of 0.5, due to their being exclusive, would represent an episode in which neither was present.

In similar vein, the fourth factor, in its drawing on activating—the exploitation of prior knowledge in the development of the current topic—and problem solving—the exploitation of mathematical knowledge in the solution of non-routine problems—was construed as reflecting alternative perspectives on the exploitation of

mathematical knowledge. In this case a score of one would represent an episode privileging activating, while a score of zero would represent problem solving as the privileged form of knowledge exploitation. Importantly, a score of 0.5, due to their being exclusive, would represent an episode in which neither was present. This factor is particularly interesting in that both poles seem to reflect different perspectives on reform. On the one hand, the activation of prior knowledge, matches reform expectations that teachers take account of learners' prior knowledge when making decisions about lesson progression (Beswick 2005; Cady et al. 2007; Peterson et al. 1989). On the other hand, reform classrooms are also characterised by an emphasis on problem solving (Cady et al. 2007; Peterson et al. 1989; Spillane and Zeuli 1999). Such findings highlight the fact that effective mathematics teaching is a complex interaction of many factors and that these may not, depending on circumstances, be complementary.

The fifth factor, which we have called privileged knowledge, sets conceptual knowledge against procedural knowledge. This factor differs from the previous two in that there were, essentially, no episodes in which neither was observed but many in which both were observed. Consequently, a score of one represents an episode in which conceptual knowledge was privileged, a score of zero an episode in which procedural knowledge was privileged, while a score of 0.5 represents a simultaneous emphasis on both. Such juxtapositions emphasise the importance of both forms of knowledge in the learning of mathematics (Kilpatrick et al. 2001).

Thus, in response to our first question, the observable learning outcomes and didactic strategies examined by the METE team interact in interesting, interpretable and largely predictable ways. However, to understand the extent to which teachers, working within particular cultural norms, adhere to a WCP it was necessary to calculate some form of factor score. In circumstances when different factors comprise different numbers of items, the arithmetical mean is appropriate for untested and exploratory data (Hair et al. 2006) and retains the scale’s metric (Di Stefano et al. 2009). Consequently, factor scores were calculated for each episode and a mean calculated for all the episodes observed in each country. This country mean was then compared, by means of *t* tests, with the means derived from all episodes from remaining countries. The results of this process are shown in Table 4.

With respect to interpreting the means, the reader is reminded that for individual episodes, factor scores for both activity progression and knowledge exploitation will lie at either end of the scale. This was due to the fact that in such episodes, with very few exceptions, only one of the two codes represented in the factor was observed. However, when taken over all episodes, the mean highlights an overall *tendency* towards one or other of the two observed events. With respect to the privileged knowledge factor, an individual episode could be scored 0, 0.5 or 1, highlighting the common simultaneous observation of both procedural and conceptual knowledge. However, taken over all episodes, the collective mean also reflects a collective tendency to privilege one form or knowledge over another.

Table 4 Mean factor scores for each country’s episodes

	Flanders	England	Hungary	Spain	All
Relational participation	0.35	0.21 _c	0.62 _c	0.37	0.37
Instrumental participation	0.35 _c	0.40 _a	0.49	0.67 _c	0.46
Activity progression	0.66	0.69	0.61 _a	0.81 _c	0.69
Knowledge exploitation	0.58 _c	0.46	0.52	0.37 _b	0.49
Privileged knowledge	0.57	0.63	0.57	0.55	0.58

Significant differences from the project mean, as identified by *t* tests, are represented by subscripts a, b and c, representing probabilities $p < 0.05$, $p < 0.001$ and $p < 0.0005$, respectively

Table 5 Factor scores and associated interpretations

Factor score (<i>S</i>)	$S \leq 0.25$	$0.25 < S \leq 0.40$	$0.40 < S \leq 0.60$	$0.60 < S \leq 0.75$	$0.75 > S$
Interpretation	Very low	Low	Moderate	High	Very high

4.2 When categorised by the cultural context in which teachers work are there patterns of interaction indicative of WCPs?

When discussing the figures of Tables 4 and 6, due to their being structured similarly, it is important that we adopt a consistent interpretation. In this regard the conventions shown in Table 5, which we acknowledge are essentially arbitrary but we believe meaningful, have been adopted throughout.

The Flemish episodes exhibited low emphases on relational participation and instrumental participation, with the latter being significantly lower than the project mean. Activity progression was managed through explaining, while knowledge exploitation was manifested in a significantly higher dominance of activating. Finally, there was a slight privileging of conceptual knowledge.

The English episodes incorporated very low levels of relational participation alongside low levels of instrumental participation, both of which were significantly lower than the project mean. Activity progression was managed predominantly by explanation and, alongside a tendency to privilege conceptual knowledge over procedural knowledge, there was, essentially, an equal exploitation of activating and problem solving.

The Hungarian episodes presented a high emphasis on relational participation alongside moderate emphases on instrumental participation, which were progressed predominantly by significantly higher emphases on explaining. Neither problem solving nor activation of prior knowledge was exploited more than the other, while conceptual knowledge was privileged, albeit minimally, over procedural knowledge.

Finally, the Spanish episodes exhibited low emphases on relational participation but significantly higher emphases on instrumental participation. Activity progression was achieved through such significantly high levels of explanation that teacher assessment was a rarity. There was a clear, significant, exploitation of problem solving alongside a slight privileging of conceptual knowledge over procedural knowledge.

In sum, the account above alludes to patterns of practice common to the four competent teachers working in each of the cultural contexts examined in this study that simultaneously marks elements of their practice as similar and different from their colleagues elsewhere. We are unable to say whether these represent national characteristics, but can offer the conjecture that, for the four teachers in each

country, there is evidence of a WCP. We return to this below.

4.3 When not categorised by the cultural context in which teachers work are there patterns of interaction indicative of CCPs?

As indicated earlier, the sorts of analyses presented above, with teacher nationality as the discriminating variable, would typically have been discussed against existing literature and a conclusion reached that teachers behave according to culturally constructed norms. However, our analyses, while alluding to the existence of WCPs, say little about the extent to which such WCPs exclude cultural outsiders or, importantly, whether patterns of interaction indicative of CCPs exist. We address these issues by means of cluster analyses that will identify groups of episodes determined by their properties rather than their cultural locations.

From the perspective of this study, an approach that allows for a predetermined number of clusters, in this case four, is required to determine whether each cluster aligned with one of the WCP profiles identified above. Therefore, a K-means approach was adopted; to each initial cluster an episode is randomly assigned as the cluster centre. Each remaining episode is then assigned to the cluster to whose centre it is closest. The codes of all the episodes in each cluster are then averaged to form a new cluster centre. The process is then repeated with each episode being re-assigned according to which of the new centres it is now closest. This process is iterated until no further movement occurs (Carver et al. 2002; Maxwell et al. 2002; Watters 2002).

Table 6 shows the mean factor scores for each of the four clusters obtained. The episodes of the first cluster privileged procedural knowledge, exhibited low levels of relational participation, high instrumental participation and were progressed exclusively by explanation. There was no particular exploited knowledge. Taken together, the episodes in this cluster seem to present a coherent perspective

on a teacher-centred, but participative, development of students' procedural skills.

The episodes of the second cluster privileged conceptual knowledge, exhibited low levels of relational participation, very low incidences of instrumental participation and were progressed more by means of assessment than explanation. Activation was the exploited form of knowledge, leading to an interpretation of these episodes as focused on a teacher-presented, but cognisant of students' prior and current understanding, development of conceptual knowledge. The impetus in these episodes was the presentation of concepts, largely independent of other concepts or related procedures.

The episodes of the third cluster privileged procedural knowledge, exhibited low levels of relational participation, high levels of instrumental participation and was progressed more by assessment than explanation alongside a tendency towards problem solving as the exploited knowledge. Such episodes, focused on the development of procedural skills, were not independent of conceptual knowledge and publicly managed in ways that accounted for students' in-the-moment responses.

The episodes of the fourth cluster privileged conceptual knowledge to the exclusion of procedural knowledge, exhibited moderate relational participation, very low instrumental participation, and was progressed by explanation alongside a tendency towards problem solving as the exploited knowledge. They were progressed very much by teacher explanation alongside the collaborative elements of relational participation—sharing and questioning—as students make connections, reason, solve problems and examine them for elegance and efficiency. In short, the episodes of this cluster appear commensurate with typical reform-related practice.

Each of the four clusters presents a different and coherent perspective on the teaching of mathematics, privileging either conceptual or procedural knowledge alongside, typically, high or very low emphases on instrumental participation. Importantly, from the perspective of this particular study, if the WCPs were unique then each would be reflected in one of the clusters. That is, were the Flemish WCP to be independent of all others then one of the clusters would comprise mostly Flemish episodes and so on. In the following we examine this more closely.

4.3.1 The national breakdown of cluster episodes

The figures of Table 7 show the national composition, both frequencies and percentages, of the episodes in each cluster. A superficial glance at the distribution indicates that each cluster draws on the episodes of each country. However, a Chi-square test ($\chi^2 = 20.9$, $df = 9$, $p < 0.01$) indicated that the distribution was unlikely to be due to chance. In this respect it can be seen that while the Flemish

Table 6 Mean factor scores for each cluster

	Cluster			
	1	2	3	4
Relational participation	0.32	0.34	0.35	0.47
Instrumental participation	0.67	0.20	0.63	0.21
Activity progression	1.00	0.34	0.37	0.93
Knowledge exploitation	0.49	0.69	0.43	0.43
Privileged knowledge	0.35	0.88	0.29	0.93
Number of episodes	99	59	107	102

Table 7 Frequency (F) and percentage (%) of each country's episodes per cluster

Cluster	Flanders		England		Hungary		Spain		
	F	%	F	%	F	%	F	%	
1	29	25	35	34	25	34	18	24	107
2	25	22	16	16	13	18	4	5	58
3	28	24	23	22	15	20	33	45	99
4	33	29	29	28	21	28	19	26	102

Table 8 Distribution of individual teachers' episodes across clusters

	Cluster			
	1	2	3	4
Flanders ($\chi^2 = 14.8$, $df = 9$, $p = 0.097$)				
F1	10	6	5	6
F2	5	9	9	14
F3	3	6	7	10
F4	11	4	7	3
England ($\chi^2 = 49.5$, $df = 9$, $p < 0.0005$)				
E1	7	4	12	9
E2	4	8	0	13
E3	4	4	5	7
E4	20	0	6	0
Hungary ($\chi^2 = 23.7$, $df = 9$, $p = 0.005$)				
H1	4	3	3	8
H2	11	6	3	1
H3	3	4	3	10
H4	7	0	6	2
Spain ($\chi^2 = 39.3$, $df = 9$, $p < 0.0005$)				
S1	11	1	4	3
S2	3	3	4	11
S3	0	0	13	2
S4	4	0	12	3

episodes were, essentially, distributed equally across all four clusters, this was not the case for the other countries, particularly Spain. This supports a conclusion that while clusters 1 and 4 appear to comprise cross national consistency clusters 2 and 3 do not. These results are important for two reasons. They indicate that the WCPs may not be unique and, importantly, allude to the existence of CCPs, or typologies of classroom practice which all teachers, irrespective of nationality, exploit at different times.

4.3.2 The distribution of teachers' episodes across clusters

When examined at the level of the individual teacher, the figures of Table 8 highlight two findings of interest. Firstly, few teachers' episodes did not fall across all four clusters. Secondly, there was little consistency between the cluster-related distributions of teachers from the same country.

Indeed, the Chi-square tests indicate that the distributions of the English, Hungarian and Spanish episodes across the clusters varied significantly. For example, while it could be argued that the episode distributions of Hungarian teachers, H1 and H3 were largely indistinguishable, the cluster-related distributions of H2 and H4 were clearly different. Similarly, while the distributions of S3 and S4 may appear equivalent, those of S1 and S2 shared little similarity. The English data showed no similarity in the episode distributions of any two teachers. Interestingly, while it could be argued that F1 and F4 appeared similar, as did F2 and F3, the overall distribution of the Chi-square test indicated that Flemish episodes were not dissimilar in their distribution.

5 Discussion

Earlier we posed three questions. The first, concerning the interactions of six generic learning outcomes and seven didactic strategies, was addressed by means of an EFA. The five factors identified presented largely unproblematic interpretations; both the participative factors were interpretable against Skemp's (1976) forms of understanding and clearly related to current perspectives on mathematics teaching and learning, while *privileged knowledge* offered a reassuring highlighting of the emphases teachers place on conceptual and procedural knowledge. The remaining two factors presented hitherto unconsidered perspectives on teachers' actions in that *activity progression* offered a strong sense that when teachers assess they tend not to explain and vice versa. In similar vein, the nature of *knowledge exploitation* was in that when teachers engage their students in problem solving they tend not to publicly activate prior knowledge and vice versa.

Importantly, we regard the factors as providing something of a foil to much of the comparative literature, which has tended to dichotomise mathematics teaching as either reform or traditional in frequently unhelpful ways (Clarke 2006). For example, *relational participation* and *instrumental participation*, reflecting particular emphases in project teachers' classrooms, were observed in all but a handful of teachers' lessons. That is, teachers' objectives, and the means by which they were achieved, reflected frequent juxtapositions of, essentially, reform and traditional practices. Moreover, all teachers, whether attending to relational or instrumental objectives, exploited teacher-centred behaviours, confounding the popular notion that teacher-centred didactics are, of themselves, reflective of poor practice.

The second question, concerning the extent to which the five factors support the notion of WCP, led to some interesting and, what seemed to be, culturally located similarities and differences in the ways that teachers

undertake their professional responsibilities. However, while we desist from making any claims about teacher representativeness or didactic generality we think it is appropriate to examine each of the four WCPs against the literature pertaining to the teaching of mathematics in their respective countries. The practices observed in the Flemish classrooms resonate with earlier Flemish studies highlighting a typically transmissive teaching (Waeytens et al. 1997) that privileges lower-order procedural skills (Janssen et al. 2002) above those of problem solving (Yoshida et al. 1997). The fragmented and overly simplified experience observed in the English episodes resonates well with earlier research (Jennings and Dunne 1996; Kaiser et al. 2006), while the Hungarian emphases on participation in general and relational participation in particular accords well with research highlighting the systematic emphasis on collective activity focused on high level learning (Andrews 2003; Szalontai 2000; Szendrei and Török 2007). Finally, the dominance of instrumental participation and explaining as the means of lesson progression in the Spanish episodes find substantial resonance with Blanco's (2003, p. 6) observation that learning in Spanish classrooms is "directly proportional to the (teacher's) capacity to manage the class and to the skill in creating and sustaining a productive discourse in the classroom". In sum, the country-based analyses indicate not only that project teachers conform to particular WCPs but that these WCPs resonate with earlier studies of their respective countries.

The third question, addressed by cluster analyses, yielded four clearly articulable CCPs, none of which showed any close resonance with any of the WCPs. Moreover, each cluster drew on the episodes of teachers from all four countries, although the proportions differed. For example, clusters one and four drew consistently from all countries' episodes, while clusters two and three did not. In sum, each CCP focused explicitly on the development of either conceptual knowledge or procedural knowledge in didactically different ways. However, as this paper is concerned less with characterising the clusters than with their existence, and limitations of space prevent it, we do not discuss further the nature of these CCPs.

Interestingly, perhaps the most telling outcome of this particular analysis is the fact that the episodes of three-quarters of all the teachers were spread across all four clusters, indicating that teachers behave differently at different times according, we assume, to decisions made concerning particular objectives at particular times. In this respect, the nature of the four clusters, with their didactically different privileging of conceptual or procedural knowledge, seems to support such a conjecture. Also, although the variation of cluster membership across a teacher's episodes was substantial, two teachers from each country except England appeared to share similar patterns

of cluster-related episode distribution, indicating that both WCPs and CCPs have a role in describing the practices of teachers from each of the four countries.

In closing we return to our final question, to what extent does the means of analysis determine the outcomes of comparative studies of mathematics teaching? Our analyses offer two perspectives. Firstly, when teachers were discriminated by country—as a proxy for their professional cultural context—the data yielded well-defined WCPs within which they worked. Importantly, while we desist from generalising from such small samples, we note that the practices embedded in each WCP resonated with the practices of teachers working in the same cultural contexts identified in the available literature. Secondly, analyses of the four clusters indicated that the ways in which these cultural roles are enacted are not so isolated, so independent of the practices found elsewhere, as to be unique. Such findings allude to the need for further research, not least because of the ambiguity surrounding the construction of the various WCPs. In this respect, Andrews (2011) has written that all teachers operate within three curricula; the intended, the idealised and the received. The intended curriculum represents the mathematical knowledge and skills that the system in which a teacher operates deems appropriate. All teachers operate within one, although its definition and extent to which didactics are mandated vary from one culture to another. The idealised curriculum pertains to the experientially formed and personal beliefs about mathematics, its teaching and learning that teachers bring to their classrooms. It refers to those subject- and didactics-specific beliefs and behaviours that distinguish individual teachers from their colleagues. Finally, the received curriculum reflects those beliefs and practices that are so "deep in the background of the schooling process" and "so taken-for-granted...as to be beneath mention" (Hufton and Elliott 2000, p. 117). The received curriculum is typically formed in childhood and persists into adulthood as the natural way of doing things; it is a collectively developed, and rhetorically warranted, pedagogy. It is our conjecture that the core of any particular group's WCP lies in the intersection of the intended curriculum and the received curriculum. Variation lies in the plurality of idealised curricula that teachers within a culture bring to their work; the more a system promotes teacher individuality, the more likely there will be divergence from the core WCP. However, this is just conjecture and further research will be necessary if we are to understand more fully the ways that culture constructs the ways in which teachers think and act.

Acknowledgments The Mathematics Education Traditions of Europe (METE) project team acknowledges the financial support of the European Union, Socrates Action 6.1 programme, project code 2002-5048.

References

- Anderson, L., Ryan, D., & Shapiro, B. (1989). *The IEA classroom environment study*. Oxford: Pergamon.
- Andrews, P. (2003). Opportunities to learn in the Budapest mathematics classroom. *International Journal of Science and Mathematics Education, 1*(2), 201–225.
- 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. (2011). The cultural location of teachers' mathematical knowledge: Another hidden variable in mathematics education research? In T. Rowland & K. Ruthven (Eds.), *Mathematical knowledge in teaching* (pp. 99–118). New York: Springer.
- Andrews, P., & Sayers, J. (2012). Teaching linear equations: Case studies from Finland, Flanders and Hungary. *The Journal of Mathematical Behavior, 31*(4), 476–488.
- Beswick, K. (2005). The beliefs/practice connection in broadly defined contexts. *Mathematics Education Research Journal, 17*(2), 39–68.
- Blanco, L. (2003). The mathematical education of primary teachers in Spain. *International Journal of Mathematics Teaching and Learning*. <http://www.cimt.plymouth.ac.uk/journal/blanco1.pdf>. Accessed 19 Sept 2012.
- Cady, J., Meier, S., & Lubinski, C. (2007). The mathematical tale of two teachers: A longitudinal study relating mathematics instructional practices to level of intellectual development. *Mathematics Education Research Journal, 18*(1), 3–26.
- Carver, F., Fuchs, A., Jantzen, K., & Kelso, J. (2002). Spatiotemporal analysis of the neuromagnetic response to rhythmic auditory stimulation: rate dependence and transient to steady-state transition. *Clinical Neurophysiology, 113*, 1921–1931.
- Clarke, D. (2006). Using international research to contest prevalent oppositional dichotomies. *ZDM The International Journal on Mathematics Education, 38*(5), 376–387.
- Clarke, D., Keitel, C., & Shimizu, Y. (2006). The learner's perspective study. In D. Clarke, C. Keitel, & Y. Shimizu (Eds.), *Mathematics classrooms in twelve countries: The insider's perspectives* (pp. 1–14). Rotterdam: Sense Publishers.
- Cogan, L., & Schmidt, W. (1999). An examination of instructional practices in six countries. In G. Kaiser, E. Luna, & I. Huntley (Eds.), *International comparisons in mathematics education* (pp. 68–85). London: Falmer.
- Di Stefano, C., Zhu, M., & Míndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation, 14*(20), 1–11.
- Everton, C. M., Anderson, C. W., Anderson, L. M., & Brophy, J. E. (1980). Relationships between classroom behaviors and student outcomes in junior high mathematics and English classes. *American Educational Research Journal, 17*(1), 43–60.
- Fabrigar, L., Wegener, D., MacCallum, R., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272–299.
- Givvin, K., Hiebert, J., Jacobs, J., Hollingsworth, H., & Gallimore, R. (2005). Are there national patterns of teaching? Evidence from the TIMSS 1999 video study. *Comparative Education Review, 49*(3), 311–343.
- Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate data analysis* (6th ed.). New York: Prentice Hall.
- Henson, R., & Roberts, J. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement, 66*(3), 393–416.
- Hiebert, J. (1999). Relationships between research and the NCTM standards. *Journal for Research in Mathematics Education, 30*(1), 3–19.
- Hiebert, J., Gallimore, R., Garnier, H., Givvin, K., Hollingsworth, H., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study*. Washington: National Center for Educational Statistics.
- Hiebert, J., Stigler, J., Jacobs, J., Givvin, K., Garnier, H., et al. (2005). Mathematics teaching in the United States today (and tomorrow): Results from the TIMSS 1999 video study. *Educational Evaluation and Policy Analysis, 27*(2), 111–132.
- Huften, N., & Elliott, J. (2000). Motivation to learn: The pedagogical nexus in the Russian school: Some implications for transnational research and policy borrowing. *Educational Studies, 26*(1), 115–136.
- Janssen, R., De Corte, E., Verschaffel, L., Knoors, E., & Colemont, A. (2002). National assessment of new standards for mathematics in elementary education in Flanders. *Educational Research and Evaluation, 8*(2), 197–225.
- Jennings, S., & Dunne, R. (1996). A critical appraisal of the National Curriculum by comparison with the French experience. *Teaching Mathematics and its Applications, 15*(2), 49–55.
- Jeon, H.-J., Langill, C., Peterson, C., Luze, G., Carta, J., & Atwater, J. (2010). Children's individual experiences in early care and education: Relations with overall classroom quality and children's school readiness. *Early Education & Development, 21*(6), 912–939.
- Kaiser, G., Hino, K., & Knipping, C. (2006). Proposal for a framework to analyse mathematics education traditions in eastern and western traditions. In F. Leung, K.-D. Graf, & F. Lopez-Real (Eds.), *Mathematics education in different cultural traditions: A comparative study of East Asia and the West, the 13th ICMI study* (pp. 319–351). New York: Springer.
- Kawanaka, T., Stigler, J., & Hiebert, J. (1999). Studying mathematics classrooms in Germany, Japan and the United States: Lessons from the TIMSS videotape study. In G. Kaiser, E. Luna, & I. Huntley (Eds.), *International comparisons in mathematics education* (pp. 86–103). London: Falmer.
- Kilpatrick, J., Swafford, J., & Findell, B. (Eds.). (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: The National Academies Press.
- La Paro, K., Pianta, R., & Stuhlman, M. (2004). The Classroom Assessment Scoring System: Findings from the prekindergarten year. *The Elementary School Journal, 104*(5), 409–426.
- Lambert, M., Williams, S., Morrison, J., Samms-Vaughan, M., Mayfield, W., et al. (2008). Are the indicators for the language and reasoning subscale of the early childhood environment rating scale-revised psychometrically appropriate for Caribbean classrooms? *International Journal of Early Years Education, 16*(1), 41–60.
- LeTendre, G., Baker, D., Akiba, M., Goesling, B., & Wiseman, A. (2001). Teachers' work: Institutional isomorphism and cultural variation in the U.S. Germany, and Japan. *Educational Researcher, 30*(6), 3–15.
- Leu, Y.-C. (2005). The enactment and perception of mathematics pedagogical values in an elementary classroom: Buddhism, Confucianism, and curriculum reform. *International Journal of Science and Mathematics Education, 3*(2), 175–212.
- Leung, F. (2001). In search of an East Asian identity in mathematics education. *Educational Studies in Mathematics, 47*(1), 35–51.
- Ma, Y., Zhao, D., & Tuo, Z. (2004). Differences within communities: How is mathematics taught in rural and urban regions in mainland China? In L. Fan, N.-Y. Wong, J. Cai, & S. Li (Eds.), *How Chinese learn mathematics: Perspectives from insiders* (pp. 413–442). Singapore: World Scientific.
- Malmberg, L.-E., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2010). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology, 102*(4), 916–932.

- Maxwell, B., Pryor, F., & Smith, C. (2002). Cluster analysis in cross-cultural research. *World Cultures*, 13(1), 22–38.
- Osborn, M. (2004). New methodologies for comparative research? Establishing ‘constants’ and ‘contexts’ in educational experience. *Oxford Review of Education*, 30(2), 265–285.
- Pakarinen, E., Lerkkanen, M.-K., Poikkeus, A.-M., Kiuru, N., Siekkinen, M., et al. (2010). A validation of the Classroom Assessment Scoring System in Finnish kindergartens. *Early Education & Development*, 21(1), 95–124.
- Peterson, P., Fennema, E., Carpenter, T., & Loef, M. (1989). Teachers’ pedagogical content beliefs in mathematics. *Cognition and Instruction*, 6(1), 1–40.
- Schmidt, W., Jorde, D., Cogan, L., Barrier, E., Gonzalo, I., et al. (1996). *Characterizing pedagogical flow: An investigation of mathematics and science teaching in six countries*. Dordrecht: Kluwer.
- Sharpe, K. (1997). The Protestant ethic and the spirit of Catholicism: Ideological and institutional constraints on system change in English and French primary schooling. *Comparative Education*, 33(3), 329–348.
- Skemp, R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 20–26.
- Spillane, J., & Zeuli, J. (1999). Reform and teaching: Exploring patterns of practice in the context of national and state mathematics reforms. *Educational Evaluation and Policy Analysis*, 21(1), 1–27.
- 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.
- Stigler, J., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study*. Washington, DC: National Center for Educational Statistics.
- Stigler, J., & Perry, M. (1988). Cross cultural studies of mathematics teaching and learning: Recent findings and new directions. In D. A. Grouws, T. J. Cooney, & D. Jones (Eds.), *Perspectives on research on effective mathematics teaching*. Reston, VA: NCTM.
- Stipek, D., Givvin, K., Salmon, J., & MacGyvers, V. (2001). Teachers’ beliefs and practices related to mathematics instruction. *Teaching and Teacher Education*, 17(2), 213–226.
- Szalontai, T. (2000). Some facts and tendencies in Hungarian mathematics teaching. *International Journal of Mathematics Teaching and Learning*. <http://www.cimt.plymouth.ac.uk/journal/tshungmt.pdf>. Accessed 4 Dec 2012.
- Szendrei, J., & Török, J. (2007). The tradition and role of proof in mathematics education in Hungary. In P. Boero (Ed.), *Theorems in school* (pp. 117–134). Rotterdam: Sense Publishers.
- Trawick-Smith, J., Russell, H., & Swaminathan, S. (2011). Measuring the effects of toys on the problem-solving, creative and social behaviours of preschool children. *Early Child Development and Care*, 181(7), 909–927.
- Waeytens, K., Lens, W., & Vandenberghe, R. (1997). *Learning to learn: How do teachers differ?* Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Watters, P. (2002). Discriminating English word senses using cluster analysis. *Journal of Quantitative Linguistics*, 9(1), 77–86.
- Wong, N.-Y. (2008). Confucian heritage culture learner’s phenomenon: From “exploring the middle zone” to “constructing a bridge”. *ZDM The International Journal on Mathematics Education*, 40(6), 973–981.
- Wong, N.-Y. (2009). Exemplary mathematics lessons: What lessons we can learn from them? *ZDM The International Journal on Mathematics Education*, 41(3), 379–384.
- Yoshida, H., Verschaffel, L., & De Corte, E. (1997). Realistic considerations in solving problematic word problems: Do Japanese and Belgian children have the same difficulties? *Learning and Instruction*, 7(4), 329–338.