International comparative studies on mathematics education: a viewpoint from the anthropological theory of didactics

Artigue, Michèle; Winsløw, Carl

Published in:
Récherches en Didactiques des Mathématiques

Publication date:
2010

Document Version
Early version, also known as pre-print

Citation for published version (APA):
Artigue, M., & Winsløw, C. (2010). International comparative studies on mathematics education: a viewpoint from the anthropological theory of didactics. Récherches en Didactiques des Mathématiques, 30(1), 47-82.

Appeared in: Recherches en Didactique des Mathématiques vol. 31 no. 1 (2010), pp. 47-82

INTERNATIONAL COMPARATIVE STUDIES ON MATHEMATICS EDUCATION: A VIEWPOINT FROM THE ANTHROPOLOGICAL THEORY OF DIDACTICS

Michèle Artigue[*], Carl Winsløw[**]

COMPARAISONS INTERNATIONALES SUR L'ENSEIGNEMENT DES MATHÉMATIQUES : UN POINT DE VUE PORTÉ PAR LA THÉORIE ANTROPOLOGIQUE DU DIDACTIQUE

Résumé – Les études comparatives internationales visent à identifier et expliquer les différences entre phénomènes homologues dans plusieurs contextes. Elles sont menées avec une variété d'objectifs et de méthodes, et leurs résultats ainsi que les interprétations de ces derniers font l'objet de débats parfois vifs, surtout pour ce qui concerne les évaluations quantitatives à grande échelle telles que PISA. Même si l'évaluation des compétences des élèves n'est pas sa préoccupation majeure, la recherche didactique ne peut manquer d'être concernée par ces débats. Dans cet article, nous présentons d'abord un modèle théorique, dérivé de la théorie anthropologique du didactique, qui nous servira à spécifier les *niveaux* auxquels une comparaison est effectuée. En nous servant de ce modèle, nous proposons ensuite une analyse synthétique d'une sélection d'études comparatives internationales en éducation mathématique (allant d'évaluations à grande échelle à des comparaisons binaires à petite échelle menées dans le cadre de thèses), dans le but spécifique de comprendre comment ces études diffèrent et peuvent s'articuler.

Mots clés: études comparatives internationales, enseignement mathématique, apprentissage des mathématiques, niveaux de codétermination didactique.

[*] Laboratoire de Didactique André Revuz, Université Paris Diderot - Paris 7, France ; artigue@math.jussieu.fr

[**] Département de Didatictique des Sciences, Université de Copenhague, Danemark; winslow@ind.ku.dk

COMPARACIONES INTERNACIONALES EN EDUCACIÓN MATEMATICA: UN PUNTO DE VISTA DESDE LA TEORIA ANTROPOLOGICA DE LA DIDACTICA

Resumen – Los estudios comparativos internacionales tienen como meta la identificación y explicación de diferencias entre fenómenos análogos en contextos variados. Se desarrollan con una diversidad de objetivos y métodos, y sus resultados como la interpretación de aquellos siempre son fuente de debates feroces, especialmente cuando se trata de comparaciones cuantitativas a gran escala como por ejemplo PISA. Incluso si la evaluación de competencias de los alumnos no es su preocupación central, la investigación didáctica no puede desinteresarse a tales debates. En este artículo, proponemos primero un modelo teórico para clarificar los *niveles* en que la comparación se hace, basándonos en la teoría antropológica de la didáctica. Utilizando ese modelo, nos proponemos enseguida un análisis sintético de una selección de investigaciones comparativas en educación matemática (desde estudios a gran escala hasta comparaciones binarias de reducido tamaño desarrolladas en tesis doctorales) con el objetivo de mejor comprender en qué se diferencian estos estudios y cómo se pueden articular.

Palabras-claves: estudios comparativos internacionales, enseñanza matemática, aprendizaje matemático, niveles de codeterminación didáctica.

ABSTRACT

Comparative studies aim to identify and explain differences of homologous phenomena in two or more contexts. Comparative studies of mathematics teaching and learning are undertaken with a variety of purposes and methods, and their results and interpretations remain the subject of fierce debates, especially in the case of large-scale quantitative surveys such as PISA. Even if the measurement of student performance is not its central preoccupation, didactic research is certainly concerned with these debates. This paper first proposes a model for clarifying the *levels* at which comparison is done, based on the anthropological theory of didactics. Using this model, we then propose a synthetic analysis of a selected body of international comparative research on mathematics education (from large-scale surveys to small-scale binary comparisons in doctoral projects) with the particular aim of understanding how the different studies differ and relate.

Key words: international comparative studies, mathematics teaching, mathematics learning, levels of didactic codetermination.

INTRODUCTION

The last decades have seen the rise of interest—political as well as academic—in large international surveys of students' mathematical performance on meant-to-be neutral test items, which may or may not take into account the variability of emphases in the participating educational systems. At the same time, the internationalisation of research on mathematics education has led to collaborations and studies that involve perspectives and practices from otherwise very distant contexts of teaching. This, as much as the large-scale surveys, has led scholars to notice, and reflect on, differences. Investigating differences and their causes is one simple definition of what comparative studies are all about.

Comparative studies may be based on different assumptions—ranging from presumably universal models of mathematical competency to in-depth comparisons of local practices in two specific contexts, in which one pays the greatest attention to, for instance, the linguistic, cultural, societal, and intellectual specificities of the two environments. One therefore ends up with comparisons that, metaphorically, are based on everything from aerial photography to microscopy, and the need arises for a common framework that permits one to relate and integrate different purposes and methods of comparison, involving different—but not necessarily unrelated—objects of study.

Didactics is the science that takes, primarily, the teaching of disciplined knowledge as its object. Such a study may equally be more or less comprehensive, and more or less focused on local conditions and constraints. In the didactics of mathematics, several theoretical frameworks are available to organise and, indeed, make possible, different kinds and levels of research. Such research may well be (indeed, is very often) taking place within fairly homogenous cultural and institutional settings, and then those settings are often subject to less attention. It is clear that a didactical theory that could help inform and organise comparative studies would have to take a less naïve viewpoint on cultures and institutions. That is why in this paper we propose—and exemplify—the use of elements from the anthropological theory of didactics (ATD) founded by Chevallard and developed over the last decades with a number of other researchers, primarily from France and Spain. In ATD research, institutional and more widely cultural conditions are always crucial.

We first introduce some of the main elements of ATD and how we make use of them to consider different kinds of comparative research (in the next two sections). Then in the next four sections, we consider a variety of recent comparative studies of mathematics education in order to demonstrate how the viewpoint of ATD helps in articulating their results and, indeed, in understanding the differences and relations between them. Finally, in the last section, we discuss briefly some more general clusters of problems that emerge from the cases considered—in particular, certain gaps and shortcomings that, according to us, require further work but to which the viewpoint of ATD could bring important resources.

THE ANTHROPOLOGICAL THEORY OF DIDACTICS: SOME PRELIMINARIES

The anthropological theory of didactics (ATD) offers a bold vision of human activity. A brief outline is given here that focuses on the specific tools needed for this study; in particular, on the *levels of didactic codetermination*. A broader and more detailed exposition has been given by Chevallard (1999) and by Bosch and Gascòn (2006).

According to ATD, the minimal unit of human activity is a *practice block*, formed by a type of task and a technique. A complete *praxeology* contains in addition a corresponding *knowledge block*: a technology describing and explaining the practice block and a theory that allows systematic reasoning about (e.g., justifying) the practice block in relation to other blocks. It is apparent that praxeologies often occur in larger systems, sharing some of the same explicit elements (i.e., knowledge blocks). A technology may serve in the context of more than one practice block; a collection of praxeologies sharing the same technology (and theory) is called a *local organisation* (of praxeologies). Such an organisation is characterised by the common technology, such as the discourse pertaining to solving polynomial equations. Finally, a theory may serve to reason about several local organisations, which then are said to form a *regional organisation* (of praxeologies), characterised by a unifying theory, such as an algebraic theory for the solution of equations.

It is normally easy to associate school mathematical practices (as the example given) with knowledge blocks and to group these according to the shared technologies or theories (leading to reference models of the corresponding mathematical organisations). It may be less straightforward for other practices—such as teaching a given mathematical organisation—even if these practices contain definite types of tasks that are carried out using techniques. This form of praxeological organisation—which may of course be considered for the teaching of any praxeology—is called a *didactical organisation*.

Didactical and mathematical praxeologies and their development cannot be understood fully without considering broader contexts than those in which they directly occur. That is because they are codetermined (i.e., determined in their mutual interaction) by a whole hierarchy of institutional levels that successively condition and constrain each other. It is an important contribution of ATD to provide a detailed model of these levels of didactic codetermination, as shown in Figure 1 (together with simple examples; for further details, see Chevallard 2002). It is important to remember that we are talking about discipline, domain, and so on, as they are realised in the educational context. Many of the conditions of teaching practice, particularly those originating at the higher levels, cannot be changed by the individual teacher; some of them may be further modified by others, such as school principals, curriculum developers, or politicians—and by what Chevallard (1985) calls the *noosphere* of the educational system.

The higher levels cover many types of content, the most familiar grand types being what is commonly referred to as *disciplines*. The conditions and principles for teaching that concern several disciplines (individually and as they interact) are subsumed in *pedagogies*. Other conditions come from the institution of teaching (*school*), come from the *society* in which the school resides, or arise even from an even larger *civilisation* (culturally homogenous group of societies).

But ATD also allows us to identify a number of *subdisciplinary levels* of codetermination of the study of a discipline; these can to some extent be associated with the unifying elements of the praxeologies they determine. Specifically, a *domain* embraces a collection of regional organisations involving several theories forming a larger part of the discipline (such as algebra within mathematics, both to be understood within the institutional context of a school). A *sector* is characterised by the study of one regional organisation (or parts of it) that comes from a family of praxeologies sharing one theory. For instance, when we study the sector of polynomials within secondary school, we will encounter praxeological organisations unified by a theory that allows us to examine precise assertions about polynomials; for example, about their number of real-valued roots. Different technologies may be in use, from informal discourses in which polynomials appear only as equations determining a curve to formal ones in which they are algebraic elements with their own composition rules. That, in turn, gives rise to different *themes*, each unified by a technology (and thus each determining the study of a local organisation). Finally, a *subject* is concentrated on a type of tasks and technique, motivated and articulated within a larger theme. Notice that these levels of codetermination cannot be entirely determined by praxeological characteristics but are specific to institutionally defined aims of studying them. In particular, themes and subjects may be highly imbued with didactical traditions, such as the theme of Thales' theorem in the French middle school or the topic of long division in most primary school systems.

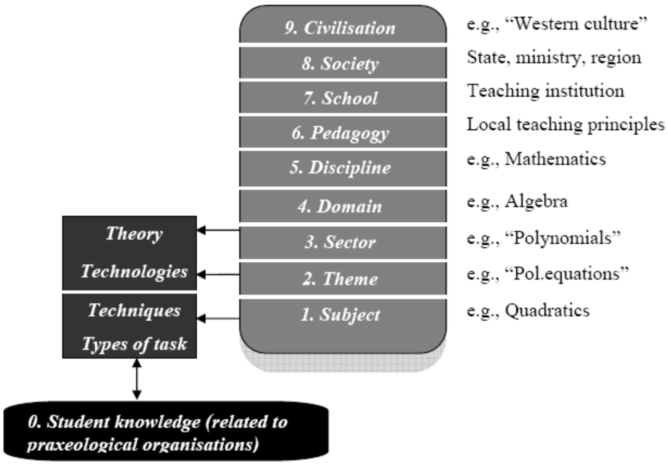

Figure 1. Three factors potentially to be compared: levels of didactic codetermination, mathematical praxeologies realised in school (through didactical organisations enacted by teachers), and the knowledge of students related to the mathematical organisations taught.

These different levels, which determine or condition the actual activity of study and teaching (didactical praxeologies), are all exterior to the teaching situation itself. Students meet didactical organisations in teaching situations, through mathematical and other praxeologies, in which the organisation seeks to engage the students. The personal knowledge (including both practical and theoretical forms) that they construct through participation in these situations may, of course, vary between students and also be different from what was intended by the teacher, the school, society, and so on. It seems reasonable and useful for our purposes in this paper to include, in our model as illustrated in Figure 1, this ultimate product of the teaching enterprise. From top to bottom, we thus have a system of processes through which didactical practices and their outcome are successively determined and realised by agents within the different institutional levels. The model provides a fine-grained, structured view of the institutional conditions for *didactical transposition*, the processes through which knowledge is transposed in and between institutions (Chevallard 1985), and ultimately reconstructed by the next generation.

LEVELS OF COMPARISON SUGGESTED BY ATD

Comparative studies address some of the levels in Figure 1, but rarely all of them. For instance, a study may consider a number of different schools within the same society and be focused on some of the underlying levels, such as pedagogies or overall aspects of students' practical knowledge in a given subject—a kind of comparison that may be of interest to parents about to choose a school for their child. Of course, comparison between different societies (understood as nations) is the hallmark of international comparison. In general, comparative studies of mathematics education may involve the following ten levels of comparison between two or more contexts (cf. Figure 2):

0. Students' knowledge in one or more specific subjects, situated and articulated within certain themes or sectors (and inevitably observed through a technology, which may to some extent be observed along with elements of students' theoretical knowledge);
1. Praxeologies of specific *subjects*, as prescribed by programs or official evaluations, or as actually found in the didactical practices of the contexts;
2. Local organisations of specific *themes* for the didactical practices of the contexts, as prescribed by programmes, described by teachers, or inferred from observation of several subjects within the theme;
3. Regional organisations of a specific *sector* in the didactical practices of the contexts, as prescribed by programmes, described by teachers, or inferred from observation of didactical practices on several themes within the sector;
4. Global organisations of specific *domains* in the didactical practices of the contexts, within a given discipline, as prescribed by programmes, described by teachers, or inferred from observation of several sectors;
5. Organisations of a *discipline* in domains or more globally, based on programmes and other evidence (including observation and assertions by teachers);
6. *Pedagogies* in the contexts, as prescribed by schools or programmes, observed, or described by teachers (this includes

principles for teaching that transcend the various disciplines taught, prescribed or observed interactions among disciplines, etc.);

7. Conditions and characteristics of specific *teaching institutions*, for example, regarding the roles, obligations, and autonomy of teachers;

8. Conditions and characteristics of whole *societies*, including in particular, the way in which schools are governed, funded, and systemically organised; and

9. Larger cultural contexts or *civilisations*, their principles for human society, in particular, as regards the role and meaning of education.

These ten levels of comparison are illustrated in Figure 2, which emphasises also that the comparison may concern both mathematical and didactical organisations, as observed (in real teaching) or prescribed (e.g., through national programmes and exams, textbooks, and so on).

It should be emphasised that concrete comparative studies may include several such levels and that a crucial feature of such studies—and the ways in which they can be used—is to what extent, and how, these levels of comparison are related; in particular, what causal relations are inferred or asserted between these levels of comparison. For instance, differences identified at Level 0 (student knowledge) could be claimed to be explained by differences at higher levels.

Moreover, it is crucial to consider whether the question of comparability is really addressed; in particular, to determine whether and how technologies found in the two contexts—at the various levels of institutional practice—are critically examined. One may, of course, find studies in which a universal technology (or just a technology particular to one context) is simply assumed, or even implicitly used; then it is a separate issue to see whether that is justified.

What we have introduced in this section is a technology to describe comparative studies in order to make such issues of comparability explicit and precise. It is of course to be tested in use on concrete comparative studies, and we proceed to do that in the following sections.

In short, to identify and characterise a comparative study on mathematics education, we propose to consider the following four questions:

a. At which of the Levels 0–9, described above, does comparison take place?

b. Assuming that the comparison includes didactical or mathematical organisations, what kind of evidence is used to study them? (E.g., direct observations, interviews with teachers, textbook analyses, etc.)

c. What methodological tools (in particular, technology in the sense of ATD) are used to interpret these observations or descriptions to ensure that the comparison makes sense at a given level? (e.g., at Level 2, how are comparable themes identified, and how is comparability argued for? In particular, how are language differences dealt with?)

d. How does the study relate *horizontal comparison* (between two contexts at the same level) to the *vertical relations* between levels found within each context? In particular, are some horizontal differences claimed to be caused by other horizontal differences (possibly situated at higher levels)?

Not all of the questions may be equally important for a given study. But in the following sections we demonstrate their pertinence to a wide selection of major comparative studies—and more broadly, in the interest of considering comparative studies from the point of view of ATD.

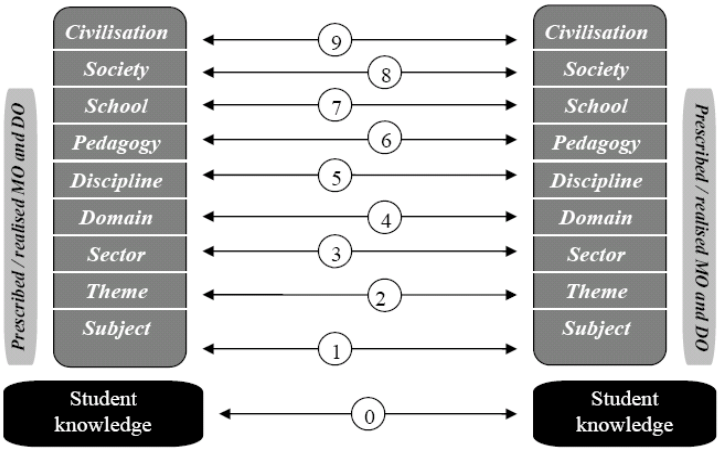

Figure 2. Possible levels of comparison of the codetermination of mathematical organisations (MO) and didactical organisations (DO) in two contexts (the contexts might coincide from some level upward, e.g., the same school).

TIMSS AND PISA

We begin by considering, in this section, two of the best-known cases of international comparative studies in mathematics education;

namely, TIMSS (Trends in International Mathematics and Science Study) and PISA (Programme for International Student Assessment). They belong to a lineage of large-scale studies that began with the First International Mathematics Study in 1964. Their main aim is to measure students' mathematical capacities across a large number of countries. In this paper, we consider neither the history nor the many critiques that have been made of such studies (see, e.g., Keitel & Kilpatrick 1999; Hoppmann, Brinek, & Retzl 2007). The complexity of the studies is such that we can provide only rough answers to our questions in this case. We base these answers primarily on the recent frameworks (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber 2005; OECD 2006) and on recent general (international) reports of the results (Mullis, Martin, Gonzales, & Chrostowsli 2008; OECD 2004).

TIMSS builds on a curriculum model with three components: intended, implemented, and attained curriculum, all studied at three grade levels (the fourth and eighth years of schooling and the end of upper secondary school). The intended curriculum is studied through official programs. Based on this curriculum, the model used to describe the implemented and attained curricula is constructed in order to "ensure that goals of mathematics and science education regarded as important in a significant number of countries were included" (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber 2005, p. 6). This model involves, in fact, four levels, which correspond neatly to the Levels 2, 3, and 4 of didactic codetermination of ATD (in parenthesis, we give illustrative examples from the TIMSS framework pertaining to the fourth year of schooling):

4. *Content domains* (Number)
3. *Topic areas* (Whole numbers)
2. *Objectives* (Compare and order whole numbers)
1. *Items* (Examples of tasks that correspond to the objectives)

TIMSS then compares national curricula at Levels 2–4 for both the intended and the implemented curriculum. As far as Level 5 is concerned, TIMSS operates with mathematics as a type of discipline, which is simply identified with actual school disciplines in participating countries; that identification gives rise to less ambiguity than for the science disciplines. Information on the implemented curriculum is collected through questionnaires given to school principals, teachers, and students; for the two former groups, information of a more general nature, concerning Levels 6 and 7, is also included; finally, public data concerning national features of education (Level 8) can also be drawn upon.

The attained curriculum—located at Level 0 but obviously influenced, if not determined, by the higher levels—is measured by submitting large samples of students in each country to a test containing items that pertain to objectives identified in the intended curriculum of that country; a significant feature of TIMSS is thus the possibility of omitting entire chunks of items based on their weak relevance to or absence from the intended curricula. These items are classified according to three levels of so-called cognitive domains (*knowing*, *applying*, and *reasoning*, with further subcategories that show that the three levels cover successively more advanced forms of mathematical activity). We notice in passing that the category *student knowledge* of our model is meant to encompass every exercise of mathematical practice and knowledge blocks, and so is much broader than *knowing* in the TIMSS sense. For all of these categories, universal validity is assumed, based on the participation of national experts in developing the framework. English is used as a common language, which implies many implicit assumptions regarding national experts' consideration of translation problems and of the degree to which the (English-language-based) framework allows researchers to take into account the nuances distinctive to each national language and context. In fact, when we write an article on mathematics education in another language than our mother tongue, we immediately realise that many expressions (even from elementary school mathematics) cannot be translated without ambiguity and approximation. Similarly, the same language may be used differently in different countries (particularly as concerns education and school mathematics). This variation implies a strong risk that the conditions of one country (such as the USA) are implicitly taken as reference model, and it becomes necessary to maintain the illusion that such a model can be objectively translated and compared with conditions found in other linguistic and cultural contexts.

The development of the PISA framework is not based on examining curricula in the participating countries. Instead, a new model of what is called *mathematical literacy* is elaborated in the project, more precisely by the so-called mathematics expert group. Given this approach to formulating a framework for describing the mathematics-related part of "how well students, at age 15, are prepared to meet the challenges they may encounter in future life" (OECD 2006, p. 7), it is interesting to note two influences that appear quite clearly in the framework itself: the Dutch research tradition of "realistic mathematics education" (cf. Lange 1996) and related German and Scandinavian research ideas on mathematical competencies, in particular related to modelling (e.g., Niss 2003). It is no coincidence that de Lange and Niss were both members of the mathematics expert group (OECD 2006, p. 187). The "mathematical

Kommentar [JK1]: See my previous comment.

literacy" model consists of three components: *situations or contexts* (in which mathematics is handled, from personal life to work or science); the *mathematical contents*, organised into *overarching ideas* (namely, *space and shape*, *change and relationships*, *quantity*, and *uncertainty*); and finally *competency clusters* (namely, *reproduction*, *connections,* and *reflection*). In addition to test items constructed to measure students' mathematical literacy along these dimensions, PISA also uses questionnaires for students and school principals (but not teachers) in order to determine certain features of the students' situation and the school context, and in some countries there is a questionnaire for parents. That is, certain variables at Levels 6 and 7 (concerning pedagogy and school) can be related to the outcome of students' responses to the test items and the questionnaire.

PISA's main aim is to confront the described model of mathematical literacy, which is supposed to be desirable in any society (Level 8), with students' performance on corresponding items. But the data give no means to evaluate whether high or low performance is caused by regulations at Levels 3–5 (such as a mathematics curriculum with different components than those assessed), levels of efficiency in teachers' realisation of the curricular goals (mainly Levels 1–3), or something else. Other studies (see, e.g., Hopmann et al. 2007) have considered this question for particular countries.

On the other hand, it is possible to analyse the PISA framework itself as suggesting certain types of mathematical praxeologies to be particularly relevant for students' future life and activity in society. The competency clusters, which at first sight resemble the cognitive domains in TIMSS, are in fact further described in terms of Niss's (2003) competency categories, which include both technical and technological components (for a further discussion of the links between competence categories and praxeologies, see Winsløw 2005). It is also possible to interpret the overarching ideas as mathematical domains and compare them with what is found or prescribed in schools. It may be that the test items can be ordered according to sectors and subjects. But since the details of such an analysis would require public access to the test items, we can only conclude that the framework itself outlines an independent structure of mathematical domains and themes that may or may not correspond to parts of the Levels 3 and 4 of didactic codetermination for a given country.

The outcome of both TIMSS and PISA is a huge collection of data pertaining to Levels 0 and 2–8 (for PISA, mainly Levels 0 and 6–8), which can to some extent be considered at and between levels in order to identify correlations across levels. In both surveys, students' performance on test items (pertaining to Level 0) is clearly the main data object, while the rest is considered background data, a term clearly suggesting that they are determinants for Level 0. The main correlations sought for are thus between students' performance on a certain group of items—such as all items related to mathematics—and other parts of the data that can be linked to the students (e.g., country, characteristics of their school's pedagogy, intended curricula, gender, and social class or ethnicity of the students). The studies clearly assume that one can interpret vertical correlations between Level 0 (student performance), on the one hand, and background data (from higher levels), on the other, in directly causal ways signalled by terms like *impact*, *influence*, *results in*, and so on (cf. Figure 3). This is already apparent in the statement of their overall purposes:

> By participating in TIMSS, countries can . . . understand the contexts in which student learn best. TIMSS enables international comparisons among the key policy variables in curriculum, instruction, and resources that result in the higher levels of student achievement. (Mullis et al. 2005, p. 10)
>
> PISA is designed to collect information through three-yearly cycles and presents data on the *reading*, *mathematical* and *scientific literacy* of students, schools and countries. It provides insights into the factors that influence the development of skills and attitudes at home and at school, and examines how these factors interact and what the implications are for policy development. (OECD 2006, p. 9)

In the presentation of results, one finds many implicit and explicit claims about how differences at Level 0 (e.g., concerning average mathematics performance) may be related to differences at higher levels (e.g., at the school level, the number of out-of-school mathematics activities the school offers):

> Schools' offering of activities to promote student engagement with mathematics, such as mathematics competitions, mathematics clubs or computer clubs related to mathematics, show a positive impact as well, over and above all other factors. Each additional such activity that is offered by schools is associated with an average performance advantage of 7 score points. (OECD 2004, p. 259)

The key word here is, again, *impact* and its synonyms. The problems with asserting such "impacts" are twofold: First, are the horizontal comparisons both meaningful and valid (related to the framework, questionnaires, items, data sampling, and treatment)? And second, assuming the horizontal differences are real, does one produce the other? It is indeed possible that both differences are instead produced by a third factor, at the same or a higher level, being similarly different for the two contexts. In the example given by the above quotation, one could think of factors like "school budget," "teachers' affective relation to mathematics," and so on. If in fact one such factor correlates strongly with both student performance and offering of activities to promote student engagement with mathematics, then the last sentence in the quotation becomes dubious, at least if it is read to say that by adding an activity of the specified type, a school could improve its students' performance on PISA items by an average of 7 points. The collection of data in PISA is so huge that it is very difficult to control for this type of covariation, even disregarding the issue of data validity.

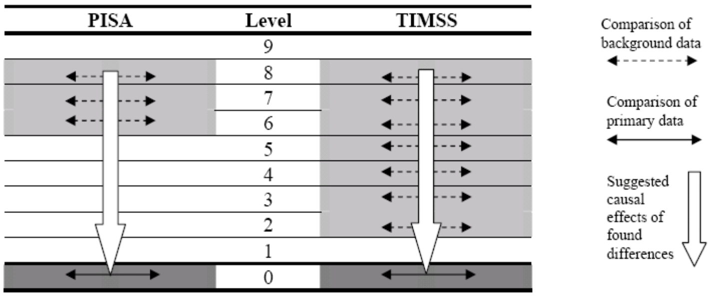

Figure 3. PISA and TIMSS compare several contexts horizontally at a given level—the primary level being student knowledge as measured by test items. For simplicity, the figure shows the comparison of just two contexts. Causal effects are sought in the background data.

An important technique to produce findings from the complex data material consists in grouping students according to one variable (typically country) and then presenting the average of another variable for the groups (like test performance or expressed interest in mathematics). In the tables that result, the assumption that the second variable somehow depends on the first may be implicit but still strongly suggested. We say "somehow" because it is not always the first variable that is seen implicitly as the cause. In the case of country (Level 8), the tendency is to view differences in performance as caused by the sum of the underlying levels, as it would obviously be meaningless to claim that, for instance, performance or interest in mathematics depends directly on nationality. National policies (regarding curricula, pedagogy principles, teacher education, and so on) appear then as candidates for causes and may be further investigated through new groupings across nations (related to variables at these lower levels). However, all correlations identified remain weak hypotheses given the number of variables available and the large variations one finds within different groups, such as the students in several of the participating countries.

Indeed, many aspects of both surveys remain highly intricate and controversial: the details of questionnaires and test items (the latter mainly secret), the data-processing methods, and so on. The secrecy surrounding test items prevents us from confronting the measurement at Level 0 directly with the mathematical and didactical practices in schools (Levels 1 and 2), which, by the way, are not observed directly in TIMSS or PISA (but are considered in some follow-up studies, like those we consider in the next section). It is therefore impossible to evaluate—within concrete mathematical themes, sectors, or domains, or for mathematics in general—the validity of claimed links between measurements of student performance and what happens in mathematics classes in schools. As we just noted, however, such links tend to be assumed in the reports, which mainly correlate student performance with higher level factors determining the mathematical practices in school classes. And whatever the item details, one may contend that answers to multiple-choice items give very limited information on students' knowledge in the sense of praxeologies they master individually, even at the level of practice blocks (since a given answer can usually be found using more than one technique). Given the results from these tests, we can infer next to nothing about the (social, institutionally situated) praxeologies students encounter in school mathematics. Conversely, to the extent information is available about the school mathematical practices (particularly at Level 1) in which the students tested have been involved, effects at Level 0 remain highly speculative.

Nevertheless, both surveys provide a rich source of material on which one may base, at least initially, more detailed discussions of the effects of didactic codetermination, and which can motivate more fine-grained studies of particularly intriguing hypotheses as far as effects on Level 0 are concerned. This possibility is not least due to the fact that the frameworks (which serve as a technology for the comparison) are very explicit, especially at higher levels. The involvement of national experts in their elaboration warrants a certain degree of compatibility with more local discourses at these higher levels. One problem that these frameworks share, and which may be hard to avoid given the number of contexts involved, is the lack of discussion of linguistic details and references to research specific to each context. Despite the involvement of national experts, the translation of frameworks, the test items, and so on are bound to be highly approximate—even as regards the technological elements of mathematical organisations as they appear in schools. We return to these questions as they appear in binary comparisons (the sections below on ICMI Study 13 and recent doctoral studies).

LARGE-SCALE INTERNATIONAL STUDIES OF CLASSROOM ACTIVITY

The 1995 TIMSS Video Study (see Kawanaka, Stigler, & Hiebert 1999) was among the first attempts to get more direct information on classroom practices in order to supplement and explain the results of surveys of the type considered in the previous section (of course, mainly TIMSS). The basic unit was the lesson; in principle, 100 random lessons in eighth-grade classes were videotaped in each country (in the 1995 study, the USA, Germany, and Japan). In practice, some compromises were made, such as just 50 lessons from Japan. The video recordings are subtitled in English so that they can be analysed by Anglophone researchers. For each lesson, a table is produced that shows the phases of the lesson, as regards:

- Time interval (within the lesson, corresponding to lesson phases)
- Organisation of interaction (class work, individual seatwork, …)
- Activity (e.g., setting up, working on tasks)
- Short description of contents (Task 1: “Find the angle…”; Student solution 1: …)

The main focus of the study—as evidenced by the coding schemes—is on structures in the lesson that are supposed to be generic for the discipline or country; that is, at Level 6 (pedagogy) or Level 5 (mathematics): How are lessons organised by time period (phases)? Who presents solutions to tasks? How many techniques are presented? To what extent are the tasks of a routine nature? Do proofs appear in the lesson or not? And so on. Striking results include, for instance, the average percentages of seatwork time that are spent on general task types: “practicing routine procedures,” “invent new solutions/think,” and “applying concepts in new situations” (the last code being defined as the negation of the first two). Praxeologies enacted in classrooms are considered under these general angles and more recently also with respect to domains (Level 4), such as algebra and geometry, combined with the general task types such as the above (Neubrand, in Leung, Graf & Lopez-Real 2006, p. 300).

In short, the aim is to find differences between general (country-specific) patterns of how mathematics lessons proceed. The hypothesis of the study is that certain aspects of pedagogy (Level 6) and of the global aspects of school mathematics (Level 5) are stable within a society (country), where they remain valid for all schools (Level 7) and, in particular, for mathematics teaching in general (Level 5 and below). This hypothesis seems to be validated by the study data; for instance, as regards general lesson scripts or the distribution and quality of students' and teachers' interventions. The methods and results of the study, particularly concerning national scripts for mathematics lessons, remain controversial (see, e.g., Keitel & Kilpatrick 1999; Clarke, Mesiti, O'Keefe, Xu, Jablonka, Chee Mok & Simizu 2007). But assuming mathematics teaching in a country has qualities that may be compared to what is found in other countries, the TIMSS Video Study has certainly been pioneering in advancing a serious and operational alternative to measuring those qualities in terms of students' responses to standardized tests—or, if you will, to complement such measurements with hypotheses for more well-founded causes of the differences found.

The Learners' Perspective Study (Clarke 2006)—LPS in the following—aims at investigating more deeply the practices in classrooms through videos of a small set of 10 consecutive lessons in each country, also in Grade 8, complemented with interviews with students and teachers, textbook analyses, and other data that permit the researchers to take into account a wide variety of perspectives. The video data itself is richer in the sense that three angles (teacher's, students', and whole class) are filmed in a given lesson. The Levels 1–4 of students' praxeologies may thus be considered for each sequence of lessons, but as they are merely samples—even if they are aimed at being representative of the curriculum—they cannot be directly compared. Also, there is no claim to national representativeness for the lessons filmed; instead, rather good teachers are deliberately selected under the assumption that one could learn more from studying their lessons. Indeed, this study is not preoccupied as the TIMSS video studies are with direct horizontal comparison between national contexts, but instead they make an in-depth analysis of the vertical relations (from Level 7 down to Level 0) within each classroom considered (cf. Figure 4). The deliberate focus on learners' practice (Level 0, related to Levels 1 and 2) is a significant feature of LPS, making it in some ways an organisation of parallel studies of single classrooms carried out in different contexts without aiming a priori at common measures. The goal, however, is clearly that one may learn from studying the material and analyses from other contexts (be it countries, societies, or civilisations). It is interesting to notice the involvement of local experts of the relevant school curriculum in analysing these relations and also the requirement that any published interpretations of data from a given country "must be validated by member researchers from the country providing the data" (Clarke 2003, p. 176).

LPS thus differs from the studies previously considered in the following sense: The horizontal comparison across countries is not a primary goal, and to the extent it is done, it is based on the primary study of vertical relations within one country or school; that is, the levels of codetermination as considered in the primary part. Indeed, many of the papers published from the project present only results from one country. Indeed, in the research questions and the research design (Clarke 2006), the main emphasis is on studying relations that are internal to the contexts: between observed practices of students and teachers (Levels 0 and 1, but also Levels 2 and 3, as coherent sequences of lessons are considered) and the relation to local curricular goals (Level 2 and higher). International differences or cultural specificities are discussed based on these parallel, in-depth studies of vertical relations within each context, but, given the few and relatively random content areas that are encountered in each country, only at the higher levels (5 and above).

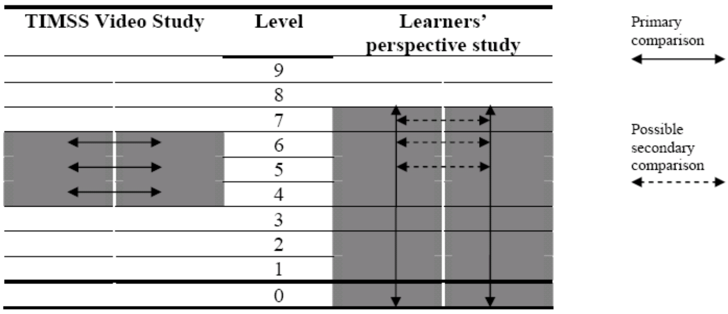

Figure 4. TIMSS video studies compare pedagogical principles from contexts directly. LPS seeks to first establish vertical relations (from data at Levels 0–8) and then possibly compare the result among contexts.

This difference between the two video studies may also help to explain the divergence as regards the *results* obtained, especially on the issue of lesson structure (see Clarke et al. 2007). In the TIMSS Video Study, looking at large numbers of random lessons, we saw that certain scripts regarding the pedagogical structure of lessons (Level 6) were identified. These structures seem to disappear in the Learners' Perspective Study, and that could, somewhat paradoxically, be seen as a consequence of the more fine-grained data available:

> An inevitable consequence of any nationally representative sample of individual lessons is to average over the distinctive lesson elements, whose location in the lesson is a direct and informative reflection of the lesson's location in the topic sequence. The detailed LPS data set supports a fine-grained analysis of both the form and the function of lesson events. (Clarke et al. 2007, p. 292)

Metaphorically, regularities that appear at great distances may disappear in a close-up look. More precisely, studies with a strong focus on relating levels vertically and on codetermination of the subdisciplinary levels may not always support conclusions from surveys aiming at horizontal comparisons at the higher levels (obtained by averaging over large data sets). Another possible reason for the lack of evident patterns in the lessons observed by LPS, suggested to us by one of the reviewers of this paper, could be the deliberate choice to consider lessons taught by excellent teachers. The results could just confirm the likely hypotheses that excellent teachers are also exceptionally autonomous with respect to how they structure their lessons.

THE CIVILISATION LEVEL: THE EAST-WEST CASE (ICMI STUDY 13)

Ever since the Second International Study of Mathematics (SIMS; see Travers 1988) was completed in 1981, it has been clear that such surveys appear to show systematic differences that go beyond countries and educational systems. The most striking of these is the superiority of students from East Asia (Japan, Hong Kong, Korea, …), according to virtually all available measures of their overall performance in mathematics. At the same time, it has been observed that culturally mixed countries like the USA display relatively large variations among students that are moreover linked to ethnicity more than to other sociological variables; so that, for instance, Asian American students' performance is quite similar to that of students in East Asian countries despite the lower average of American students in general (cf. Clarke in Leung et al. 2006, pp. 354ff.). This finding leads to the hypothesis that Level 9 of our model—civilisation, in the sense of larger cultural context—may account for some of the differences observed at lower levels of codetermination and, through them, in students' performance:

> Our contention is that cultural divisions are much more meaningful than political or geographic divisions in explaining differences of educational practices in mathematics. East Asia and the West in this study are therefore cultural demarcations rather than cultural divisions, roughly identified as the Chinese/Confucian tradition on one side, and the Greek/Latin/Christian tradition on the other. (Leung et al. 2006, p. 4)

The assumption that one civilisation or culture dominates a society (built into the model we use here) is not invalidated by the existence of multicultural societies. However, the results alluded to above suggest that the determining force of cultural factors may well be independent of those from what could be called official society, and conceivably bypass most of the levels in Figure 2 to exercise its influence more or less directly on students' activities and learning outcomes. To investigate such hypotheses clearly requires deeper studies of well delimited forms of civilisation, and that was indeed what the International Commission on Mathematical Instruction (ICMI) comparative study of East Asia and the West set out to do—following the call from which the above quotation comes. The comparison of mathematics education in East Asia and in the West has certainly been studied elsewhere, but to demonstrate the use of our model, we focus on the ICMI study, which we find particularly rich and also quite representative of the field as such.

The study had a reasonable balance of authors coming from East Asian and Western countries. Nevertheless, the overall impression is that viewpoints of a Western origin (which certainly permeate the present paper!) tend to dominate—perhaps in part as an effect of a Western language (English) being the norm for all chapters. Part of the comparative challenge, indeed, is related to the different sets of analytic frameworks that scholars from different cultural traditions bring with them—in fact, frameworks from the two sides are combined in some chapters, particularly in those with authors from different traditions. It is evident that comparison of civilisations cannot be done from a higher, neutral viewpoint; measuring one in terms of the other inevitably leads to a colonial position, which does not further our understanding. As a consequence, we (as Western readers) can perhaps learn most from those parts of the study that are substantially based on East Asian models for describing and analysing differences; these are therefore privileged in the following sample.

The first two chapters are written from manifest East Asian perspectives; they provide portraits of cultural and historical tendencies linked to education in Japan (clearly at Level 9). In particular, Hirabayashi (in Leung et al. 2006, pp. 51–65) points out the importance of Japanese notions like *dô* (道, roughly translated as "way") and *justsu* (techniques), which in the Japanese tradition are important factors in educational goals. Considering mathematics as *jutsu* emphasises it as a set of efficient tools; considering it as *dô* emphasises it as a specific, coherent intellectual enterprise. Hirabayashi contends that the prevailing role of *jutsu* over *dô* may be responsible for what he sees as central shortcomings related to students' motivation in modern Japanese mathematics education (implicitly at Level 5 and lower). A related institutional structure—

that of *iemoto* (家元, roughly “spiritual family house,” evolving around common goals)—is described by Emori and Winsløw (in Leung et al. 2006, pp. 553–566). They use it to interpret observations from secondary mathematics classrooms and compare them with Western counterparts to show the pertinence of this idea. Although the main conclusions concern Level 5 (the meaning of mathematical sign systems and discourse, learning goals of mathematics), the cases considered are clearly situated within subjects and themes (Levels 1 and 2). Similar lines of thought were developed by Wong (in Leung et al. 2006, pp. 111–121) for the Chinese context (but probably within the same civilisation): again, a notion of “way” (translated here as “right way to do something”) is central to describe what education is supposed to initiate the learner into; some examples at Levels 1 and 2 are considered to show the meaning of the deeper principles (at Level 6) for such initiation: repeating and varying techniques related to a subject.

Several contributions to the study are concerned with investigating teachers’ values and beliefs related to mathematics (codetermination Level 5) and to pinpoint systematic differences between two cultural traditions (such as China and Australia). An interesting variation from the more obvious ways to organise such studies is presented by Tiong and Bishop (in Leung et al. 2006, pp. 523–535), who report on a qualitative study with two East Asian mathematics teachers who moved to Australia and are now teaching there. They identify a number of value differences related to the way mathematics is taught (Level 5) and what they see as underlying factors in schools and in societies (mainly Levels 6 and 7); the authors identify these with certain universal dimensions of cultural variability such as “individualism/collectivism.” The opinions of the two teachers are interesting not because they are representative of anything but because of the teachers’ privileged viewpoint on the boundaries between the two cultures of teaching, and the cultures more broadly. A similar idea can be found in the chapter by Isoda, McGrae, and Stacey (in Leung et al. 2006, pp. 397–408); they report on the benefits and challenges experienced by Japanese and Australian school kids as they discussed their ideas and methods for solving a set of mathematics problems. Such an experiment is obviously difficult to set up in a reasonable way, especially given the language barriers; but it could provide an interesting supplement to the methods found in studies of videotaped lessons.

Another set of studies concern the analysis of global differences between textbooks and curricula in East Asia and the West. As a strong example, Park and Leung (in Leung et al. 2006, pp. 227–238) compared major eighth-grade textbooks in China, Japan, Korea, the UK, and the USA and found significant differences between the Asian and Western books (and commonalities within each group). For instance, they found that the Western books reflected a more individualised view of the learner, with a stronger emphasis on attractive design, real-world examples to motivate new topics, and the provision of different levels of approach. Asian textbooks at this level are more terse or minimalist in their presentation, as “East Asian culture believes in orthodoxy, and students are believed to adhere to the orthodoxy despite their individual differences” (p. 236). Here the word *orthodoxy* comes, for the Westerner, with a number of connotations that are not likely to have been intended by the authors; the text as a whole suggests that it refers to the larger cultural idea of teaching a common “way,” mentioned above, corresponding to the value of collectivism and the ideal of social harmony in society. Thus the manifest differences of textbook styles seem to fit into a larger picture of determining forces coming from Level 9.

Finally, several of the studies remind us that East Asian and Western cultures, societies, and systems of education have interacted in various ways and cannot, therefore, be considered as independent systems. Even in East Asian countries like Japan, which was never colonized, Western methods in mathematics teaching have been at times systematically imported, and of course subsequently adapted, as explained, for example, by Ueno (in Leung et al. 2006, pp. 65–79). Moreover, within each of the two civilisation groupings, one can find enormous variations that cast further doubt on any easy generalisations. Kaiser, Hino, and Knipping (in Leung et al. 2006, pp. 318–350) proposed eight “empirically grounded” variables for comparing mathematics teaching in different countries (all at Level 5 or 6) which they describe in some detail for four countries (England, France, Germany, and Japan). Their discussion is mainly based on previous studies of the first three countries and more limited evidence from Japan (due “mainly [to] time and capacity restrictions”). For most of the variables, they found that England and France had “polar approaches, with Germany and Japan in between” (p. 343). This is, for instance, the case as regards the role of real-world examples, which was found to be strongest in England and weakest in France, with intermediate positions occupied by Germany and Japan.

In fact, Western mathematics teaching may not be a meaningful category. Based on work that we shall consider in more detail in the next section, Bessot and Comiti (in Leung et al. 2006, pp. 159–179) provide a multilevel comparison of mathematics teaching in Vietnam and France: It goes from a brief historical sketch of how mathematics curricula developed in the two countries (Level 8 and below) to a study of how graphs may be used by students in investigating properties of functions (Level 3 or 2). The study of the genesis of educational practices in these countries shows how competing ideologies and philosophies (at Level 6 or 8) as well as institutional systems (Level 7) have succeeded each other in ways that cannot be accounted for in terms of a generalised East Asian or Western civilisation.

BINARY COMPARATIVE STUDIES IN RECENT DOCTORAL THESES

Up to now, we have considered the diversity of widely published comparative studies on mathematics education and its outcomes and have put to the test the use we propose of ATD for making sense of this diversity. Most of the studies evoked concern ambitious projects carried out by international groups of researchers, and they have been developed without any theoretical connection with ATD. In the frame of our metastudy, we decided to complement this test by exploring the potential of ATD for making sense of and comparing research projects offering quite different characteristics, in terms of scope, ambition, and theoretical environment. For that purpose, we decided to focus on French comparative doctoral theses. These objects are indeed more modest in scope than the projects evoked in the previous sections, they obey different constraints, and they are achievements of a single researcher. Moreover, being prepared in France and supervised or cosupervised by French researchers, their authors have been immersed in an educational culture for which ATD is a core component. For this part of our metastudy, we selected a corpus of eleven comparative doctoral theses prepared in France. More precisely, the sample consists of two subsamples. The first subsample is composed of comparative doctoral theses prepared at the University Paris Diderot – Paris 7 in the DIDIREM team, which is the team having the largest number of doctoral students in mathematics education in France. The second subsample is composed of comparative theses prepared at the University of Grenoble 1, in the context of a long-term collaboration established with Ho Chi Minh Pedagogical University in Vietnam.

In this section, we consider the following questions: To what extent is the complexity of comparative studies, as illustrated in previous sections, approached in these doctoral theses? How do the characteristics of the doctoral academic exercise influence the type of comparative research carried out, and is our framework able to capture that influence? What do these theses add to the kind of comparative research evoked so far? What has been the influence of the cultural proximity with ATD on these doctoral theses, if any, and what can we learn from the analysis of such influence?

Our sample is quite limited and may not even be considered a representative sample of comparative thesis research carried out in France. However, we think it is appropriate for approaching these questions. A first reason is that the corpus at our disposal, even if limited, presents an important diversity in terms of the countries involved and the way comparative research themes are constructed and methodologically addressed. Moreover, it contains both a few isolated theses from a variety of countries, those prepared at DIDIREM, and a group of related theses, those prepared at the University of Grenoble 1. Figure 5 provides a general presentation of the sample.

Author (year)	**Countries**	**Research theme (rough outline)**
Amra (2004)	France-Palestine	The teaching of functions in upper secondary schools
Celi (2002)	France-Italy	Teaching of geometry for 11–16-year-old students and its effects
Cabassut (2005)	Germany-France	Reasoning and proof in secondary mathematics education
Stoelting (2008)*	Germany-France	Functional thinking of students 10 to 16 years old and its teaching in France and Germany
Châu (1997)*	France-Vietnam	The teaching of vectors at the 10th-grade level
Hai (2001)*	France-Vietnam	The teaching of 3D geometry in upper secondary school
Tien (2001)*	France-Vietnam	The relationships between functions and equations in upper secondary school
Thanh (2005)*	France-Vietnam	Introduction in secondary schools of algorithms and programming with calculators
Quoc (2006)*	France-Vietnam	The solution of quadratics in secondary education
Khanh (2006)*	France-Vietnam	The notion of integral at upper secondary school
Trung (2007)*	France-Vietnam	Relationships between the teaching of the limit notion and the “decimalization” of real numbers in a calculator environment

Figure 5. The 11 theses of our sample. Theses with an * had two cosupervisors, one in each country. Most of the theses may be accessed online at the database http://ccsd.cnrs.fr

One first characteristic emerging from the table is that the theses focus on the comparison at the level of some particular mathematical domain or sector (Level 3 or 4), or around a type of paradigmatic practice of the mathematical discipline, such as proof (Level 5). The titles of the theses confirm this claim, being, for instance, (in our English translation): “Proof, Reasoning, and Validation in the Secondary Level Teaching of Mathematics in France and Germany,” “A Didactical and Epistemological Study of the Teaching of Vectors in Two Institutions: 10th Grade in Vietnam and France,” and “The Contributions of Comparative Didactical Analysis of the Solution of

Quadratic Equations in Secondary Education in Vietnam and France." These titles also suggest that *horizontal comparison* is the main aim. But in order to understand the reasons motivating the comparisons undertaken, how they are achieved, how the relationships between the different levels of codetermination are taken into account, and up to what point comparability is critically examined, one needs to enter the details of the comparative work. For this purpose, we have constructed a specific grid of analysis shown in Figure 6. For the synthetic presentation of the results, we split the corpus into two subcorpuses, the French-Vietnamese theses (forming a coherent series) and the rest. We then globally discuss the role played by ATD in these comparative theses.

<u>Basic data</u> of thesis (author, title, year of defence, university, supervisor or supervisors)
<u>Motivation</u> of the comparison (outline)
<u>General characteristics</u>: countries compared, levels of teaching considered, scale of contexts of data, the objects compared, main level at which comparison is done (cf. Figure 2), secondary levels (if relevant), methods used to collect and compare data.
<u>Theoretical and methodological dimensions</u>: theoretical frameworks and constructions on which the comparison is founded, the most important notions and their function in the project; the methodological tools specifically elaborated within the project
<u>The results obtained</u>: situated, as far as possible, at the levels of comparison (cf. Figure 2) and the extent to which comparisons at different levels are related, possibly causally; difficulties encountered in the course of comparison

Figure 6. The framework used to analyse the theses listed in Figure 5.

1. The French-Vietnamese theses.

Bessot and Comiti were at the origin of the French side of the French-Vietnamese collaboration underlying the theses we consider here. In a recent article (Bessot & Comiti 2009), they complement the information and results provided by them in the ICMI study (and presented at the end of the section on that study). The article provides a synthetic vision of the way comparison has been approached in these theses and points out the coherence of the whole enterprise. It is made clear, for instance, that within this collaboration, the choice of favouring comparative issues for doctoral thesis was made quite early. Its main aims were as follows: (i) to reinvest and put to the test didactical knowledge built in the French context, (ii) to answer the needs of the Vietnamese educational system, and (iii) to favour the integration of new doctors into the scientific community of their country. Moreover, comparative studies were seen as a privileged means for questioning the tendency we all have to consider our own particular context as normal or natural—the phenomenon that Chevallard (date?) describes as naturalization of mathematics and didactical praxeologies. These comparative studies thus have specific characteristics expressed by the two authors in the following terms:

Kommentar [JD2]: A date is needed

> Far from a "cognitivist" approach centred on the particularities of students and teachers, this type of research situates the problem of students' and teachers' difficulties in the totality of conditions and constraints which bear upon the teaching of mathematics (Bessot & Comiti 2009, p.; our translation)

Kommentar [JD3]: Complete pages

This quotation explains why institutional analysis, relying on curricular documents and textbooks, plays a fundamental role in the methodology developed for these comparisons. In general, institutional analysis and comparison has both a synchronic and a diachronic dimension. The importance given to the diachronic dimension is consistent with the reference made to the notion of didactical transposition, understood as a long-term process. In the diachronic work, specific importance is given to the moments of curricular reform. At these points of change, one may more clearly observe the technological and theoretical discourse produced by the *noosphere* in order to explain and justify the choices made in terms of both mathematics and didactical organizations.

Bessot and Comiti (2009) provide an example that illustrates the type of result obtained through such diachronic comparative analysis. It deals with the different status given to approximation in the Vietnamese and the French teaching of analysis. This difference was identified first in the thesis by Lê Van Tien, who studied the relationships between functions and equations (our Level 3) in upper secondary school. Differences observed at this level were interpreted by Lê Van Tien by considering more globally the historical changes in the epistemological vision of analysis (Level 4) and in the status given to experimental activities (Levels 5 and 6) in the two countries. Notice that theses of Nguyen Chi Thanh, Tran Luong Cong Khanh, and Lê Thai Bao Thien Trung build substantially on the thesis by Lê Van Tien, which adds a cumulative effect to the work within this group of theses (with its virtually novel research context).

Institutional analysis of mathematical and didactical praxeologies plays a predominant part in the methodologies used in the theses, but connections with implemented and achieved curricula are also present. On the one hand, students' questionnaires are used for clarifying the impact of identified differences in mathematics praxeologies on students' practices. On the other hand, institutional comparison can lead to the conception, implementation and analysis of didactical engineering products (for use in the Vietnamese system).

This is the case, for instance, in the two theses by Nguyen Chi Tanh and Lê Thai Bao Thien Trung, which are less focused on comparison than the others. In their projects, the institutional comparison is put at the service of didactical design; as expressed by Bessot and Comiti (2009):

> The realisation and analysis of a didactical engineering [based on such an institutional comparison] in one of the systems, allows not only to raise questions about this system but also on the other system, and it favours thus the emergence of generic questions whose pertinence are suggested by the other system. (p.)

Kommentar [JD4]: Complete page

It is worth noticing that a direct study of the implemented curriculum through the observation and analysis of classroom practices comparable to what has been described in section about large-scale international studies is not part of the methodologies used in this subcorpus.

This set of theses thus offers a quite coherent picture. The main focus is generally situated at Levels 3 and 4. The comparison explicitly takes into consideration various levels of codetermination, mainly up to Level 7, even if societal and cultural factors are regularly invoked, especially through reference to history and to the different foreign countries that have successively imposed their views on Vietnamese education, beyond the long-lasting influence of the Confucian heritage (cf. the section about the civilisation level).

Theoretical frames are mainly provided by ATD, including the theory of didactical transposition, and the theory of didactical situations when a dimension of design is present. The use of ATD naturally leads to consider upper levels of codetermination even when this notion is not explicitly used. What is at stake is the understanding of an ecological system and its possible dynamics. In accordance with the theoretical framework, students and teachers are not considered in their individuality (as pointed out in the first quotation above). Questionnaires and experimentations involving them are generally small-scale qualitative studies. There is no direct study of the implemented curriculum. We suspect that some characteristics of the Vietnamese context contribute to this methodological choice: the existence of one official textbook and a detailed teacher guide, plus the dominant culture of transmissive pedagogy (cf. Bessot & Comiti, in Leung et al. 2006, p. 169). They can lead researchers to consider that the implemented curriculum is rather uniform and close to the intended curriculum. Moreover, even when comparison is the apparent focus of the study, the goal of the enterprise is a better understanding of the rationality of the two systems at stake and of their possible evolution.

Finally, it should be noted that in this group of theses, the two countries are far from playing a symmetric role. The doctoral projects were all done by Vietnamese students, and their final goal was to serve the educational system of their country. On the other hand, theoretical frameworks and methodology mainly come from the French tradition of research in didactics.

2. The DIDIREM comparative theses.

The second part of the corpus consists of isolated theses, all of which were prepared with the DIDIREM research team at the University Paris-Diderot; however, among the authors, just one (Cabassut) is French. The overall image we can give of these theses is not as homogeneous as with the first subcorpus, but some common traits are nevertheless observed.

As in the first subcorpus, curricular analysis carried out through the study of syllabi and textbooks plays a fundamental role. It generally follows an epistemologico-mathematical analysis of the domain at stake and a research review that supports the framing of the didactical analysis. Except for one thesis (Stölting), the main theoretical framework is again ATD; however, complemented by more local constructs considered especially relevant for the mathematical domain at stake or more globally complementing the constructs provided by ATD. From this point of view, it is worth noticing that even if ATD has a semiotic dimension structured around the dialectics between ostensive and non-ostensive, semiotic phenomena are mainly addressed through the constructs and categories provided by Duval’s (1995) more elaborate theory concerning semiotic systems of representations.

As with the first subcorpus, the results of the institutional analysis carried out at levels from 3 to 5 are often interpreted by referring to higher levels of codetermination. For instance, to explain the great differences between French and Palestinian mathematical and didactical praxeologies related to the notion of function, Amra points out the diverse colonial influences that have shaped the Palestinian mathematics syllabi and also the transmissive vision of pedagogy that prevails there. Similarly, in order to interpret the differences she observed, Celi refers to general characteristics of the curricular organisations in Italy such as the greater autonomy given to teachers, the different role given to textbooks, and the historical culture of Euclidian geometry in mathematics education.

Finally—again much as in the first subcorpus—one can observe a jump from the study of intended curriculum to the study of achieved curriculum, implemented curriculum being the missing link. Institutional analysis appears as a first and necessary step in the comparative work, necessary for making sense of students' behaviour. The constraints of doctoral study (developing one's first personal research work to be carried out in a few years) do not permit one to seriously address both the implemented curriculum and the achieved curriculum. In the current state of development of comparative studies, it is understandable then that the prevalent choice is to examine the effects on students. Within that general trend, differences are nevertheless observed in the importance given to this dimension of the research (from one chapter to an importance comparable to that given to institutional analysis) and in the methodologies developed (self-constructed questionnaires, interviews, use of data coming from other sources such as PISA, national or regional studies, and competitions involving two different states).

This is also a place in which issues of comparability often crystallize. For instance, in Celi's thesis, selecting a theme appropriate both for French and Italian tenth-grade students and formulating the precise questions posed to the students raised difficult problems because of the difference between France and Italy in geometrical approaches and tools, and also the fact that in contrast to Italian students, French students are used to very detailed and guiding formulations of tasks. The situation is even worse in Amra's thesis: The French approach to functions is covariational, whereas the Jordanian approach is set theoretical. This difference led Amra to construct tasks that are supposedly unfamiliar to students in both countries and to look at the way they make sense of unfamiliar tasks.

Contrary to what might have been expected, linguistic issues do not necessarily emerge as a major concern. Cabassut's thesis is an exception in this subcorpus. In order to cope with the diversity of terms currently used for addressing issues related to validation and proof in the educational communities in France and Germany, he develops a systematic comparative analysis of the vocabulary used in the two countries. Moreover, motivated by the differences between French and Germany with regard to didactical contracts related to the formulation of proofs, he develops a specific tool for comparing what he calls the *discursive expansion of proofs*.

One of the theses (Stölting) in this group has supervisors from two countries. It is interesting to notice that this situation has an evident impact on the theoretical frameworks used. The domain at stake is that of functions, and it is approached through the notion of *functional* *thinking*, introduced in Germany at the beginning of the 20th century. The study of functional thinking in the thesis combines general constructs coming from the two didactical cultures: the notion of conceptual field due to Vergnaud (1990); semiotic registers in the sense of Duval; *Grundvorstellungen* (fundamental ideas) due to Hofe (1995); and *Grundkenntnisse*, a notion created in the thesis. A similar combination or integration of theories can be observed in Celi's thesis, where notions that are part of the educational discourse in Italy (e.g., the distinction between an attitude of rational intuition and a rational attitude) are connected and combined with the notion of geometrical paradigms due to Houdement and Kuzniak (2000). These examples show that comparative studies also have to address the issue of compatibility and connection between the theoretical frameworks and notions supporting didactical research in the countries involved, and more globally the ideas involved in their educational discourses. In our opinion, it is not a coincidence that this quality is especially visible in studies that compare the teaching in countries with well-established research cultures in didactics, although the institutional mechanisms through which research may influence teaching practices vary considerably (for the example of France and Japan, see Miyakawa & Winsløw 2009, §2).

3. The role of ATD.

As pointed out in the introduction to this section, the corpus selected presents the particularity of consisting of doctoral theses prepared in French universities. Consequently, we could suppose that ATD was part of the theoretical background shared by their authors, unlike the situation for the comparative studies analysed in the previous sections. As shown by the analysis developed above, this difference has an evident impact on the way doctoral students define their problematics and in their methodological approach to comparison. In all these doctoral theses but one, the main theoretical framework is ATD. This theoretical choice induces that institutional comparison cannot be considered as part of the background data and analysis, as is the case in many other international studies. On the contrary, institutional comparison constitutes the core of the study. Moreover, the analysis shows that through the notion of praxeology and of didactic transposition, ATD has provided operational tools to these doctoral students for carrying out the institutional analyses they aimed at. There is no doubt also that this operationality has progressively increased as far as the same tools were productively used with similar purpose. When complemented with didactic constructs and epistemological analysis specific to the mathematical domains at

Kommentar [JK5]: I don't understand what the blank space here was meant to signify.

stake, as is the case in the theses of this corpus, such an approach results in interesting and strong institutional analyses, enlightening the conditions and constraints imposed on the teaching and learning of mathematics in the different contexts at stake, and the origin of these. This characteristic gives an evident originality to the corpus of theses when compared with the comparative studies analysed in the previous sections. It also shows that beyond the role given to it in this article, that of tool for the meta-analysis of comparative studies, it can be engaged directly as a productive tool for developing comparative studies. Let us add that the notion of level of didactic codetermination that we extensively use in this article is less present in the corpus. That is not surprising given that this notion is a more recent construct of the theory. That being said, we believe that many results of the comparative theses can be usefully rephrased using this notion, so one can anticipate its increasing use in comparative research inspired by ATD.

As pointed out in the analysis developed above, however, this strength in the institutional analysis goes along with some limitations. They concern, first, the level of the implemented curriculum, which is never addressed. Analysing the first subcorpus, we made the hypothesis that some characteristics of the Vietnamese context could partially explain this phenomenon. But such a hypothesis does not survive the analysis of the second corpus, where we observe the same phenomenon despite the diversity of contexts. The neglect of the implemented curriculum is most likely an effect of the methodological difficulty of combining strong and convincing complementary analyses within the constraints of a doctoral thesis. The choice of ATD as theoretical framework leads researchers to anchor comparative studies in institutional analyses, and consequently limits the place that can be given to other facets of the comparative work. Within these conditions, it is not surprising that we observe a relative neglect of the implemented curriculum to focus on the achieved curriculum, which seems methodologically more accessible. But, even regarding the achieved curriculum, as was pointed out in the analysis, the studies are generally small-scale and qualitative, and the interpretation of the results obtained within a comparative perspective remains thus hypothetical.

FINAL REMARKS

International comparative studies of regulations and practices of mathematics education come in many forms, and they are carried out with a great variety of purposes and means. In this paper, we have presented what seems to us a promising way to relate and contrast the domains that such studies investigate and compare, as well as the relations they establish or claim to find between them. Although the large-scale studies considered in the section on TIMSS and PISA (and to some extent also the section on large-scale international studies of classroom practice) have certainly drawn more public attention than the close-up studies of mathematics teaching in a few (often just two) countries considered in the preceding two sections, it seems clear to us that the deeper questions involved in comparison require the affordances of research with a sharp focus on particular contexts, including the institutional analysis of how mathematical domains, sectors, and themes are understood and articulated with each educational culture. This focus is, in particular, salient for the purpose of comparing contexts in which radical differences exist at every level (from civilisation to the way mathematical subjects are approached), and where the risk of drawing simplistic conclusion as to their mutual relations or impacts is almost inevitable without such a closer consideration of concrete conditions at every level. The model we have proposed in this paper offers a means of identifying missing elements in a given analysis but also of situating it with respect to studies that could supply those elements.

More specifically, we wish to point to four problematic observations emerging from our work with the research products analysed in this paper:

1. Although the integration of studies with data from several levels should certainly be possible, there is a tendency of most studies to rely entirely on their own data, perhaps as a consequence of particular aims but possibly also out of the lack of a framework that could help identify what other research it could be important to draw on.

2. There is a clear tendency in large-scale studies to average across the lower levels (below Level 5 in particular) in ways that are rarely justified by the data; moreover, these studies generally present snapshots, with little or no view of how conditions develop over time. On the other hand, researchers conducting binary, small-scale studies sometimes draw hasty conclusions for higher levels of codeterminations based on close-up studies of fairly peculiar contexts, with particular sectors or themes in focus and often also with a reference to the diachronic dimension. Researchers are clearly inclined to draw maximal conclusions from their own research, but combinations of data of the kind alluded to in the previous point may help close this gap between evidence and inferences.

3. Language issues—which affect the formulation of conditions at all levels—are rarely taken into serious consideration. Given the cultural and institutional situatedness of mathematical and didactical organisations in the context of teaching, this neglect may lead to very partial or even misconstrued conclusions that are not free of colonialist tendencies in the (frequent) case where the models of one culture are taken for granted as a pseudo-neutral reference.

4. As most studies are concerned with more than one level of codetermination, and indeed with hypotheses of how they interact with and determine each other, it is very important that these levels be explicit and that the aims of comparison—vertical and horizontal—are made clear. The way in which results are used and interpreted cannot be entirely controlled by researchers, but we should at least not be naïve when it comes to the massive interests that are attached to the objects of research—in the first place, at the level of society, but when it comes to the conclusions that may be drawn, at each and every of the other levels as well.

When studying didactical and mathematical practices in a cultural-institutional light—which is all the more inevitable when considering such phenomena in different institutional and cultural contexts—we are necessarily constrained by our own experiences and contexts. Although ATD as a whole is certainly not free of markings from the (European, more precisely Francophone) context in which it arose, we think that not only does it provide a useful means of identifying important assumptions and traditions well beyond its origins, but also its adaptation and development for such an enterprise will necessarily require collaborative efforts of the kind exemplified by some of the studies considered in this paper.

REFERENCES

BESSOT A., COMITI C. (2009) Apport des études comparatives aux recherches en didactique des mathématiques : le cas Vietnam / France. In Hache C., Coulange L. (Eds.) (pp.) *Actes du Séminaire National de Didactique des Matématiques.* Paris: IREM de Paris7.

Kommentar [JD6]: Complete pages

BOSCH M., GASCÒN J. (2006) 25 years of the didactic transposition. *ICMI Bulletin* 58 51–65.

CHEVALLARD Y. (1985) *La transposition didactique. Du savoir savant au savoir enseigné*. Grenoble: La Pensée Sauvage.

CHEVALLARD Y. (1999) L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en Didactique des Mathématiques* 19(2) 221–265.

CHEVALLARD Y. (2002) Organiser l'étude 3. Écologie & régulation. In Dorier J.-L., Artaud M., Artigue M., Berthelot R., Floris R. (Eds.) *Actes de la 11e*

école de didactique des mathématiques (pp.41–56). Grenoble: La Pensée Sauvage.
CLARKE D. (2003) International comparative research in mathematics education. In Bishop, A. et al., *Second international handbook of mathematics education* (pp. 143–184). Dordrecht: Kluwer.
CLARKE D. (2006) The LPS research design. In Clarke D., Keitel C., Shimizu Y. (Eds.) *Mathematics in twelve countries: The insider's perspective* (pp. 15–36). Rotterdam: Sense Publishers.
CLARKE D., MESITI C., O'KEEFE C., XU L., JABLONKA E., CHEE MOK I., SHIMIZU Y. (2007) Addressing the challenge of legitimate international comparisons of classroom practice. *International Journal of Educational Research* 46 280–293.
DUVAL R. (1995). *Semiosis et Pensée Humaine*. Paris : Peter Lang.
HOFE R. VOM (1995) *Grundvorstellungen mathematischer Inhalte*. Heidelberg: Spektrum Akademischer Verlag GmbH.
HOPMANN S., BRINEK G., RETZL M. (Eds.) (2007) *PISA according to PISA – Does PISA keep what it promises?* Wien: LIT Verlag.
HOUDEMENT C., KUZNIAK A. (2000) Formation des maîtres et paradigmes géométriques. *Recherches en Didactique des Mathématiques* 20(1) 89–116.
KAWANAKA T., STIGLER J., HIEBERT J. (1999) Studying mathematics classrooms in Germany, Japan and the United States: Lessons from the TIMSS Videotape Study. In Kaiser G., Luna E., Huntley I. (Eds.), *International comparison in mathematics education* (pp. 86–103). London: Falmer Press.
KEITEL C., KILPATRICK J. (1999) The rationality and irrationality of international comparative studies. In Kaiser G., Luna E., Huntley, I. (Eds.), *International comparison in mathematics education* (pp. 241–256). London: Falmer Press.
LANGE J. DE (1996) Using and applying mathematics in education. In Bishop A. et al. (Eds.) *International handbook of mathematics education,* (Part 1, pp. 49–97). Dordrecht: Kluwer.
LEUNG F., GRAF K.-D., LOPEZ-REAL F. (Eds.) (2006) *Mathematics education in different cultural traditions : A comparative study of East Asia and the West.* New York: Springer Science.
MIYAKAWA T., WINSLØW C. (2009) Un dispositif japonais pour le travail en équipe d'enseignants : étude collective d'une leçon. *Education & Didactique* 3 (1), 77–90.
MULLIS I., MARTIN M., RUDDOCK G., O'SULLIVAN C., ARORA A., ERBERBER E. (Eds.) (2005) *TIMSS 2007 assessment frameworks.* Boston: TIMSS & PIRLS International Study Centre.
MULLIS I., MARTIN M., GONZALES E., CHROSTOWSKI S. (2008) *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades.* Boston: TIMSS & PIRLS International Study Centre.
NISS M. (2003) Quantitative literacy and mathematical competencies. In MADISON B. L., STEEN L. A. (Eds.), *Quantitative literacy: Why numeracy matters for schools and colleges* (pp. 215–220). Princeton, NJ: National Council on Education and the Disciplines.
OECD (2004). *Learning for tomorrow's world: First results from PISA 2003.* Paris: OECD.

OECD (2006) *Assessing scientific, reading and mathematical literacy. A framework for PISA 2006*. Paris: OECD.

TRAVERS K. (1988) Introduction to the Second International Mathematics Study. In Robitaille D., Garden F. (Eds.), *The IEA Study of Mathematics II: Contexts and outcomes of school mathematics* (pp. 1–16). London: Pergamon Press.

VERGNAUD G. (1990) La théorie des champs conceptuels. *Recherches en Didactique des Mathématiques*, 10(2/3) 133–170.

WINSLØW C. (2005) Définir les objectifs de l'enseignement mathématique : La dialectique matières-compétences. *Annales de didactique et des sciences cognitives* 10 131–156.