RESEARCH IN MATHEMATICS EDUCATION THROUGH A KEYHOLE: TASK PROBLEMATIZATION

ANNA SIERPINSKA

This article developed out of a talk given at the 2003 Canadian Mathematics Education Study Group (CMESG) meeting [1].

There is a dream, among mathematics educators, for a common direction that would unite teachers, policy makers and researchers, with researchers collaborating to accumulate knowledge and thus contribute to our understanding and improvement of school practice. But this dream becomes less and less realistic in view of the tendency of mathematics education to multiply and diversify its epistemological foundations, theoretical approaches, methodologies and research questions.

> In the past, mathematics education researchers drew largely from psychological frameworks and theories, but contemporary researchers are increasingly demonstrating the insights that may be learned from additional frameworks. Central among the frameworks now drawn upon are sociology […], sociocultural theory […], politics […], mathematics […], philosophy […], history […], anthropology […]. But scholarship is anything but simple, and whilst we may agree that breadth of thinking is critical to the evolution of ideas, and that different frameworks should be considered and employed, we must also be wary that mathematics education is a relatively new and young field and that *too much breadth and diversity will cause a scattering of focus and preclude opportunities for consolidation and identity* […]. (Boaler, 2002, p. 7, *emphasis added*).

The "breadth and diversity" is growing not only through importing theories from neighbouring domains, but also as a result of the productivity of mathematics educators themselves.

> [I]t has become the norm rather than the exception for researchers to propose their own conceptual framework rather than adopting or refining an existing one in an explicit and disciplined way. This prolific theorising might be represented as the sign of a young and healthy scientific discipline. But it may also mean that theories are not sufficiently examined, tested, refined and expanded. A theory may be used mainly by its […] creators and their students rather than by a large number of independent and experienced researchers. It may be used for only one particular type of research study, of population, of methodology or of context. Equally, essentially the same issues and research questions may be being described and analyzed by a multiplying array of parallel theories. One concern may be that such proliferation of theories, influences and frameworks may lead to mathematics education becoming a tower of Babylon, where many strive, with excellent intentions, to provide light for their colleagues, but few listen, read and take into account their colleagues' ideas and work. (ESM Editors, 2002, p. 253)

In response to these concerns and challenges, mathematics educators started engaging in 'meta-studies' such as attempts at outlining a definition of the domain (*e.g.*, most papers collected in Sierpinska and Kilpatrick, 1998), more or less comprehensive syntheses of research approaches and theories (*e.g.*, Sierpinska, 1995; 1996; 2002; Niss, 1999), comparison of theories (*e.g.*, Boero *et al.*, 2002), and analyses of articles published in leading journals and conference proceedings over the past 10 or more years (*e.g.*, Boaler, 2002; Lerman *et al.*, 2002 [2]; Hanna and Sidoli, 2002).

My previous meta-studies focused mostly on theoretical frameworks and referred to a personally biased selection of papers. For the 2003 CMESG conference, I decided to take a more empirical approach and have a good look at a more or less 'random sample' of papers, in the sense that they were not specially selected to represent this or that issue, theme, theoretical perspective or mathematical topic. And thus I picked one of the volumes piled up on my desk; it happened to be *Volume 4* of the *PME 26 Proceedings* (Cockburn and Nardi (eds), 2002), containing 55 research reports.

Editors of the proceedings put the reports in alphabetical order with respect to the presenting authors' names; therefore there was no bias with regard to topics or research issues. At least from this point of view, the choice of the texts was random. But it is a very small sample, compared to the number of papers that are published worldwide. Moreover, the papers were written for a PME conference, which is a very special kind of conference, with its own interests and biases. Choosing papers published in a variety of journals, or presented at different conferences could produce very different results. This is why the most I could obtain from the study of the 55 PME research reports is a limited vision of mathematics education research, "through a keyhole", and not the 'big picture' of where mathematics education is going.

In reading the reports, I had several questions in mind:

- What was the research problem?
- What were the results of the research?
- What were the tools of the research?

In particular, were mathematical tasks used as tools in the research? How were they presented? Was their choice justified and discussed? Were these tasks *problematized*, *i.e.*, were alternative formulations or settings of the task taken into account and the possible reasons for and consequences of these variations studied or, at least, discussed?

A natural question would be: *In what kind of theory was the research grounded?* This is an important question, because addressing it in depth is necessary if we were to realize the dream of unification or systematization of our domain. But this job is presently being rather successfully done by other colleagues, acknowledged above, and therefore I will not devote much attention to the issue of theories in this article.

In this article, I will initially give an overview of the results from this empirical study, and then elaborate a little more on one particular issue (namely, *task problematization*). I will give examples of task problematization from the 55 PME papers, but I shall also go beyond the sample and speak more broadly on various approaches to this kind of work, known to me from other publications.

Some results from an analysis of the research reports

I will refer to the papers in the *Volume 4* of the *PME 26 Proceedings* either in general terms, without mentioning the particular papers, or by their ordinal number of appearance in this volume, from 1 to 55, and not by their authors' names. This is not to be interpreted as a sign of lack of respect for the authors' work but as my strategy of reinforcing the idea that the choice of texts has been rather arbitrary.

Research questions: overview

There were five categories of research questions according to whether they focused on the teaching of mathematics (24 reports, ~44%), the learning of mathematics (24, ~44%), research methodologies (5, ~9%), mathematics teacher education (1, ~2%) or methods and instruments of assessment (1, ~2%) (see Figure 1).

The focus on teaching and learning with little attention devoted to assessment and policy issues in mathematics education research has also been noted in Lerman *et al.* (2002).

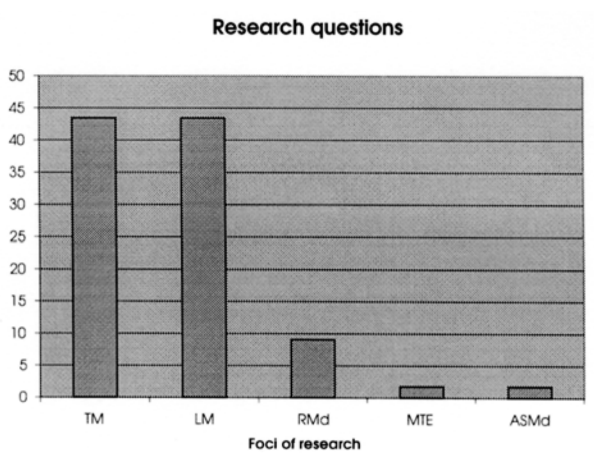

Figure 1: Foci of research questions (TM indicates teaching mathematics; LM, learning mathematics; RMd, research methods; MTE, mathematics teacher education and ASMd, assessment methods).

In the next two subsections, I will give some more details for the questions related to teaching and learning of mathematics.

Questions about the teaching of mathematics

There are, in principle, four approaches to questions about teaching a subject:

- *descriptive* (5 papers): How is the subject actually being taught?
- *analytic* (7 papers): What are the factors that influence teaching practices?
- *experimental* (12 papers): What happens if the subject is taught differently?
- *prescriptive* (1 paper): How best to teach the subject?

Ten out of the twelve teaching experiments were done on a small scale, involving a few students or a couple of classes, with four of the ten looking at the potential of technology in the teaching of mathematics.

Questions about the learning of mathematics

There were three kinds of questions about learning in the sample of articles:

- *theories* (8 papers): How people learn (a) mathematics in general (4 papers), (b) specific mathematical concepts or processes (4 papers)?
- *descriptions* (9 papers): How and/or how well a particular person or group of persons learned a particular mathematical idea?
- *analyses* (7 papers): What are the social, cognitive and/or affective factors that influence/interfere with the learning of mathematics in general or specific mathematical concepts or processes?

Among the papers devoted to theories about learning mathematics in general, two could be said to try "adopting or refining an existing [theory] in an explicit and disciplined way" (as if following the advice of the ESM Editors, 2002, p. 253). The author of one looked at a number of theories about cognitive development and proposed a (selective) synthesis of these for the purposes of explaining and predicting processes of mathematics learning (calling it the theory of "fundamental cycles of cognitive growth" (report 6). One of the theories was the SOLO Model (Structure of the Observed Learning Outcome; Biggs and Collis, 1982), which was also used in another paper (a resurge in popularity?). The other made an attempt at adopting, for mathematics education research, the neuro-psychological notion of "flexible thinking" as the "ability to use long-term declarative knowledge in novel situations" (report 45). The remaining two papers about theories of learning mathematics in general worked with 'home grown' conceptual frameworks developed by the authors themselves (sometimes in collaboration with colleagues).

Theories of learning specific mathematical concepts or processes were more interesting for the development of our domain, in that *they could not be obtained in any other domain* of knowledge such as neuropsychology or general sciences of education. These theories were based on analyses of specifically mathematical content and thinking involved in, for example, representing relations in word problems using algebraic equations, solving linear equations, avoiding the pattern of proportional thinking in solving certain probabilistic problems, and reasoning mathematically in justifying or explaining.

There were quite a few reports offering detailed descriptions of learning. I presume that the papers would not have been accepted for publication were they not couched in terms of a definite theory. Indeed, each of these papers contained a section about the "theoretical framework" of the research, making all the right references and introducing concepts, which were then sometimes (*sic*!) used in the descriptions.

Results of research: overview

Most papers reported on research in progress and their answers to the research questions were tentative. But researchers drew conclusions from their research, described the various tools of their research and these could also be considered as 'results'. I counted 71 such results in the papers, 35 of which I called 'conclusions', and 36 – 'products'. Some conclusions confirmed previously obtained results or common knowledge (12 or ~17%); some appeared to be new findings (13 or ~18%), and some were refutations of previously obtained results or common knowledge (10 or ~14%). Some products were 'practical' such as teaching proposals (12 or ~17%) and material products, *e.g.*, software or videos (4 or ~6%). Other products (20 or ~28%) were theoretical contributions to: theories of teaching and/or learning (14 or ~19%); methodology of research (4 or ~6%); epistemology of mathematics (1 or ~1%) and philosophy of mathematics education (1 or ~1%) (see Figure 2).

The relations between research questions and research results are represented in Figure 3. If I may risk a conjecture, this table suggests that the most interesting results about teaching, *i.e.*, refutations, are more likely to come from a search for factors influencing teaching rather than from descriptions of teaching or teaching experiments. However,

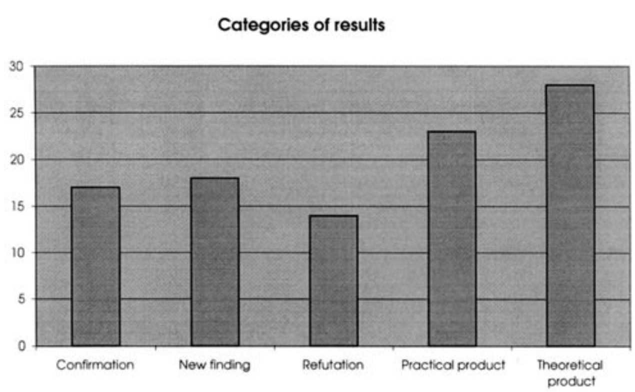

Figure 2: Categories of results.

descriptions of learning processes turned out to be quite fruitful in refutations. This could perhaps be explained by the fact that, in our domain, theories about learning are more developed than theories about teaching mathematics and therefore our descriptions of learning are more analytical, while descriptions of teaching are more phenomenological. A search for factors influencing teaching practices brings in an analytic approach that sharpens our understanding of the processes and yields more interesting results.

Confirmation results

Are large-scale studies that only confirm what we know from ordinary teaching experience worth the effort? An example of such common sense knowledge is: if a student has the necessary declarative knowledge, good judgment about mathematical arguments, and a scientifically oriented mind then he or she is better prepared for proving in mathematics than if he or she has important gaps in declarative knowledge, cannot tell a correct proof from an incorrect one, has a non-scientifically oriented mind and cannot solve even simple word problems related to plausible situations. This sounds trivial, yet the conclusions, from one study involving 669 grade 7 students in Germany, said not much more than that (report 15).

Refutation results

Results that put into question our previous beliefs or findings have a more revitalizing effect on a domain of research than those that merely confirm them. At least they stir some discussion and provoke an elaboration of our ways of thinking about a phenomenon.

In a study of 41 English high-school students' definitions and images for the concept of definite integral, it was found that students who learn the concept in a conceptually and experientially oriented curriculum (involving classroom discussions of conceptual ideas, experiencing them in use and going back to refine the conceptual ideas) seem to have no less difficulty with the concept than those who learn it through a procedural approach (report 12). Another paper reported a disappointment with a progressive curriculum (Realistic Mathematics Education in Dutch schools), which has not produced better performance in students, this time, in the expected practical, investigational skills using mathematizing (report 42).

It is unfortunately also not true that more complex assessment items necessarily lead to teachers doing more complex mathematics in class; buried is thus the hope that, by changing assessment, changes in teacher practices can follow. This finding suggests that, if 'teachers teach to the test', it is not because teachers will teach to any test, but because *tests reflect the ways teachers most commonly teach.* When a new type of assessment is introduced, after some years of its use, it is modified to fit with the teachers' practice better: tests are transformed to match teaching, not the other way round (report 25).

New findings

Many findings, if not exactly confirming some well-known statements, were not very surprising and modest in scope.

Identification of *factors* responsible for bringing about the desired cognitive behaviors of students figured in conclusions of several papers, but the identified social, emotional and cognitive factors were not new, and the evidence supporting the impact of the environmental factors was not very strong. For example, in one report, researchers compared two learning environments, for grade 8 students in Italy, that aimed to engage them in conjecturing and proving: one using only mechanical linkages and the other both the mechanical linkages and their Cabri-geometry models. The report contained a description of the behaviors of two pairs of students each working in one of the two environments. Conjecturing and correct proofs were obtained more easily in students who worked with both the mechanical and the computer models. From this observation the authors concluded that the use of Cabri made a "substantial contribution" (report 40) to students' engagement with conjectures and proofs. Thus, although the two learning situations were different in many ways, and only two pairs of students were observed, the authors pointed to the use of Cabri as the main factor of success of one of the pairs.

Mathematical tasks as tools in teaching and research

I use the expression *mathematical task* in a broad sense to refer to any kind of mathematical problem, with clearly formulated assumptions and questions, known to be solvable in predictable time by students. A mathematical task may be open (*e.g.*, not all information may be given; there may be many solutions or even interpretations), but it is not an unsolved research mathematics problem for which mathematics experts have not found a solution yet.

The notion of *mathematical task* has had a long and tortuous history in mathematics education. There have been debates related to the meaning of the term, and this meaning has changed over the years and with fashions in educational discourse. For some time it even seemed to be banned from researchers' vocabulary as if reflecting a backward philosophy of mathematics and mathematics education. It was replaced by expressions such as 'problem-situation', 'activity', 'didactical situation' and 'teaching/ learning situation'. The term *task* evoked the dark ages of *task analysis* and *task variables* in a mathematics education without the cognitive subject, never mind the child, the person, the student, the learner, the teacher, the socio-cultural group, or the community of practice (see *e.g.*, Goldin and McClintock, 1979, representing progress by mentioning "context" and "situation" variables, but still giving no account of students' reactions to the tasks whose variables were analyzed).

Questions about mathematical tasks reflect my own particular bias as a reader of mathematics education papers, researcher and practitioner. I consider the design, analysis and empirical testing of mathematical tasks, whether for the purposes of research or teaching, as one of the most important responsibilities of mathematics education.

Different tasks are needed for different purposes. Students' responses may be very sensitive to even small changes in a formulation of a task, or its mathematical, social, psychological and didactic contexts. This is why I think it is so important to justify the choice of the mathematical tasks used in a research, not just in terms of the general goals and theoretical framework of the research, but in terms of the *specific* characteristics of the task. A task may be set in different contexts, and formulated in different ways; it is important to be aware of *the possible variants* and reflect on the influence of the choice of a variant on teaching, learning or research results.

Questions ↓ **Results** →	C	F	R	TD	MT	E	P	MP	TP
Teaching practices	*32*	*32*	*52*	*18; 21; 43*					
Factors of teaching		*8; 29*	*8; 23; 33; 25*	*8; 27*					*8*
Teaching experiments.	*2; 4*	*38*	*17*					*54*	*2; 3; 4; 7; 19; 22; 28; 36; 49; 54*
How best to teach									*20*
Theory of learning, general				*6; 34; 45; 48*					
Theory of learning, specific concepts	*11; 39*			*11; 14; 41*					
Description of learning	*9; 26; 35; 37*	*35; 46; 55*	*1; 5; 12*			*26*		*5*	
Factors of learning	*10; 15; 51*	*16; 40; 47*		*24; 47*	*51*				
Methodology		*53*			*13; 31; 53*		*44*	*50*	
Mathematics Teacher Education		*30*						*30*	
Assessment		*42*	*42*						

Figure 3: Table showing research questions and results (C indicates confirmation; F, new finding; TD, theory development; MT, methodological tools; E, epistemology of mathematics; P, philosophy of mathematics education; MP, material product and TP, teaching proposal).

Research overview: mathematical tasks

Of the 55 *PME 26* research reports, 47 used one or more mathematical tasks, for various purposes. The tasks were stated, but they were also sometimes justified and problematized to some degree. In classifying these reports, I used the following codes:

11TJ: The given task is justified, *i.e.*, it is explained how the task allows pedagogical or research goals to be reached. The justification is based on explicit principles or theory and refers to the specific features of the task.

10TJ: A general rationale for the task is provided, *i.e.*, it is explained what goals the task was expected to achieve, but it is not explained exactly which features of the particular task were essential in achieving these goals.

00TJ: The task is not justified.

1TP: The task is problematized, *i.e.*, variations on the task are debated and there is a discussion of the effects of such variations on the learning or on the research results.

0TP: The task is not problematized.

Figure 4 shows a table giving the frequencies of the six categories of presentations in the reports, in absolute values of number of papers and percentages relative to the 47 papers with tasks. Thus 33 (70%) of the 47 reports contained at least some justification of the tasks, but only 18 (38%) problematized the tasks. However, a complete, specific justification and problematization of tasks was found only in 14 (30%) of the reports.

	TP		
TJ	0	1	Totals
00	13 (28%)	1 (2%)	14 (30%)
10	10 (21%)	3 (6%)	13 (28%)
11	6 (13%)	14 (30%)	20 (43%)
Totals	29 (62%)	18 (38%)	47 (100%)

Figure 4: Table showing frequencies of different task presentations in all forty-seven reports with tasks.

Of course, you might say that it is not possible to justify fully the choice of a task in a short, eight-page report for a conference. This is true, but the fact that the authors, in making choices about what to write about in a report, massively decided to sacrifice exactly this justification, does say something about what we value and what we do not in the mathematics education research community.

In the 47 reports with tasks, 40 (85%) used non-routine tasks; 31 (66%) referred to out-of school contexts; in 35 (75%) the tasks were exploratory; 30 (64%) allowed for multiple interpretations and 29 (62%) were to be done in a computer environment.

In the remaining part of this article I will focus on the notion of task problematization.

Examples and a non-example of task problematization

Task problematization must be distinguished from the problematization of the procedures of assessment of students' solutions of tasks. Confusion is possible because task problematization is often triggered by problems related with grading students' solutions (*e.g.*, ambiguity or impossibility to discriminate between various levels of understanding or performance). For example, in an analysis of the results of Dutch students on certain TIMSS items (report 42) the authors were putting into question the procedures for marking students' solutions: they noticed, for example, that on some items the inter-scorer agreement on the correctness of answers was as low as 52%. But they did not discuss the particular features of the items themselves, and did not study the effects that possible variations of these items could have on the results.

On the other hand, the "discussion" part of a study on students' erroneous procedures in solving algebraic inequalities, especially their tendency to divide both sides of an inequality by a not necessarily positive parameter (report 37), contained a thorough problematization of the following set of tasks used in the research:

Task I. Examine the following claim: for any a in R, $ax < 5 \Rightarrow x < 5/a$;

Task II. Examine the following statement: for any $a \neq 0$ in R, $ax < 5 \Rightarrow x < 5/a$;

Task III. Solve the inequality $(a - 5)\,x > 2a - 1$ where x is a variable and a is a parameter.

The choice of the tasks was *problematized* in the sense of a discussion of possible alternatives to this choice and their consequences for the research. The authors argued how misleading it would have been if only *Task I* was given to the students. Here is the beginning of this detailed discussion:

> [E]xamination of […] solutions to Task I revealed almost no intuitive, erroneous ideas. Most students correctly rejected the statement […] [M]any students used zero as their counterexample […] Had we stopped here, we might have assumed that most students have a good formal understanding of such parametric inequalities. However, an examination of their responses to Tasks II and III revealed that this was not the case. […] (report 37)

In mathematics education, from time to time, a task appears that has an international career. The task is used in research by researchers working independently. The interesting thing is that there are normally several variations on the task. Analyzing the nature of these variations and their effect on students' reactions to them is part of the work of problematization of the task and would be a subject for yet another independent research study. In the 55 PME 26 reports, the following interesting task appeared in two versions:

Version 1: [In Cabri-geometry II] Draw a quadrilateral ABCD. Construct the perpendicular bisectors of each

> side. Label the four points (say, M, N, P, Q) at the intersection of the perpendicular bisectors from adjacent sides. If you move ABCD what happens to the inner quadrilateral MNPQ? (a) Investigate the relationship between the internal angle at A and the internal angle at M. (b) For what types of external quadrilateral is MNPQ a parallelogram? (c) Find a quadrilateral ABCD that is similar [to] its inner quadrilateral MNPQ, i.e. an enlargement/reduction. (report 28)
>
> *Version 2:* [In Cabri-geometry II] You are given a quadrilateral ABCD. Construct the perpendicular bisectors of its sides: a of AB, b of BC, c of CD, d of DA. H is the intersection point of a and b, M of b and c, L of c and d, K of a and d. Investigate how HMLK changes in relation to ABCD. Prove your conjectures. (report 2)

Multiple interpretations of the situation are possible in both versions. In *Version 1*, no assumptions about the quadrilateral ABCD are made, therefore different students could start from a very different initial figure (someone could draw a non-convex quadrilateral) and this could lead to different paths of reasoning. A concave ABCD is implicitly ruled out by talking about "inner" and "external" quadrilaterals, but this occurs only in question (c). Also, the mutual position of points A and M is subject to multiple interpretations, so students may be answering different questions (b).

On the other hand, while the task in *Version 2* still contains the possibility of multiple interpretations with respect to the quadrilateral ABCD, there is no longer the ambiguity of the positions of the vertices of the quadrilateral obtained from intersections of perpendicular bisectors. It would be interesting to investigate the consequences of these differences for research, for students' learning, and for assessment purposes.

Possible approaches to task problematization

I believe that 'task' is a useful notion in mathematics education research. It cannot be replaced by 'problem' if we assume a 'law' of economy of intellectual effort: students will not invest more intellectual effort than strictly necessary for completing the mathematical task given to them. I mean "completing the task" and not "solving a problem" because students – *qua* students – complete their tasks; they do not solve problems. If a person engages in solving a problem in our classroom, this person no longer acts as a student – a subject of an educational institution and our subject – but as an independent intellectual. Therefore, in analyzing a task, we are asking not what kind of mathematical knowledge and thinking it makes *possible*, but what kinds of mathematical knowledge and thinking are *strictly necessary and sufficient to complete the task.* If these are not what we aim at in giving the task to the students, then we have the problem of identifying the aspects of the task that were responsible for that.

Task problematization as a by-product of research

A situation where a task given to students or research subjects fails to satisfy our expectations forces us to engage with the work of task problematization. This situation often takes us by surprise and the work is undertaken reluctantly, as a punishment for not having known better prior to the lesson or study. Problematization is then a by-product of research rather than an explicit aim in an important and planned phase of it. It is the messy part of doing research, and there is no place for it on the neatly arranged plate of final results, served in a published report.

Task problematization in teaching design

For mathematics educators who are engaged in the design of educational materials, problematization of tasks is a daily routine. But if they consider design as 'development' and not 'research', they do not engage in writing papers about their task problematization experience.

There is, of course, no reason why design should not count as research (see, *e.g.*, Wittmann, 1995, for a sound theoretical argument). There are individuals who are working under this assumption (*e.g.*, Sinclair, 2003) and there are whole teams and communities of researchers who do that, too, having developed solid theoretical foundations for the work of task problematization. I am thinking, in particular, of *Realistic Mathematics Education* research (see, *e.g.*, Presmeg and van den Heuvel-Panhuizen, 2003) and the *Theory of Didactical Situations in mathematics* (TDS) (Brousseau, 1997) underlying the design methodology called "didactical engineering". In didactical engineering based on TDS, research is particularly focused on the construction and empirical verification of tasks for students, where the tasks are taken not as isolated mathematical problems but part of the complex structure of the classroom "milieu" (Margolinas, 1998; Bloch, 2002). The so-called 'a priori analysis' of a planned classroom situation is a justification, in very detailed and specific terms, of all relevant choices made with regard to the mathematics aimed by the tasks. Already at this point, tasks are problematized. They are further problematized during the 'a posteriori analysis', done after an experimentation in class (see, *e.g.*, Maurel and Sackur, 2002).

Task problematization in describing ordinary teachers' practices

Task problematization as a research problem need not occur only in the context of design of classroom activities and materials. One can also identify and problematize tasks used in everyday mathematics teaching in ordinary mathematics classes. This kind of study is at the center of interest of researchers working within the "anthropological approach" whose main ideas were developed by Chevallard (1999; 2002). In this research program, task problematization is part of studying and describing practices related with the teaching and learning of mathematics in various institutions. The results of such studies are "praxeologies" or theories of practices.

From this perspective, a mathematical task, given to students in a specific type of school institution, is part of the practice of a teacher who chooses it to meet certain curricular goals, but whose choice is heavily constrained by decisions made at the level of the institution. The identification of these institutional task variables and their impact on what students eventually learn is an important part of the study of teachers' praxeology.

Task problematization in the study of educational archives

The study of the archives of mathematics education, in particular, curricular documents, textbooks, tests, students' exercise books are a mine of information for anybody interested in educational praxeology. It is quite fascinating, for example, to compare the past and more recent school-leaving tests, looking at the relation between the kinds of thinking and mathematical knowledge actually necessary and sufficient for their solution and the grand curricular goals of teaching mathematics.

It could turn out that the curricular objectives of "mathematization", "modeling", "problem-solving" or "study of situations", translate into test questions with long stories about real life characters involved in real life problems, but whose solution requires very little mathematization or modeling activity. The student is told which variables to look at and often an equation describing the relationship between them is explicitly given. A doctoral student of mine, Jakobsson-Åhl from Sweden, is presently engaged in a study of this kind. Comparing two national school leaving tests, one from 1973 and the other from 2003, she wrote recently:

> From what I have seen so far, [...] the differences between the tests are not very deep. For example, just as, in 1973, the formal aspects of mathematical activity were very superficial and not really a solution to students' problems, so were the mathematical modeling aspects in 2002. Students were spectators of formalization in 1973 and spectators of applied mathematical modeling in 2002. Models were given, they did not have do be constructed by students.

In the work of task problematization you can take into account little or none of the classroom context in which students would work on it (which is unavoidable in historical studies of tests like the one mentioned above), or make the material, social and cultural contexts an inseparable part of the task (as in Maurel and Sackur's study). Both have their merits.

Conclusions

Theory – breadth without depth. In my analysis of the 55 research reports, I have not focused specifically on the underlying theoretical frameworks. But it was difficult not to notice that theory occupied a lot of space. A great multitude of theories were mentioned in the papers and contributions to theory constituted about 19% of the results. However, theories mentioned in the introductions or 'theoretical frameworks' sections did not always play an essential role in the research. Of course, in principle, a consistent use of a theoretical framework to, say, analyze a teaching episode, does make a difference in what is seen and how (see, *e.g.*, Steinbring, 1998; or Even and Schwartz, 2003). The problem was, however, that theories mentioned in the 'theoretical framework' appeared to perhaps inspire the research but they were not necessarily consistently applied in the research. It would be interesting to conduct an empirical study to describe the role that theories, mentioned in the 'theoretical framework' sections of mathematics education papers actually play in the research. As mentioned in the introduction, work in this direction has already begun.

Weakness of results. For all the theory production in mathematics education, conclusions from research remain shaky or weak. Many reports merely confirmed that abstract mathematical thinking is difficult for many students. Do we really need more empirical evidence for this? Perhaps we could now use what we know about the distribution of thinking styles among students in designing our curricula, textbooks and teaching methods so that they reflect this distribution rather than the distribution that the "mathematicians in us" dream about (Amitsur in an interview, Sfard, 1998).

Only a few results obtained in the 55 research reports could count as strong, that is, both interesting (surprising) and well justified or documented. Many results were teaching proposals whose value is mainly inspirational, for teaching and for research. But they often did not contain information that would allow the reader to judge what the essential variables of the teaching situations were. Moreover, few researchers took into account the constraints of the teacher's work and factored these constraints into their teaching proposals.

Scarcity of studies on assessment: We also need to be more constructive (rather than merely constructivist) in producing and putting to the test the results that bridge theory and practice. One area in dire need of this kind of constructive research is assessment. This activity is one of the main activities of a teacher, and it is also the least pleasant and the most difficult one. Without a serious consideration of the problem of assessment, most innovative approaches and teaching proposals never make it through the phase of initial experimentation by enthusiastic researchers and committed teachers. Yet, relatively to other issues, very little research has been devoted to assessment in mathematics education.

Call for a come-back of task analysis: No matter what emotive meanings mathematical tasks evoke and what they are called, they remain the very fabric of both mathematics teaching and learning and research. Mathematical tasks can be regarded as tools of research on a par with methodological tools such as statistics or coding schemes for qualitative data analysis. A teaching proposal contains tasks, and tasks are also needed to probe the effectiveness of a teaching approach. Studies in the psychology of mathematics learning require specially designed tasks.

In the sample of research publications I looked at, people were doing interesting teaching experiments but they chose to say very little about the reasons underlying the choice of the tasks and what difference it would make if they changed the tasks in this or that respect. This puts into question the possibility of their reproduction in different teaching or research conditions, because it is not known in the tasks what is arbitrary and what is essential and in what way.

Tasks in the reports were mostly 'exploratory', or part of initial activities used to introduce a new idea. This choice is understandable from the point of view of research, but then this research is very far from the everyday concerns of the teachers. If we assume that knowing is based on a system of connections, then we must allow students the time to build and consolidate these connections. The reports are silent on the problem of the construction of systems of tasks that aim at consolidation of knowledge.

In general, the tasks were not the kind of tasks that would be used in day-to-day teaching or assessment. In research, we need tasks capable of revealing students' most hidden conceptions and misconceptions; we do not merely test students' knowledge; we *put it to the test*, we probe it, we want to know all about it. In school assessment we use tasks that allow students to prove *that* they have a sufficient grasp of the basic required material and skills. We are not even interested in knowing too much about *what* each particular student knows and even less in *how* he or she knows it.

Two aspects of teaching could explain this: time constraints (evaluation of several classes of students in our charge must be done in a timely fashion) and the fact that school is interested in proving its success and not its failure. Research, on the other hand, has a lot of time to analyze students' work and is interested in proving that 'traditional' teaching (any current or ordinary teaching is deemed 'traditional') is ineffective and therefore in need of reform. This justifies research. Each institution defends its reason of being.

The need for analyzing and problematizing tasks could be explained from the point of view of the methodology of research, but also from the point of view of our relations with two professional groups, teachers and mathematicians. These relations are not always good (Bartolini Bussi and Bazzini, 2003; Goldin, 2003) and there is a lot of misunderstanding about what research in mathematics education is actually doing. But mathematical problems, of which tasks for learners constitute a subset, are a common ground on which all these three groups meet.

Mathematics education could substantially contribute to the discussion, if only it took the problem of tasks and task variables seriously again, capitalizing on all that has been found in the last twenty to thirty years about students' learning and teaching practices.

A final remark

We need a lot more critical studies that would question and find weak points in our research methodologies and theories, as well as conclusions and other products of research. We need to be cognitively stimulated in mathematics education research. There needs to be more debate, more sharp criticism, more hard analytic thinking about the phenomena we study and about the validity of our claims about them.

Notes

[1] For the original version of this talk, "Research in mathematics education through a keyhole", I produced a detailed analysis of the reports, with a justification of my qualitative and quantitative claims. It is included in the Proceedings of the CSMESG (Canadian Mathematics Education Study Group) 2003 conference in Wolfville, Nova Scotia, and will be available as a pdf-file on the FLM web-site http://www.flm.math.ca.

The expectations of the program committee, worded by one of its members, were:

> I would love it if she were to talk about the privileging and under-privileging of different forms of research in math. ed. […]. I'd love to hear her talk on that kind of meta-level, of what we need math. ed. research to try to be. I want to be stirred up, angered, forced to recognize how partial my [or rather the standard] outlook is […]. Invite Anna to tell us if there's anything resembling a big picture out there, for math. ed. research. Or is it all just a tug-of-war over curriculum between the Good Guys and the Bad Guys, with the good-hearted but ineffectual math. education researchers spiraling around the power brokers, like the electrons in an atom of uranium?

[2] A collection of papers by Lerman and his colleagues related to this project is available from: http://myweb.lsbu.ac.uk/~lermans/ESRProjectHOMEPAGE.html.

References

Bartolini Bussi, M. G. and Bazzini, L. (2003) 'Research, practice and theory in didactics of mathematics: towards dialogue between different fields', *Educational Studies in Mathematics* **54**(2-3), 203-223.

Biggs, J. and Collis, K. (1982) *Evaluating the quality of learning: the SOLO taxonomy (Structure of the Observed Learning Outcome)*, New York, NY, Academic Press.

Bloch, I. (2002) 'Différents niveaux de modèles de milieu dans la théorie des situations', in Dorier, J.-L., Artaud, M., Artigue, M., Berthelot, R. and Floris, R. (eds), in *Actes de la 11e École d'Été de Didactique des Mathématiques* – Corps – 21-30 Août 2001, Grenoble, France, La Pensée Sauvage éditions, pp. 125-139.

Boaler, J. (2002) 'Exploring the nature of mathematical activity: using theory, research and 'working hypotheses' to broaden conceptions of mathematical knowing', *Educational Studies in Mathematics* **51**(1-2), 3-21.

Boero, P., Dreyfus, T., Gravemeijer, K., Gray, E., Hershkovitz, R., Schwartz, B., Sierpinska, A. and Tall, D. (2002) 'Research Forum 1. Abstraction: theories about the emergence of knowledge structures', in Cockburn, A. and Nardi, E. (eds), *Proceedings of the 26th Conference of the International Group for the Psychology of Mathematics Education*, Norwich, UK, School of Education and Professional Development, University of East Anglia, **1**, pp. 133-140.

Brousseau, G. (1997, eds and trans., N. Balacheff; M. Cooper; R. Sutherland; V. Warfield) *Theory of didactical situations in mathematics*, Dortrecht, The Netherlands, Kluwer Academic Publishers.

Chevallard, Y. (1999) 'L'analyse des pratiques enseignantes en théorie anthropologique du didactique', *Recherches en Didactique des Mathématiques* **19**(2), 221-266.

Chevallard, Y. (2002) 'Organiser l'étude. 1. Structures et fonctions', in Dorier, J.-L., Artaud, M., Artigue, M., Berthelot, R. and Floris R. (eds), *Actes de la 11e École d'Été de Didactique des Mathématiques* – Corps – 21-30 Août 2001, Grenoble, France, La Pensée Sauvage éditions, pp. 3-22.

Cockburn, A. and Nardi, E. (eds) (2002) *Proceedings of the 26th Annual Conference of the International Group for the Psychology of Mathematics Education* , Norwich, UK, School of Education and Professional Development, University of East Anglia.

ESM Editors (2002) 'Reflections on Educational Studies in Mathematics', *Educational Studies in Mathematics* **50**(3), 251-257.

Even, R. and Schwartz, B. (2003) 'Implications of competing implications of practice for research and theory in mathematics education', *Educational Studies in Mathematics* **54**(2-3), 283-313.

Goldin, G. (2003) 'Developing complex understandings: on the relation of mathematics education research to mathematics', *Educational Studies in Mathematics* **54**(2-3), 171-202.

Goldin, G. and McClintock, C. (eds) (1979) *Task variables in mathematical problem solving*, Columbus, Ohio: ERIC Clearinghouse for Science, Mathematics and Environmental Education.

Hanna, G. and Sidoli, N. (2002) 'The story of ESM', *Educational Studies in Mathematics* **50**(2), 123-156.

Lerman, S., Xu, G. and Tsatsaroni, A. (2002) 'Developing theories of mathematics education research: the ESM story', *Educational Studies in Mathematics* **51**(1), 23-40.

Margolinas, C. (1998) 'Le milieu et le contrat, concepts pour la construction et l'analyse de situations d'enseignement', in Noirfalise, R. (ed.), *Analyse des pratiques enseignantes et didactique des mathématiques, Actes de l'Université d'Été de La Rochelle (4-11 juillet 1998)*, IREM de Clermont-Ferrand.

Maurel, M. and Sackur, C. (2002) 'La Presqu'île. Une introduction aux fonctions de deux variables en DEUG. Analyse en termes de structuration du milieu d'une situation de classe ordinaire', in Dorier, J.-L.,

Artaud, M., Artigue, M., Berthelot, R. and Floris, R. (eds), *Actes de la 11e École d'Été de Didactique des Mathématiques – Corps* – 21-30 Août 2001, Grenoble, France, La Pensée Sauvage éditions, pp. 167-175.
Niss, M. (1999) 'Aspects of the nature and state of research in mathematics education', *Educational Studies in Mathematics* **40**(1), 1-24.
Presmeg, N. and van den Heuvel-Panhuizen, M. (eds) (2003) 'Realistic Mathematics Education Research: Leen Streefland's work continues: a PME Special Issue', *Educational Studies in Mathematics* **54**(1).
Sfard, A. (1998) 'A mathematician's view of research in mathematics education: an interview with Shimshon A. Amitsur', in Sierpinska, A. and Kilpatrick. J. (eds), *Mathematics education as a research domain: a search for identity. An ICMI Study*, Dortrecht, The Netherlands, Kluwer Academic Publishers, pp. 445-458.
Sierpinska, A. (1995) 'Some reflections on the phenomenon of *French didactique*', *Journal für Mathematik-Didaktik* **16**(3-4), 163-192.
Sierpinska, A. (1996) 'Whither mathematics education?', in Alsina, C., Alvarez, J., Niss, M., Pérez, A. and Sfard, A. (eds), *Proceedings of the 8th International Congress on Mathematics Education*, Seville, Spain, 14-21, July 1996, pp. 21-46.
Sierpinska, A. (2002), 'Perspectives sur les recherches en didactique des mathématiques', *Zentralblatt für Didaktik der Mathematik* **34**(4), 164-174. (English version available on the web: http://alcor.concordia.ca/~sierp/downloadpapers.html.)
Sierpinska, A. and Kilpatrick, J. (eds) (1998) *Mathematics education as a research domain: a search for identity. An ICMI Study,* Dortrecht, The Netherlands, Kluwer Academic Publishers.
Sinclair, M. (2003) 'Some implications of the results of a case study for the design of pre-constructed, dynamic geometry sketches and accompanying materials', *Educational Studies in Mathematics* **52**(3), 289-317.
Steinbring, H. (1998) 'From "Stoffdidaktik" to social interactionism: an evolution of approaches to the study of language and communication in German mathematics education research', in Steinbring, H., Bartolini Bussi, M. G. and Sierpinska, A. (eds), *Language and communication in the mathematics classroom*, Reston, VA, NCTM, pp. 102-119.
Wittmann, E. (1995) 'Mathematics education as a 'design science'', *Educational Studies in Mathematics* **29**, 355-374.

I suggest to you, now, that the word "idea," in its most elementary sense, is synonymous with "difference." Kant, in the *Critique of Judgement* – if I understand him correctly – asserts that the most elementary aesthetic act is the selection of a fact. He argues that in a piece of chalk there are an infinite number of potential facts. The *Ding an sich,* the piece of chalk, can never enter into communication or mental process because of this infinitude. The sensory receptors cannot accept it; they filter it out. What they do is to select certain *facts* out of the piece of chalk, which then become, in modern terminology, information.
I suggest that Kant's statement can be modified to say that there is an infinite number of *differences* around and within the piece of chalk. There are differences between the chalk and the rest of the universe, between the chalk and the sun or the moon. And within the piece of chalk, there is for every molecule an infinite number of differences between its location and the locations in which it *might* have been. Of the infinitude, we select a very limited number, which become information. In fact, what we mean by information – the elementary unit of information – is a *difference which makes a difference,* and it is able to make a difference because the neural pathways along which it travels and is continually transformed are themselves provided with energy. The pathways are ready to be triggered. We may even say that the question is already implicit in them.

(Bateson, G. (2000, first edition 1972) 'Form, substance and difference', in *Steps to an ecology of mind,* Chicago and London, The University of Chicago Press, p. 459.)