

Questioning the world by questioning the exam

BRITTA EYRICH JESSEN¹

Resumen. En este trabajo se analiza el papel del examen y su impacto en la enseñanza de las matemáticas utilizando el concepto de niveles de codeterminación didáctica de la teoría antropológica de lo didáctico. Se argumenta sobre la necesidad de discutir el papel del examen en relación con los recorridos de estudio e investigación si se quiere cambiar el actual paradigma pedagógico, de la “visita de monumentos” al paradigma del “cuestionamiento del mundo”.

Abstract. This paper discusses the role of the exam and its impact on the teaching of mathematics using the notion of levels of codetermination from the anthropological theory of the didactics. It is argued that a discussion of exam in relation to study and research paths is crucial in order to succeed in changing the current teaching paradigm from “visiting monuments” to “questioning the world”.

1. Introduction

Researchers in mathematics education have argued for decades that common classroom activities in mathematics lack the potential to teach students more coherent mathematical knowledge, autonomous mathematical thinking and creativity based on problem posing activities (e.g. Singer, Ellerton, & Cai, 2013). Instead, common teaching activities lead students to imitate the teachers’ or textbooks’ solution strategies for standard exercises (Bergqvist, 2004). Alan Schoenfeld (1988) characterises common practice of mathematics teaching as activities where a rule is presented by the teacher and to be subsequently rehearsed by the students in standard exercises. Within the Anthropological Theory of the Didactics (ATD) this kind of teaching has been characterised as “visiting monuments” (Chevallard, 2015). The students are presented with a monument (e.g. the Pythagorean Theorem, the division algorithm etc.) and how it can be used in order to solve exercises. As such, mathematics teaching is similar to a tourist excursion with a guide who explains or highlights the specific features of the monuments in order for the visitors to appreciate it. This monumentalistic teaching is often aiming at preparing students for a high stake exam, which raises the attention to the monuments but often have a negative influence such as surface oriented strategies developed to cope with external demands. When final exams consist of standard exercises which the students encounter progressively and passively in the teaching, the exam defines the monuments of the teaching – often more than the official curriculum and even textbooks. The result of this kind of teaching is problematic in the sense that “students are almost reduced to mere spectators” (Chevallard, 2015, p. 3).

In ATD, another teaching paradigm has been proposed: the paradigm of “questioning the world”. The idea is that teaching situations should be based on a generating question that leads

¹ Department of Science Education, University of Copenhagen, Denmark – britta.jessen@ind.ku.dk

students to engage in posing questions and develop answers for these questions. The answers should be developed through a deconstruction and reconstruction of knowledge based on students' praxeological equipment, search and investigation of media, etc. Principles for designing this kind of teaching have been named study and research courses or lately study and research paths (SRP) (Chevallard, 2006, 2015). The potentials of the dynamics between students' research and study processes have been discussed and experimented in a wide range of contexts (see García & Hiugeras, 2005; Serano, Bosch, & Gascón, 2010; Barquero, Bosch, & Gascón, 2013; Jessen, 2014; 2015; Rasmussen, 2016; Florensa, Bosch, & Gascón, 2016). The experiments reported on in these papers show the potentials of using SRPs as design tool. Several designs address modelling explicitly, usually involving different kinds of functions. The goals of the experiments range from modelling with linear and exponential functions or with differential equations, engaging in bidisciplinary work, developing teachers' professional knowledge, learning statistics and even improving engineering education. However not much research discusses the relation between SRP and high stakes examinations. For some of the above mentioned studies, it might not be relevant in the context. Nevertheless, if the paradigm of questioning the world should become a viable alternative to the paradigm of visiting monuments in contexts where high stakes examinations form part of the institutional constraints, we need to address how these examinations could work with the new paradigm.

In fact, the two studies by Jessen (referenced above) relate to assessment in the context of exam situations as they are prescribed by existing constraints and condition of Danish secondary mathematics teaching. Below it is pointed out what can be gained from these experiments, keeping in mind that students and teachers tend to return to the prevailing teaching paradigm even after a successful implementation of SRP based teaching, as pointed out by Berta Barquero and Marianna Bosch (2015, p. 268).

We note here that in the general mathematics education literature, there is no shortage of papers which document the impact exam and assessment has on teaching (e.g. see Webb, 1992). The present paper intends to analyse, from such a broader perspective, the need for addressing the relationship between the teaching paradigm of "questioning the world", SRP based teaching and different forms of high stakes exams. Specifically the paper addresses the research question:

What is needed from school systems in general and exams in particular for this new paradigm of "questioning the world" to be a viable alternative to the paradigm of "visiting monuments"?

The mechanisms of "backwash" of exam requirements on teaching and learning practices are analysed using the notion of levels of codetermination from ATD.

2. Why discuss exams at all?

In the ATD context there is not much mentioning of exams as part of a changed teaching paradigm. However, the scale of levels of codetermination can be used to indicate where potential constraints and conditions for the teaching stem from. We recall that the scale consists of the following levels (Chevallard, 2002):

civilisation – society – school – pedagogy – discipline –
domain – sector – theme – question

As we will see in this section, the higher levels of codetermination affect the exit examinations and through them, the classroom activities and teaching paradigms. In this section, a minor

literature review is presented, which indicates that exams must be taken into account if a change of paradigm should become reality.

2.1. A long term challenge

For the last 200 years or so, school examinations have been regarded alternatively as encouragement, grading and “outright corruption in allocation of scarce opportunities” (Cheng & Curtis, 2004, p. 7). Ever since written exams were introduced in Boston schools in 1845, it has been discussed what these tests are measuring and in what sense (if any) they are objective in some sense. Throughout the 20th century, psychologically founded studies have investigated effects and results of testing (Romberg, 1987). In the 1980s and 90s more research was done on the impact of exams on the teaching and learning of mathematics. Schoenfeld reports on a study where a class was well taught, but the students still lacked justification of the techniques they used for solving problems: “By virtue of obtaining the correct answer, the students indicated that they had mastered the procedures of the discipline. However they had clearly not mastered the underlying substance” (Schoenfeld, 1988, p. 148). The study shows that the teacher kept lectures to a minimum and time was spent on students working with problems closely related to curriculum. It is further noted that the

Examination is the primary measure of both teacher and student success [...] [and further that] Examinations were well established and quite consistent from year to year. Thus the amount of attention to give to each topic, and way to teach it (for “mastery” as measured by the exam), were essentially prescribed (Schoenfeld, 1988, p. 153).

This is an example of the exam defining *de facto* the monuments of the teaching and thereby the actual contents of the teaching itself. The incentive for the teacher to let the test affect the teaching is clear; the teacher’s performance is often measured by the students’ results. Another study by Norman Webb (1992) shows that “large-scale assessment has an influence on what is taught in classroom [...]” (p. 678). The examination becomes a crucial constraint of the teaching since both students and teachers are measured based on the scores from the exam. Webb discussed purposes of assessment, and regarding final exams he points out that:

[the exam] is to provide information to decision makers, including those within the educational system, governmental policy makers, and others. At this level, assessment results are used by parents, administrators, school boards, and taxpayers as the basis of judgements about the effectiveness of the educational program in general and, in some cases about the relative skill and ability of individual teachers [...] (Webb, 1992, p. 663).

This was repeated by Christine Suurtamm and colleagues (2016) in the commission for the topic group discussion of assessment at ICME13. Suurtamm et al. (2016) emphasise that the exam is most certainly a strong part of the constraints put on teaching from civilisation, society, and school (the higher levels of co-determination). Webb points out the role (and constraints) played by society by claiming that “public has, in large part, been resistant to reform efforts to expand assessment beyond norm referenced testing [...]” (Webb, 1992, p. 664). Hence, these influences of society in general might cause challenges for changes at the level of pedagogy and be at the root of the resistance reported by Barquero & Bosch (2015).

The literature referred above mainly discusses the situation of national exit exams, however the backwash effect of exams are not restricted to mathematics education nor restricted to

national mathematics examinations (Cheng & Curtis, 2004; Bergqvist, 2007). Ewa Bergqvist investigated what kind of reasoning is needed for good performances at the exam of introductory courses on calculus at the beginning of undergraduate programmes at four Swedish universities. It is argued that emphasis is put on imitative reasoning and the use of algorithmic procedures in exam exercises, and this can prevent students from learning relational reasoning. Reasons for this are primarily the impact of assessment on teaching and instruction focused on relational reasoning is time-consuming, according to teachers (Bergqvist, 2007).

In the Danish context similar findings have been reported: “it seems that the circumstances and traditions governing university mathematics teaching make it difficult to assess more than the use of standard techniques or the passive knowledge of textbook material” (Grønbæk, Misfeldt, & Winsløw, 2009, p. 85). Grønbæk et al. are referring to written as well as oral exams in an introduction course on real analysis at the University of Copenhagen. The study of Grønbæk, Misfeldt and Winsløw investigated the possibilities of altering assessment and exam situations in order to “minimize regrettable effects of assessment, while retaining a visible incitement for students to meet necessary work requirements as well as a credible declaration of the results of this work” (Grønbæk et al., 2009, p. 85). They further argue that if they want to improve quality and scope of students’ work, they need to alter the exam, the frames of the teaching and the role of the teacher (Grønbæk et al., 2009, p. 91). The course was changed by experimenting with a new design of the oral exam, based on “thematic projects”. A thematic project is a written synopsis covering mathematical tasks, which are relatively open. Some of the tasks were more complex and theoretical than standard exercises (Grønbæk et al., 2009, p. 93). This format of oral exams has been introduced in upper secondary mathematics in Denmark, and below we will explain how this can function as assessment format in relation to SRP or SRA based teaching.

2.2. Questioning the world, SRP and test design

When Yves Chevallard (2015) proposed the teaching paradigm of questioning the world, he emphasised the role of questions to make the teaching of mathematics truly educational and adapted to present-day society. The aim of the paradigm of questioning the world should be for students – and citizens – to become *herbartian*, which means having a “receptive attitude towards yet unanswered questions and unsolved problems, which is normally the scientist’s attitude in his field of research” (2015, p. 178). In this sense the paradigm relates to problem posing, with respect to which Silver (2013) argues that:

Not only do we need more attention to issues of reliability and validity in the measures we use to assess the impact of mathematics instruction infused with problem-posing activities, but we also need to explore ways in which problem-posing tasks might be used as assessments of desired mathematics learning outcomes. (Silver, 2013, p. 161).

Both issues remain largely open in the mathematics education research literature; to address them from a task design perspective might promote a change of teaching paradigm putting questions and problem posing in front. Furthermore, Suurtamm and colleagues (2016) argue that the backwash effect of high stake exams is not uniquely negative: “if countries have nationally organised exit examinations these may drive (or hinder) reforms [of mathematics teaching]” (Suurtamm et al., 2016, p 20). Further, Malcolm Swan and Hugh Burkhardt (2012) suggest that

“teaching to the test” can have positive outcomes as long as the test items are designed to align with and nurture the intended reform of teaching. In terms of the scale of levels of codetermination, if we want to change the dominant teaching paradigm – hence make changes at the level of pedagogy – we need to take into account the higher levels of codeterminations (school, society and civilisation), including the preferences situated at these levels for evaluating the outcomes of mathematics teaching through high stake exams. In doing so we need to explicate how these high stake exams should be designed in order to promote the change of teaching paradigm. And, of course, be clear about what is it that high stake exams should promote.

3. SRP and current exams

In order to discuss the paradigm of questioning the world and SRP in relation to examination, a more detailed presentation seems appropriate. The paradigm of questioning the world is proposed through teaching designed as study and research paths (SRP). A SRP is initiated when a group X (of students) starts to study a generating question Q_0 , assisted by another group Y (one or more teachers) (Chevallard 2006, 2015). A generating question is formulated in a form so that students understand the question, but are unable to answer the question without engaging in a study and research process. Study and research processes are entangled and not necessarily possible to separate. The study process is characterised as students’ study of *works* created by others (textbooks, papers, online videos, etc.) – including other students. The knowledge studied is deconstructed and reconstructed as an answer to the generating question or one of the derived questions in what is characterised as the research process. The de- and reconstruction of knowledge is carried out on the basis of the students’ praxeological equipment – their previously acquired knowledge – which might be combined and related to create new knowledge. The process often requires that students pose derived, perhaps more manageable, questions Q_i (Barquero & Bosch, 2015). When studying works in order to answer Q_0 , students supposedly formulate questions addressing the nature of minor rules or notions, which need further study in order to develop a coherent answer to Q_0 . As such the question Q_0 generates numerous paths and side-tracks in the process of developing an answer to Q_0 (Bosch & Winsløw, 2016). Hence, the learning process of a SRP is characterised as a dialectics between research and study (Winsløw Matheron et Mercier, 2013), which the design of a strongly generative Q_0 supports.

It is evident that this way of teaching might lead to relatively different SRPs, meaning students develop different praxeologies and therefore different mathematical knowledge from working with a SRP. Moreover this calls for ideas regarding orchestration and guidance of the study and research process. The guidance have been approached by didactic researchers in ATD with different means (see Barquero et al, 2013; Serrano et al, 2010; Rasmussen, 2015; Florensa et al. 2015; Jessen 2014, 2015). When individual or group findings are shared, to some extent it can be expected that students develop similar but not identical answers. This is a challenge since national exit examinations are often based on individual exercises testing certain techniques which are the main monuments of curriculum. In this situation, common classroom activities and imitative reasoning might prepare the students equally well or even better for the examination exercises. This leads to the question if there exist exam formats where SRP prepare

students better, which are as reliable as existing formats, and which test curriculum in more satisfactory ways. Below we discuss these parameters in relation to the studies conducted by Jessen (2014, 2015). But before we can discuss these cases we give a short account of the context of these.

3.1. The context of the two studies

In Denmark, upper secondary mathematics education is divided into three levels. Level C takes one year and is intended for students with a main academic interest in humanities or social sciences. Level B takes two years and most students attend this. Level A lasts all three years of upper secondary school and is required for higher education in natural sciences, engineering, economy etc. Students' mathematical knowledge (specified in terms of competencies) is tested in three different types of assessment depending on the level of mathematics the students are studying.

The most common type of assessment is written exams which students sit for 4 or 5 hours in a gym hall, solving a number of exam exercises (Danish Ministry of Education, 2013). The first hour is dedicated to exercises to be solved with just pen and paper. During the remaining time there is a full allowance of calculators and computers, including Computer Algebra Systems (e.g. Geogebra, TI Nspire, wordmath, Maple etc.). Communication with others during examination period is not permitted, but students are allowed to re-visit webpages which have been used prior the exam by the students. This causes a variety of technical challenges including student fraud, which are pointed out in a Danish report on upper secondary mathematics teaching (Jessen, Holm, & Winsløw, 2015), but which will not be elaborated further here.

The second form is oral exams. Students draw a question, which they are familiar with and have prepared at home. They are given 24 or 30 minutes for preparation and examination. Questions often require students to present a notion, rule or theorem and how it can be applied in contexts outside mathematics. For levels B and A, students are expected to provide some proof or justification of the rules or theorems they present. A significant amount of the questions should be based on "thematic projects" – the format designed for the calculus course mentioned earlier and presented in Grønbæk et al. (2009). The ministry of Education was inspired by Grønbæk and colleagues and required the format used in secondary education. However the format was alien to most secondary teachers and represents a challenge for most teachers to design.

The last examination type is bidisciplinary projects, where students combine two disciplines they have studied at A- or B-level. The teachers of the two disciplines formulate a problem or set of questions, which the student must answer employing knowledge from the two disciplines. The answer must take the form of a 15-20 pages report, which is evaluated and given a grade by an external evaluator. This exam counts twice in the total score of the students' performance and is therefore very important to their final grade from upper secondary education. The individual student decides whether mathematics should be one of the disciplines in their bidisciplinary project.

All exams in Denmark are on draw, which means that each spring the Ministry of education randomly determines which students are attending what exams. Hence, students can graduate without attending any mathematics examination – and others are tested both written and orally.

Below are the rates of students from each level attending the oral and written examinations in the academic year 2013/2014:

Level of mathematics (A high)	Percentage of students attending oral exam	Percentage attending written exam
A	21 %	101 %
B	30 %	74 %
C	20 %	-

Table 1. The table shows how many percent of students at each level is attending oral and written exam respectively at each level in Danish upper secondary education (see (Jessen et al, 2015, p. 37).

The reason why more than 100% of the students attend the written exam at level A is that students can attend this exam even if they did not attend upper secondary education, but attended upgrade courses for higher education. It is evident that the probability of attending the written exam is much higher than the oral exam. Jessen et al. (2015) reports how this affects the activities in the classroom, according to both teachers and students: a large part of the teaching aims at students rehearsing their ability to answer the written exam exercises.

4. SRP as design for bidisciplinary written projects

The high stakes bidisciplinary written exam in Denmark runs around Christmas in the last year of upper secondary. Students choose two disciplines from their study line. Upper secondary education is organised around study lines with a number of common disciplines and three disciplines defining the study line, e.g. Biology, Mathematics and Social science. Students choosing this particular study line often choose higher education in the field of health and medical sciences.

The bidisciplinary exam results in a written report, which the students produce during two weeks without teaching. Together with the teachers of the two disciplines, the student finds an area of interest and the teachers formulate a problem or some questions the student must answer in the 15-20 pages report. The problem should cover both disciplines and the student is supposed to write a report combining the disciplines in a meaningful way.

The first case study we refer to was conducted in a study line as the one described above. During the second year the students wrote a minor bidisciplinary report (10-15 pages) on a problem formulated by the biology teacher and the mathematics teacher of the class. The students had two days without teaching and were expected to carry out most of the work after school hours. This second year report served as rehearsal for the high stake exam in the third year.

In 2012, a rehearsal problem was designed in terms of a generating question covering mathematics and biology (Jessen, 2014). The topic was how to administer painkillers in order to relieve a patient from his pain. This requires developing and activating praxeologies from the fields of differential equations and nerve physiology. Students formed minor study groups posing questions, sharing different media and discussing each others' hypotheses for answers to derived questions (Jessen, 2014, p. 218). The teachers followed the process through email correspondences with the students. Students were asked to pose questions on email so the teachers could keep track of the process. This gave insight into the sharing and the group dynamics of the class during the study and research process. Further it forced the students to be

precise in formulating the challenges they encountered and the formulations of the ideas they were working on. SRPs proved to be a suitable model for the activities and the a priori analysis secured the connections between disciplines. Moreover students engaged themselves in an inquiry of the fields involved, leading to rich answers regarding the generating questions. However some students did not see the need of mathematics in the project. A reason could be insufficient details in the initial questions handed out to the students. It was concluded that SRPs and the design principles were recommendable for this kind of exam.

As implication for the design of SRP related examinations, one may thus recommend more open projects, which require students to pose derived questions and engage in study and research processes, as input to the second issue formulated by Silver (2013), regarding test designs including problem posing and reflecting the classroom activity – under the assumption that students were taught mathematics through SRPs beforehand. However, it is hard to control if written reports were formulated by others than the stated author – practices of organised fraud were recently reported on in a documentary broadcasted by Danish national television. As a result, it has been proposed that the study line project should be defended at an oral exam, as with university dissertations. An alternative would be to require that students did the work in a gym hall, which would limit the study process, and in view of time constraints imposed by that, students might not be asked equally broad generating questions. For sure this type of examination requires further elaboration in design and orchestration before being fully applicable for large-scale mathematics examinations. Also, Swan and Burkhardt (2012) argue that mathematical processes might not require time-consuming project examinations. They refer to examples of test items being “rich tasks in the 10-20 minutes range [...] that is fair to students” (Swan & Burkhardt, 2013, p. 31). We will discuss a concrete alternative below.

5. SRA based teaching and oral exam

Jessen (2015) used study and research activities (SRA) at level C mathematics (cf. above) to teach the basics of exponential functions. The study shows potentials regarding students’ problem posing and development of answers to these. But also students’ reading of mathematical texts and their engagement in the study of different media was furthered. Moreover the study shows potentials in supporting the students’ de- and reconstruction of mathematical knowledge in autonomous working processes. The teaching was based on a sequence of generating questions like those of a SRP, however there was an specific target mathematical organisation for the students to develop, corresponding to the “monuments” of curriculum (for further details see Jessen, 2015, to appear). The design had similarities with the sequence of SRAs presented by Barquero, Serrano, & Ruiz-Munzón (2016). The first generating question leading to the first SRA was the following:

Q₁: Grandparents start a savings account for their new-born grandchild by putting 5,000 dkr into an account of annual rate of interest of 2.5%. Bank regulations say that the balance may not exceed 50,000 dkr. Will this be a problem?

The groups chose different solution strategies ranging from repeated multiplication with the factor 1.025, to graphic representations and to formulate the question as an equation, which the students needed to study new techniques in order to solve. The groups were only given 5-7 minutes for each question before presenting preliminary results. This assured the classroom

discussion on strategies, which encouraged the students to pose derived questions, discuss mathematics and justify their ideas.

The orchestration in the classroom was relatively strict. The class was divided into groups of three students with similar praxeological equipment, to the extent this could be estimated based on their previous performances. The groups were required to present preliminary answers to Q_0 and their own derived questions both orally and in writing at their groups' field of the whiteboard. They were not allowed to erase anything. Within this relatively tight framework, the students did share their work and questioned each other's strategies, and in fact they developed more coherent answers to the generating questions than those expected and found in curriculum and standard textbook materials (Jessen, to appear). The sequence of SRAs prepared the students for producing a thematic project, on which the oral exam was based. As we shall now see, this created a clear alignment between the in-class working format and the format of the final exam.

5.1. Thematic project exam and its link to SRAs

The thematic project on exponential functions, which the students answered as homework, were based on the following questions:

Q_a : Give a short presentation of the notion of exponential function. Apply the notion of rate of growth in your presentation.

Q_b : Given two points on a graph of exponential growth, which can be described by the expression $y = b \cdot a^x$, show how you can find the rate of growth a and the graphs intersection with the y -axis, b .

Q_c : Show how to model exponential growth based on a word problem and how it can be used for forecasts.

Q_d : What information is contained in the doubling time of an exponential function and how to calculate it ?

Q_e : What is exponential regression ? Give an example where you have used it and discuss the choices you made.

The students were explicitly encouraged to draw on their answers for the generating questions from the sequence of SRAs. The students wrote the thematic projects in the same groups as those they worked in during the SRAs. They knew the thematic project would be the basis of one or two questions at the oral exam. Therefore, the students genuinely engaged in studying the answers of the other groups during class sessions. Everybody wanted to perform at their best at a potential examination, and this generated an external motivation for the students to be able to justify the rules they might have to present at the exam. This can be a reason for the observed autonomous questionings by students during classes as reported in (Jessen, to appear). The draft projects were handed in to the teacher. The teacher gave comments on the projects, as in what parts were well written and well justified and what parts could need further elaboration. Based on the comments, the groups revised their projects before the oral exam.

This is similar to the management of the thematic projects described by Grønbaek, Misfeldt and Winsløw (2009). In the university course, students handed in six reports (4-5 pages) covering different topics. At the oral exam, students drew one report to present, instead of a theorem from the textbook as in the more usual practice. The reports were worked out as group

work. For each project, a date was fixed for handing in to the tutors, in order to get feedback on the reports. Tutors pointed out what was good and which parts needed elaboration. Lectures were partly dedicated to the topics of the projects, and the lecturer was also answering questions about them during lectures (Grønbaek et al., 2009).

Even if only 20% of the level C mathematics students attended the oral exam in 2014, according to table 1, it turned out to be 70 % of the class where the sequence of SRAs had been experimented. The entire curriculum on functions was taught on the basis of SRA sequences, and finalised in thematic projects. At the exam the students performed slightly above average, and better than previous classes in the same study line. The good grades were mainly given to students who drew questions relating to the notion of functions. This indicates that students were better prepared for the oral examination through the teaching based on open questions. At first, the external examiner expressed a great deal of concern regarding the SRA teaching, which he found “crazy”. After the last student had left the examination room he admitted, that it might not be so crazy after all. The reason for mentioning this episode is that this kind of resistance represents an obstacle for changing the teaching paradigm. Even if enthusiastic teachers wish to change the way they teach, they might meet resistance from other parts of the educational system, including colleagues and external examiners.

6. Discussion

The paradigm of questioning the world introduced through teaching designs based on SRP and SRA certainly shows potentials for teaching students not only the use of techniques, but also elements of the rationale of the techniques. Worth noticing is that the need for justification comes from the students, when they are trying to develop coherent answers rather than justification as another monument to visit (Jessen, to appear). Despite the good potentials and already documented results (e.g. in higher education), it is still a challenge to persuade teachers (who rarely have the motivations of a didactical researcher) to continue teaching based on SRPs (Barquero & Bosch, 2015).

The literature on large scale assessment and exit examinations indicates massively the crucial role played by assessment in the teaching paradigm, which is why this paper investigates the possibilities of designing stronger links between SRP or SRA based teaching and examinations. One can argue that the teaching paradigm will not change by just producing good examples of its realisation. An infrastructure of tools and regulations may have to complement the good examples.

The study of Jessen (2014) indicates that SRP is a strong tool regarding design and evaluation of the students' performances in project based examinations. The high stakes bidisciplinary written reports might be a rare case of examination, outside Denmark. However the study supports the suggestion of using larger projects for assessment in mathematics education. Similar ideas have been proposed by other researchers (e.g. see Niss, 1993; Frejd, 2013; Suurtamm et al., 2012).

Others, like Swan and Burkhardt “contest the assumption that tasks assessing mathematical processes need to be of project length” (2012, p. 31). They argue that high stakes exam exercises testing processes can be designed appropriately for written exams lasting e.g. four hours. Bearing in mind that project examinations never really have gained ground in

mathematics education, the idea of assessing the result of SRP based teaching in yet another SRP might not convince actors and decision makers at the higher levels of codetermination on how to design exit examinations.

With respect to oral exams, the thematic projects are promising formats which deserve further experimentation and design. It could be an interesting endeavor to cover all teaching for the one year mathematics through sequences of SRAs, with all exam exercises then relying on thematic projects. The study by Jessen (2015, to appear) exemplifies the idea that the change of examination drive the change of teaching, as proposed by Suurtamm et al. (2016), in the sense that writing the best possible thematic project delivered yet another incentive for students to engage in the SRAs.

This directly relates to the research question of this paper and what is needed from school systems and exam regulations, in order for the paradigm of questioning the world to be a viable alternative to the paradigm of visiting monuments. One answer could be to let examination rely only on oral examinations based on thematic projects. In the real analysis course redesigned by Grønbaek and colleagues (2009) the students' performance was measured by the oral exam as well as by a more traditional written exam and the two performances were given one grade. This could be another solution, to retain the attention to basic written skills. Here we must point out that a strong constraint on national examinations is the cost of the exam; it would be very costly to require all students to attend oral exams, and it is far from the current situation (see table 1).

This leads to the idea of designing test items for written exams, which are aligned with SRP and SRA based teaching. For teaching based on sequences of SRAs, the students are supposed to develop certain praxeological organisations which could serve as a basis for design of test items. Following the design presented by Jessen (2015), students could be tested reasonably by standard exam exercises regarding exponential functions. A classic example is exercises of the form: given a function describing the relation between the age and weight of a fish, how old is a fish of a given weight? How old is the fish when it has doubled its weight? A fish of a certain age had this weight, how does that fit the model? What are the limitations of the model?

The paradigm of questioning the world might prepare students to answer this exercise, but the exercise can hardly drive a change of teaching paradigm into questioning the world. In a teaching paradigm aiming at students developing *raison d'être* of the techniques they employ, justification should be a core element of the test items. For this purpose development of exercises as the following is needed:

Given the equality: $(3y - 6x) : 5 = 3$, argue if it represents a linear relation between the variables y and x .

This type of task requires that the students apply basic algebra techniques to manipulate the expression. Further, the students must know what is meant by a linear relation between the variables. Hence, the exercise tests a larger praxeological organisation, rather than isolated techniques. There is a need for developing a large variety of these types of tasks aiming at students' mathematical processes and justifications, as argued by Swan and Burkhardt (2012). Moreover, the praxeological analysis of ATD could become a strong tool for assessment design, in the sense that the explicit praxeological references models could guide what connections between point praxeologies should appear in the test, and what is actually tested in different type of tasks. Praxeological analysis of tasks has been developed in detail by Chaachoua (2010),

who conducted detailed analysis of type of tasks and techniques in order to develop an applet in terms of a computer algebra system aiming at students' solution strategies. This approach to praxeological analysis might also be useful for the design of more advanced test items.

In relation to SRP based teaching, the a priori analysis of possible paths and side-tracks including the final point praxeologies of students could also serve to develop more advanced items than the types of task usually covered in the written exam exercises.

Redesign of exam exercises might not be an easy path to changing the teaching paradigm. The studies of Romberg (1987), Schoenfeld (1988), Webb (1992), Cheng and Curtis (2004) and Suurtamm et al., (2016) all indicate that radical changes of the exam are not easy, due to the influence of policy makers, societal stake holders and parents, who all hold strong views regarding the nature of mathematics exams. Representatives of the higher levels of codetermination need to be taken into account when promoting a changed teaching paradigm.

Additionally, the teachers should be prepared for changes through professional development. Barquero, Bosch and Romo (2015) and Rasmussen (2015) present two different approaches to introducing and assisting teachers in designing SRP based teaching. Barquero et al. (2015, p. 813) have developed online in-service courses, and documented some of the barriers for teachers to actually change the way they teach. Rasmussen (2015) developed part of a pre-service teachers' course as an SRP, and encountered resistance from some students, who challenged the changed didactic contract of SRP based teaching, compared to more common classroom activities. Prevailing didactic contracts instill in students, including future teachers, some very solid beliefs about what mathematics teaching should be and how its outcomes must be assessed.

Further, if the paradigm of questioning the world should be anchored in the school systems, teachers need support at the lower levels of codetermination: discipline, domain, sector, theme and subject. The teachers need frames for professional development and tools on how to design and manage everyday SRP based teaching. This calls for paradidactic infrastructures. Here the Japanese tradition of lesson study, shared preparation and reflection upon shared teaching experiences, could serve as inspiration (for further details in paradidactic infrastructures see (Winsløw, 2012)).

7. Concluding remarks

It seems evident that to change the teaching paradigm of school mathematics requires changes at many levels of the school system and at the higher levels of codetermination. However, leaving aside the more political dimension of the question, there are intriguing needs for didactic researchers to pursue a number of design problems. This paper mainly offers indications of ideas on how to implement SRP and SRA based teaching in upper secondary schools.

There is, however, no doubt that ATD researchers should address more intensely the backwash effects of exams on the classroom activities, by designing and experimenting modified forms of high stakes examinations. The aim of this work should be to design examination formats which can further and adequately assess teaching according to the paradigm to questioning the world. This endeavour includes solving open questions on how to assess problem posing activities, to capture and value students' mathematical thinking, their creativity and study processes.

The format for oral exams offered by the notion of thematic projects seem promising and can drive the study and research process of teaching based sequences of SRAs. Much more open is the question regarding alternatives to or variations of standard written exams. There are potentials and shortcomings in written project as examination format. Maybe the answer to this question lies within new and more advanced uses of praxeological analysis.

References

- Barquero, B., Bosch, M., & Gascón, J. (2013). The ecological dimension in the teaching of mathematical modelling at university. *Recherches en didactique des mathématiques*, 33(3), 307-338.
- Barquero, B., & Bosch, M. (2015). Didactic engineering as a research methodology: From fundamental situations to study and research paths. In *Task Design In Mathematics Education* (pp. 249-272). Springer International Publishing.
- Barquero, B., Bosch, M., & Romo, A. (2015). A study and research path on mathematical modelling for teacher education. In *CERME 9-Ninth Congress of the European Society for Research in Mathematics Education* (pp. 809-815).
- Bergqvist, E. (2007). Types of reasoning in university exams in mathematics. *Journal of Mathematical Behavior*, 26, 348-370.
- Chaachoua, H. (2010). La praxéologie comme modèle didactique pour la problématique EIAH. Etude de cas: la modélisation des connaissances des élèves (Doctoral dissertation, Université de Grenoble).
- Cheng, L., & Curtis, A. (2004). Washback or backwash: A review of the impact of testing on teaching and learning. *Washback in language testing: Research contexts and methods*, 3-17.
- Chevallard, Y. (2002). Organiser l'étude. 3. ecologie & regulation. in J.-L. Dorier et al. (eds.), *Actes de la 11e Ecole d'Ete de didactique des mathématiques*, pp. 41-56. Grenoble: La Pensée Sauvage.
- Chevallard, Y. (2006). Steps towards a new epistemology in mathematics education. In *Proceedings of the 4th Conference of the European Society for Research in Mathematics Education (CERME 4)* (pp. 21-30).
- Chevallard, Y. (2015). Teaching mathematics in tomorrow's society: a case for an oncoming counter paradigm. In *The Proceedings of the 12th International Congress on Mathematical Education* (pp. 173-187). Springer International Publishing.
- García, F. & Higuera, L. R. (2005). Mathematical praxeologies of increasing complexity: variation systems modelling in secondary education. *WORKING GROUP 13 Applications and modelling*, 1645.
- Grønbaek, N., Misfeldt, M., & Winsløw, C. (2009). Assessment and contract-like relationships in undergraduate mathematics education. In *University science and mathematics education in transition* (pp. 85-105). Springer US.
- Florensa, I, Bosch, M., Gascón, J. (2016). SRP design in an elasticity course: the role of mathematic modelling. In *First conference of International Network for Didactic Research in University Mathematics*.
- Frejd, P. (2013). Modes of modeling assessment – a literature review. *Educational Studies in Mathematics*, 84, 413-438.

- Jessen, B. E. (2014). How can study and research paths contribute to the teaching of mathematics in an interdisciplinary settings?. In *Annales de didactique et des sciences cognitives*, 19, pp 199-224.
- Jessen, B. E. (2015). What role do study and research activities play in the relation between research and education? In: M. Achiam & C. Winsløw (Eds.), *Mathematics and Science: The relationships and disconnections between research and education. Papers from a doctoral course at the University of Copenhagen*. Copenhagen: Department of Science Education.
- Jessen, B. E., Holm, C. & Winsløw, C. (2015). Matematikudredningen. (Evaluation report). Copenhagen: Ministry of Education.
http://www.ind.ku.dk/publikationer/inds_skriftserie/2015-42/
- Jessen, B. E. (to appear). How to generate autonomous questioning in secondary mathematics teaching?
- Ministry of Education (2013). *Bekendtgørelse nummer 776 af 26/06/2013 – Bilag 37, Matematik C*. Copenhagen: Ministry of Education.
- Niss, M. (1993). Assessment in mathematics education and its effects: An introduction. In M. Niss (Ed.) *Investigations into assessment in mathematics education: An ICMI study (1-30)*. Dordrecht: Kluwer.
- Rasmussen, K. (2016). The direction and autonomy of interdisciplinary study and research paths in teacher education. *Journal of Research in Mathematics Education*, 5(2), 158-179.
- Romberg, T. A. (1987) Measures of mathematical achievement. In T. A. Romberg & D. Stewart (Eds.), *The monitoring of school mathematics: Background papers: Vol. 2. Implications from psychology outcomes of instructions*. Madison: Wisconsin Center for Education Research.
- Schoenfeld, A. H. (1988). When good teaching leads to bad results: the disasters of “well-taught” mathematics courses. *Educational Psychologist*, 23 (2), 145-166.
- Serrano, L., Bosch, M., & Gascón, J. (2010). Fitting models to data: The mathematising step in the modelling process. In *Proceedings of CERME* (Vol. 6, pp. 2-185).
- Singer, F. M., Ellerton, N., & Cai, J. (2013). Problem-posing research in mathematics education: new questions and directions. *Educational Studies in Mathematics*, 83(1), 1-7.
- Suurtamm, C., Thompson, D. R., Kim, R. Y., Moreno, L. D., Sayac, N., Schukajlow, S., ... & Vos, P. (2016). Assessment in mathematics education. In G. Kaiser (Ed.), *ICME-13 Topical Surveys Book Series* (pp. 1-38). Springer International Publishing.
- Swan, M., & Burkhardt, H. (2012). A designer speaks: Designing assessment of performance in mathematics. *Educational Designer: Journal of the International Society for Design and Development in Education*, 2(5), 1–41.
- Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. *Handbook of research on mathematics teaching and learning*, 661-683.
- Winsløw, C., Matheron, Y., & Mercier, A. (2013). Study and research courses as an epistemological model for didactics. *Educational Studies in Mathematics*, 83(2), 267-284.
- Winsløw, Carl (2015). Mathematical analysis in high school, a fundamental dilemma. In C. Bergsten & B. Sriraman (Eds.): *Refractions of Mathematics Education: Festschrift for Eva Jablonka*, pp. 197-213. Charlotte, NC, USA.
- Winsløw, C. (2012). A comparative perspective on teacher collaboration: The cases of lesson study in Japan and of multidisciplinary teaching in Denmark. In G. Gueudet, B. Pepin & L.

Trouche (Eds.), *From text to 'lived' resources: Mathematics curriculum materials and teacher development* (pp. 291-304). Dordrecht: Springer.