Mathematics and Mathematics Education
Two Sides of the Same Coin

Mathematics and Mathematics Education

Two Sides of the Same Coin

Some Results on Positive Currents Related to Polynomial Convexity and *Creative Reasoning in University Exams in Mathematics*

Ewa Bergqvist

Doctoral Thesis No. 36, 2006,
Department of Mathematics
and Mathematical statistics,
Umeå University

Department of Mathematics and Mathematical statistics
Umeå University
SE-901 87 Umeå, Sweden

ISSN 1102 -8300
ISBN 91-7264-208-4
Typeset by the author using LaTeX 2ε
Printed by Print & Media, Umeå universitet, Umeå, 2006: 2002532

To Pontus
and all our future children

Mathematics and Mathematics Education Two Sides of the Same Coin

Some Results on Positive Currents Related to Polynomial Convexity and *Creative Reasoning in University Exams in Mathematics*

DOCTORAL DISSERTATION
by
EWA BERGQVIST

Doctoral Thesis No. 36, Department of Mathematics and Mathematical statistics, Umeå University, 2006.

To be publicly discussed in the lecture hall MA 121, the MIT-building, Umeå University, on Friday, December 1, 2006, at 1.15 p.m. for the degree of Doctor of Philosophy.

Abstract

This dissertation consists of two different but connected parts. Part A is based on two articles in mathematics and Part B on two articles in mathematics education.

Part A mainly focus on properties of positive currents in connection to polynomial convexity. Earlier research has shown that a point z^0 lies in the polynomial hull of a compact set K if and only if there is a positive current with compact support such that $dd^cT = \mu - \delta_{z^0}$. Here μ is a probability measure on K and δ_{z^0} denotes the Dirac mass at z_0. The main result of Article I is that the current T does not have to be unique. The second paper, Article II, contains two examples of different constructions of this type of currents. The paper is concluded by the proof of a proposition that might be the first step towards generalising the method used in the first example.

Part B consider the types of reasoning that are required by students taking introductory calculus courses at Swedish universities. Two main concepts are used to describe the students' reasoning: imitative reasoning and creative reasoning. Imitative reasoning consists basically of remembering facts or recalling algorithms. Creative reasoning includes flexible thinking founded on the relevant mathematical properties of objects in the task. Earlier research results show that students often choose imitative reasoning to solve mathematical tasks, even when it is not a successful method. In this context the word *choose* does not necessarily mean that the students make a conscious and well considered selection between methods, but just as well that they have a subconscious preference for certain

types of procedures. The research also show examples of how students that work with algorithms seem to focus solely on remembering the steps, and researchers argue that this weakens the students' understanding of the underlying mathematics. Article III examine to what extent students at Swedish universities can solve exam tasks in introductory calculus courses using only imitative reasoning. The results show that about 70 % of the tasks were solvable by imitative reasoning and that the students were required to use creative reasoning in only one of 16 exams in order to pass. In Article IV, six of the teachers that constructed the analysed exams in Article III were interviewed. The purpose was to examine their views and opinions on the reasoning required in the exams. The analysis showed that the teachers are quite content with the present situation. The teachers expressed the opinion that tasks demanding creative reasoning are usually more difficult than tasks solvable with imitative reasoning. They therefore use the required reasoning as a tool to regulate the tasks' degree of difficulty, rather than as a task dimension of its own. The exams demand mostly imitative reasoning since the teachers believe that they otherwise would, under the current circumstances, be too difficult and lead to too low passing rates.

Sammanfattning på populärvetenskaplig svenska – så gott det går

Avhandlingen består av två ganska olika delar som ändå har en del gemensamt. Del A är baserad på två artiklar i matematik och del B är baserad på två matematikdidaktiska artiklar.

De matematiska artiklarna utgår från ett begrepp som heter polynomkonvexitet. Grundidén är att man skulle kunna se vissa ytor som en sorts ”tak” (tänk på taket till en carport). Alla punkter, eller positioner, ”under taket” (ungefär som de platser som skyddas från regn av carporttaket) ligger i något som kallas ”polynomkonvexa höljet.” Tidigare forskning har visat att för ett givet tak och en given punkt så finns det ett sätt att avgöra om punkten ligger ”under taket.” Det finns nämligen i så fall alltid en sorts matematisk funktion med vissa egenskaper. Finns det ingen sådan funktion så ligger inte punkten under taket och tvärt om; ligger punkten utanför taket så finns det heller ingen sådan funktion. Jag visar i min första artikel att det kan finnas flera olika sådana funktioner till en punkt som ligger under taket. I den andra artikeln visar jag några exempel på hur man kan konstruera sådana funktioner när man vet hur taket ser ut och var under taket punkten ligger.

De matematikdidaktiska artiklarna i avhandlingen handlar om vad som krävs av studenterna när de gör universitetstentor i matematik. Vissa uppgifter kan gå att lösa genom att studenterna lär sig någonting utantill ur läroboken och sen skriver ner det på tentan. Andra går kanske att lösa med hjälp en algoritm, ett ”recept,” som studenterna har övat på att använda. Båda dessa sätt att resonera kallas *imitativt resonemang*. Om uppgiften kräver att studenterna ”tänker själva” och skapar en (för dem) ny lösning, så kallas det *kreativt resonemang*. Forskning visar att elever i stor utsträckning väljer att jobba med imitativt resongemang, även när uppgifterna inte går att lösa på det sättet. Mycket pekar också på att de svårigheter med att lära sig matematik som elever ofta har är nära kopplat till detta arbetssätt. Det är därför viktigt att undersöka i vilken utsträckning de möter olika typer av resonemang i undervisningen. Den första artikeln består av en genomgång av tentauppgifter där det noggrant avgörs vilken typ av resonemang som de kräver av studenterna. Resultatet visar att studenterna kunde bli godkända på nästan alla tentorna med hjälp av imitativt resonemang. Den andra artikeln baserades på intervjuer med sex av de lärare som konstruerat tentorna. Syftet var att ta reda på varför tentorna såg ut som de gjorde och varför det räckte med imitativt resonemang för att klara dem. Det visade sig att lärarna kopplade uppgifternas svårighetsgrad till resonemangstypen. De ansåg att om uppgiften krävde kreativt resonemang så var den svår och att de uppgifter som gick att lösa med imitativt resonemang var lättare. Lärarna menade att under rådande omständigheter, t.ex. studenternas försämrade förkunskaper, så är det inte rimligt att kräva mer kreativt resonemang vid tentamenstillfället.

Preface

Seven years ago I left the department of mathematics at Umeå university just after completing my licentiate degree. I started studying maths in 1989 and spent the following ten years first as a student and then as a doctoral student (doing research and teaching). I felt quite at home but my self-confidence wasn't at its peak. I didn't believe that I had very much to contribute to research, and I could not help myself from constantly comparing my achievements with anyone more successful than myself. In my eyes that was everyone. In 1999 I finally decided to leave. I wasn't happy doing research in mathematics, although I loved both mathematics and teaching, and I knew that it was time to do something new. I just wasn't sure what.

By pure chance I got a job at a small company, Spreadskill, that provided testing platforms, tests, and test development to other organisations. I entered a world quite different from what I was used to. At Spreadskill I started to learn a thing or two about assessment and testing and my interest for measurements began to sprout. (I also learnt a lot about participating in and leading projects, handling customers, giving technical support, being a consultant, revising test items, and drinking coffee.) As it turned out, I wasn't stupid after all (Yey!), and I started to feel a lot more confident.

After a few years at Spreadskill I wanted to combine my new interest in assessment with my old love for mathematics. I got a temporary position in the project group constructing Swedish National tests in mathematics at the department of Educational Measurement (EM). I was responsible for the E course item bank test for upper secondary school for two years. During my time at EM I read more about item response theory, standard setting, and other aspects of assessment and also participated in the writing of a couple of reports. After a while I realised that I had a lot of questions concerning university exams in mathematics from a perspective of mathematics education and assessment. Research in mathematics education felt very tempting, but the idea of being a PhD student again scared the ba-jeebas out of me. Really. Fortunately, I told Johan Lithner that I was interested and then he took over. He applied for funding, talked to the right people, and suggested a preliminary research plan. Actions that made me realise that I could and really wanted to do this. He believed in me, and I slowly started to do the same. In january 2005 I started working in Johan's project and a couple of months later I was accepted as a PhD student at the faculty of teacher education. And I was back at the Mathematics department (now: the Department for mathematics and mathematical statistics). Doing research and doing math. Benefiting from my math studies, from what I learnt at Spreadskill and EM, and—most of all—having fun!

Coming full circle.

Acknowledgements

First of all, I would like to thank my supervisors in mathematics education, **Johan Lithner** and **Peter Nyström**. They have offered invaluable help during the planning, implementation, and documentation of my two studies in mathematics education (Article III and IV). I have learnt a lot about doing and presenting research from our discussions. Thank you! An extra thanks to Johan since I literary wouldn't be at this point if it weren't for him!

I would also like to thank **Urban Cegrell**, my supervisor in mathematics, for introducing me to polynomial convexity. Thanks also to **Magnus Carlehed**, **Klas Markström**, and particularly **Anders Fällström** for valuable discussions and for reading and commenting on the mathematical articles.

A special thanks to the **six teachers** that patiently let me interview them. Article IV would not have been the same without you. Thanks!

Agneta Halvares-Palmqvist, **Anders Rehnman**, and **Peter Nyström** have given valuable comments on the language. **Tomas Bergqvist** have read and commented the contents of the mathematics education part of the thesis. Thanks, all of you!

The research group in mathematics educations has provided a forum for discussions and seminars that have been both fun, challenging, and interesting. Thank you all!

My colleagues at the Department of mathematics and mathematical statistics have made me feel welcome back and once more a natural part of the staff. Your support has been really important for me. Thanks! Special thanks to **Berith Melander** for valuable help with the practical chores preceeding the printing of this thesis.

My former colleagues at the Department of educational measurement have also been encouraging and supportive. I really appreciate it. Thank you! A special thanks to **Anna**—I've really appreciated our lunch discussions that periodically have been my only contact with the rest of the world!

My former colleagues at Spreadskill. You taught me more than I think you know. I particularly remember learning about language and gender (**Maria**), about projects and herbs (**Marie**), and about drinking coffee (**Krister**). Thank you all!

I would also like to thank all my other friends—**Mattias, Cicci, Oliver, Anders, Minea, Angelica, Helena, Marie, Lisa, Jocke**, and many others—for being my friends, and especially for putting up me these last months. I'll be back!

My Rokugan subordinates, **Tsiro, Sadane, Garu**, and **Kojiro**. I trust you with my... or, rather... I kind of think... well, at least you try! Arigato!

Mom and **dad**, thank you for always making me feel loved. You're the best!

To the rest of my extended family—**my brothers, Berit & Sven-Olov, my brothers- and sisters-in-law, Kerstin, my nephews and nieces**—thank you all for being so easy to get along with and for always being supportive!

At last I want to thank my partner, soul-mate, and best friend **Pontus**, who by pure coincidence also happens to be the funniest and most attractive man alive. I truly could not have done this without you. I love you! (And the little weird alien and the black furry skurk that live with us.)

"Follow your bliss!"
Joseph Campbell

Contents

Chapter 1

Introduction

This thesis is divided into two separate, but still related, parts. Part A is based on two articles in mathematics within the research area of several complex variables, Articles I and II. The articles concern polynomial convexity and the connection to positive currents. Part B is based on two articles in mathematics education with focus on the required reasoning in calculus exams at Swedish universities, Articles III and IV. All four articles are placed at the end of the thesis.

Although there is no direct relation between the contents in Part A and B, they are strongly connected. One obvious connection is that: it is all about mathematics. Different aspects of mathematics, granted, but still mathematics. Another connection is simply that Part B would not exist without Part A. This is true in several ways. First, I would not have been able to carry out the analysis of exam tasks (as in Article III) if I had not had the competence in mathematics gained from, among other things, writing Part A. (The analysis included not only solving the exam tasks, but also e.g. assessing the possibility of other realistic solutions to each task.) Second, as a doctoral student in mathematics I did a lot of teaching and exam construction myself. My experiences as a university teacher in mathematics were crucial when planning the interview study presented in Article IV. My background also facilitated my communication with the university teachers during the interviews, since we e.g. had similar vocabulary and experiences. Also, on several occasions during the interviews I discussed different aspects of the teachers' own exam tasks, e.g. my classification of a task or different possible solutions of a task, discussions where a firm competence in mathematics was crucial. This connection between the two parts of the thesis is the reason I see mathematics and mathematics education as two sides of the same coin.

Part A

Chapter 2

Part A

Part A of the thesis consists of an introduction, a brief background (Section 2.1), and Articles I and II (placed after Part B). Both articles are based primarily on results by Duval and Sibony, presented in "Polynomial Convexity, Rational Convexity, and Currents" (Duval and Sibony, 1995), and the second article is also based on results by Bu and Schachermayer (1992). The Background contains no definitions, since it is merely intended to give a brief historical setting for the two articles. Necessary definitions can be found in Section 2 of the second article. This introduction, the background, and the two articles are almost identical to the contents of the author's Licentiate thesis defended on September 14, 1999.

2.1 Background

There are good reasons why polynomial convexity is an object of intense study. Every function analytic in a neighbourhood of a polynomially convex set is the uniform limit, on the set, of polynomials. So the notion of polynomial convexity is frequently used in connection to uniform approximation. In one complex variable, every simply connected set is polynomially convex, so the setting is fairly uncomplicated. In several complex variables there is no simple geometrical or topological interpretation of the concept.

Polynomial convexity is connected to maximal ideal spaces as well. Every finitely-generated function algebra can be realized as the uniform closure of the polynomials on a compact set. The maximal ideal space of this function algebra is exactly the polynomial hull of the set, see Gamelin (1969), Chapter III.

The local maximum modulus principle by Rossi also concerns polynomial convexity (Rossi, 1960; Oka, 1937). If $\hat{K}$ denotes the polynomial hull of K, Rossi's local maximum modulus principle tells us that

$$\text{if } Y \subset \hat{K} \setminus K, \text{ then } Y \subset \widehat{\partial Y}$$

where ∂Y is the topological boundary of Y in $\hat{K}$. Stolzenberg (1963b) asked if this principle is anything more than the classical maximum modulus principle for analytic functions on an analytic variety. "In particular, does the set $\hat{K} \setminus K$ consist of (or, at least, contain) positive dimensional analytic varieties?" If almost every point of the radial limit of an analytic disc lies in a compact set, the whole image of the disc lies in the polynomial hull of the set. Could it be that the "analytic structure" of $\hat{K} \setminus K$ guarantees that it always contains some analytic disc? This question has been answered by both Stolzenberg and Wermer, and the answer is no. The set $\hat{K} \setminus K$ need not contain an analytic (Stolzenberg, 1963a) or even a continuous (Wermer, 1982) disc. But $\hat{K} \setminus K$ has an analytic structure of some kind, and Duval and Sibony (1995) finally came up with a clarifying result. Inspired by the fact that integration over an analytic set of complex dimension p can be seen as a positive (p,p)-current, they proved the following: A point lies in the polynomial hull of a set if and only if there exists a specific positive $(1,1)$-current, associated to the given point in the polynomial hull.

2.2 Summaries of the articles in Part A

2.2.1 Article I

Title: Non-uniqueness of Positive Currents Describing Polynomial Convexity

This article contains a result concerning non-uniqueness of a current, as mentioned in the title. Duval and Sibony (1995) showed that a point lies in the polynomial hull of a compact set if and only if there exists a positive current directly related to the set and the point. The main result in the first paper is that such a current need not be unique. The paper also contains a discussion on positivity for currents, especially in $\mathbb{C}^2$, and some examples of currents of the type that Duval and Sibony proved existence for.

2.2.2 Article II

Title: Positive Currents related to Polynomial Convexity

In this article the work with positive currents is continued. The paper contains two different constructions of currents of the type found by Duval and Sibony. One construction merely gives, for a specific given set, a simple example constructed with a global method. The other construction is more general since there is no restrictions as to which set to consider. The method is local and leans heavily on results by Bu and Schachermayer (1992) concerning approximation of

Jensen measures using analytic discs. The paper is concluded by the proof of a proposition that might be the first step towards generalising the method used in the first, global, example.

Part B

Chapter 3

Part B

I believe that algorithms are an important part of mathematics and that the teaching of algorithms might be useful in the students' development. I also believe that a too narrow focus on algorithms and routine-task solving might cut the students off from other parts of the mathematical universe that also need attention, e.g. problem solving and deductive reasoning. This belief is to some extent supported by both national and international research results, e.g. Lithner (2000a,b, 2003, 2004); Palm et al. (2005); Leinwand (1994); Burns (1994); McNeal (1995); Kamii and Dominick (1997); Pesek and Kirshner (2000); Ebby (2005). Most of the international results mentioned do however concern school years K-9 and often young children in situations concerning arithmetic calculations. The situation at the university level, especially in Sweden, is not thoroughly studied, and there are still many important questions to be asked.

This thesis contains two articles within the research field of mathematics education, Article III and Article IV. The articles both focus on university exams in mathematics and the reasoning the students have to perform to solve the tasks in the exams. The two studies lie within the scope of a large project, "Meaningless or meaningful school mathematics: the ability to reason mathematically"[1] (abbreviated *Meaningful mathematics*), carried out by the research group in mathematics education at Umeå University. The project primarily focused on upper secondary school and university students. The purpose of the project was to identify, describe, and analyse the character of, and reason for, students' main difficulties in learning and using mathematics. This all-embracing purpose was naturally too extensive to be expected to be fully attained within the scope of the project, but describes quite well the setting for the Article III and IV in this thesis. The Meaningful mathematics project was based on empirical

[1]The project was mainly financed by the Swedish Research Council

findings implying that students' difficulties with solving problems[2] and developing mathematically might be connected to their inability to use, and perhaps unaccustomedness with, creative reasoning.

Part B consists of a theoretical background to the articles in mathematics education, a presentation of the context including related research, a short description of the Swedish system, a section on method, a presentation of the Swedish system, a summary of the two articles, a concluding discussion, and the Articles III and IV (placed at the end of the thesis).

3.1 Theoretical background

The aim of this section is to provide the reader with relevant definitions and a theoretical foundation more extensive than is possible within the scope of each article. Presented in this section are four different theoretical frameworks.

The framework focusing different types of mathematical reasoning (Lithner, 2006), together with the framework outlining general mathematics competences (NCTM, 2000), form the base of the analyses performed in the studies. Why these particular frameworks were chosen is discussed in Section 3.4.2. The third framework presented (Schoenfeld, 1985) is used to structure the arguments for the value of creative reasoning in problem solving as done e.g. in Section 3.2.4. The last framework (Vinner, 1997) is briefly presented here because it includes several interesting ideas and theories directly related to the issues concerning mathematical reasoning discussed in this thesis. It is referred to in Section 3.2 in connection to a couple of different phenomena.

3.1.1 The framework for analysing reasoning

There are several theoretical comprehensive frameworks for mathematical competences but not as many detailed frameworks specifically aimed at characterising mathematical reasoning. The concept "reasoning" (or "mathematical reasoning") is in many contexts used to denote some kind of high quality reasoning that is seldom, or vaguely, defined. It is reasonable that general frameworks, in order to be comprehensive, do not use too much detail in defining concepts. The consequence is, however, that they will not be very useful in detailed analyses of the type and quality of mathematical reasoning used by students or demanded by school tasks (textbook exercises, test tasks, and so on). A framework that particularly consider mathematical reasoning is developed by Lithner (2006). In this fairly detailed framework, *mathematical reasoning* is defined as any type of

[2]A *problem* is in this thesis defined as "a task for which the solution method is not known in advance", see p. 51 in NCTM (2000).

reasoning that concerns mathematical task solving, and the quality of a specific case of reasoning is expressed by different reasoning characteristics.

The reasoning framework (Lithner, 2006) was developed within the "Meaningful mathematics" project and is based on results from several empirical studies aiming at analysing traits of, and reasons for, students' learning difficulties, e.g. Lithner (2000a,b, 2003, 2004). These studies' main conclusion is that "the students are more focused on familiar procedures, than on (even elementary) creative mathematical reasoning and accuracy." A specific example is students' reasoning when solving textbook exercises. The found dominating strategy was that the students identified similarities between exercises and examples in the textbook, and then copied the steps of the solution. Within the various studies included in "Meaningful mathematics", this and several other different types of reasoning were identified. The need to characterise and compare these reasoning types with reasonable consistency led to the development of the reasoning framework.

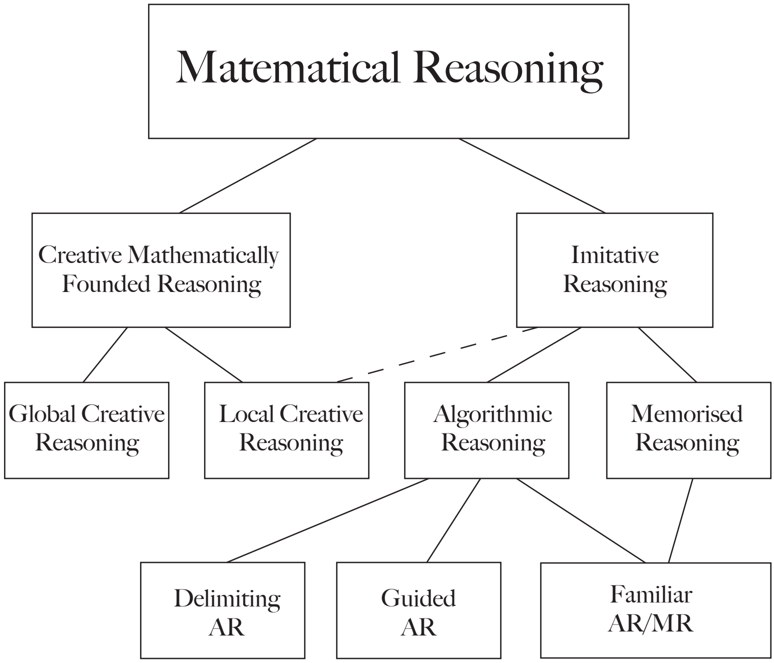
During the empirical studies, two basic types of mathematical reasoning were defined: *creative mathematically founded reasoning* (CR) and *imitative reasoning* (IR), where IR primarily consists of *memorised reasoning* (MR) and *algorithmic reasoning* (AR). There exist three main empirically established versions of imitative reasoning—familiar MR/AR, delimiting AR, and guided AR—which could give the impression that IR has a richer characterisation than CR. The reason that more versions of IR has been identified is, however, not that IR is richer but simply that IR is so much more common in the empirical data (Lithner, 2006). Figure 3.1 on page 16 gives an overview of the different types of reasoning in the framework. All the types of reasoning in the figure are presented in the following subsections, except for local and global CR that are presented in connection to the classification tool (Section 3.4.1).

Creative mathematically founded reasoning

The reasoning characteristics that separate creative mathematically founded reasoning from imitative reasoning are creativity, plausibility, and mathematical foundation. Creativity is the key characteristic that separates the two, and the concepts plausibility and mathematical foundation are necessary to separate mathematical creativity from general creativity.

Creativity. Models presenting common process steps for successful problem solvers have been presented by several researchers, e.g. Pólya (1945); Schoenfeld (1985); Sriraman (2004), but these models do not specifically define features that are necessary for a reasoning to be characterised as creative. Haylock (1997) claimed that there are at least two mayor definitions of the term creativity within the research literature. One definition considers thinking that leads to a product considered magnificent by a large group of people. A typical example is a grand work of art. The other definition is more modest and considers thinking that is

Figure 3.1: Overview of reasoning types in the framework



divergent and avoids fixation. This structure is aligned with what Silver (1997) suggests: that mathematics educators may benefit from considering creativity an inclination toward mathematical activity possible to promote generally among students, instead of a trait solely associated with geniality and superior thinking. The framework uses these ideas of Haylock and Silver to see creativity as primarily characterised by *flexibility*, and *novelty*. Creative reasoning is flexible since it avoids fixation with specific contents and methods, and it is novel because it is new to the reasoner. Creative reasoning is in this context not connected to exceptional abilities.

Plausibility. The students solving school tasks are not required to argue the correctness of their solutions as rigorously as professional mathematicians, engineers, or economists. Within the didactic contract (Brosseau, 1997) it is possible for the students to guess, make mistakes, and use mathematically questionable ideas and arguments. Inspired by Pólya (1954) and his discussion of strict reasoning vs. plausible reasoning, plausibility is therefore in the framework used to describe reasoning that is *supported by arguments that are not necessarily as* *strict as in proof.* The quality of the reasoning is connected to the context in which it is produced. A lower secondary school student arguing for an equality by producing several numerical examples confirming its validity might be seen as performing quite high quality reasoning, while the same reasoning produced by a university student would be considered of poor quality.

Mathematical foundation. The arguments used to show that a solution is plausible, can be more or less well founded. The framework defines task *components* to be objects (e.g. numbers, functions, and matrices), transformations (e.g. what is being done to an object), and concepts (e.g. a central mathematical idea built on a related set of objects, transformations, and their properties). A component is then seen to have a *mathematical property* if the property is accepted by the mathematical society as correct. The framework further distinguishes between *intrinsic* properties—that are central to the problematic situation—and *surface* properties—that have no or little relevance—of a particular context. An example: try to determine if an attempt to bisect an angle is successful. The visual appearance of the size of the two angles is a surface property of these two components while the formal congruency of the triangles is an intrinsic property. The empirical studies showed that one of the main reasons for the students' difficulties was their focus on surface properties (Lithner, 2003). Similar results were obtained by Schoenfeld (1985) who found that novices often use 'naive empiricism' to verify geometrical constructions: a construction is good if it 'looks good' (as in the example with the bisected angle mentioned above). In the reasoning framework, a solution have mathematical foundation if "the argumentation is founded on intrinsic mathematical properties of the components involved in the reasoning" (Lithner, 2006).

To be called ***creative mathematically founded reasoning***, the reasoning in a solution must fulfil the following conditions:

- *Novelty* The reasoning is a new (to the reasoner), or a re-created forgotten, sequence of solution reasoning.
- *Flexibility* The reasoning fluently admits different approaches and adaptations to the situation.
- *Plausibility* The reasoning has arguments supporting the strategy choice and/or strategy implementation, motivating why the conclusions are true or plausible.
- *Mathematical foundation* The reasoning has arguments that are founded on intrinsic mathematical properties of the components involved in the reasoning.

Creative mathematically founded reasoning is for simplicity usually only denoted *creative reasoning* and is abbreviated CR. Whether a task demand CR of the students or not is directly connected to what type of tasks the students have practised solving. Suppose that a group of students have studied *continuity* at a calculus course. They have seen examples of both continuous and discontinuous functions, they have studied the textbook theory definitions, and they have been to lectures listening to more informal descriptions of the concept. The students have also solved several exercises asking them to determine whether a function is continuous or not. They have not, however, encountered a task asking them to construct an example of a function with certain continuity properties. Under these circumstances an example of a CR task could be: "Give an example of a function f that is right continuous, but not left continuous, at $x = 1$."

Imitative reasoning

Imitative reasoning can be described as a type of reasoning built on copying task solutions, e.g. by looking at a textbook example or through remembering an algorithm or an answer. An *answer* is defined as "a sufficient description of the properties asked for in the task." A *solution* to a task is an answer together with arguments supporting the correctness of the answer. Both the answer and the solution formulations depend on the particular situation where the task is posed, e.g. in a textbook exercise, in an exam task, or in a real life situation.

Several different types of imitative reasoning are defined within the framework. The two main categories, memorised and algorithmic reasoning, will be presented here, together with the empirically established versions: familiar memorised reasoning, familiar algorithmic reasoning, delimiting algorithmic reasoning and guided algorithmic reasoning. Guided algorithmic reasoning is usually not possible to use during exams and is therefore only presented briefly. Familiar and delimiting algorithmic reasoning are however highly relevant versions in the classification study. These versions' foundation on the students' familiarity of an algorithm, or a set of algorithms, is the basis for the classification of tasks using the theoretical classification tool presented in Section 3.4.1.

Memorised reasoning

The reasoning in a task solution is denoted ***memorised reasoning*** if it fulfils the following conditions:

- The strategy choice is founded on recalling a complete answer by memory.
- The strategy implementation consists only of writing down (or saying) the answer.

Typical tasks solvable by memorised reasoning are tasks asking for definitions, proofs, or facts, e.g. "What is the name of the point of intersection of the x- and y-axis in a coordinate system?"

Even long and quite complicated proofs, like the proof of the Fundamental Theorem of Calculus (FTC), are possible to solve through memorised reasoning. In a study described in the reasoning framework, many students that managed to correctly prove the FTC during an exam were afterwards only able to explain a few of the following six equalities included in the proof.

$$F'(x) = \lim_{h \to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0} \frac{1}{h} \left(\int_a^{x+h} f(t)dt - \int_a^x f(t)dt \right) =$$

$$= \lim_{h \to 0} \frac{1}{h} \int_x^{x+h} f(t)dt = \lim_{h \to 0} \frac{1}{h} hf(c) = \lim_{c \to 0} f(c) = f(x)$$

The students' inability to explain these equalities, most of which are elementary in comparison to the argumentation within the proof, imply that they memorised the proof and did not understand it. A task asking the students to prove a theorem is however only possible for the students to solve using memorised reasoning if they are informed in advance that the proof might be asked for during the exam. It is common in Sweden that the teachers hand out a list of theorems, proofs, and definitions that might appear on the exam. Memorised reasoning is abbreviated MR.

Algorithmic reasoning

An algorithm, according to Lithner (2006), is a "set of rules that if followed will solve a particular task type." An example is the standard formula for solving quadratic equation. The definition also includes clearly defined sequential procedures that are not purely calculational, e.g. using a graphing calculator to approximate the solution to an equation through zooming in on the intersection. Even though these procedures are to some extent memorised there are several differences between algorithmic and memorised reasoning. The most apparent is that a student performing memorised reasoning has completely memorised the solution, while a student using algorithmic reasoning memorises the difficult steps of a procedure and then perform the easy ones. That algorithmic reasoning is dependent on the order of the steps in the solution also separates algorithmic from memorised reasoning where the different parts could mistakenly be written down in the wrong order.

The reasoning in a task solution is denoted ***algorithmic reasoning*** if it fulfils the following conditions:

• The strategy choice is founded on recalling by memory a set of rules that will guarantee that a correct solution can be reached.

• The strategy implementation consists of carrying out trivial (to the reasoner) calculations or actions by following the set of rules.

The word *trivial* is basically used to denote lower level mathematics, i.e. standard mathematical content from the previous stages of the students' studies.

Algorithmic reasoning is a stable solution method in cases of routine task solving when the reasoner has encountered and used the algorithm several times and is completely sure of what to do. Still, the studies mentioned earlier indicate that students also try, unsuccessfully, to use algorithmic reasoning in problem solving situations. Using algorithmic reasoning is not a sign of lack of understanding, since algorithms are frequently used by professional mathematicians. The use of algorithms save time for the reasoner and minimises the risks for miscalculations, since the strategy implementation only consists of carrying out trivial calculations. Algorithmic reasoning is however *possible* to perform without any understanding of the intrinsic mathematics. Algorithmic reasoning is abbreviated AR.

Familiar memorised or algorithmic reasoning

Imitative reasoning is often based on familiarity of tasks, i.e. that the reasoner recognise a particular task type. If the task is familiar, the reasoner either remembers the complete answer or recalls the correct algorithm. These two different versions are called *familiar memorised reasoning* and *familiar algorithmic reasoning* respectively.

An uncomplicated version, described by Hegarty et al. (1995), is when elementary school pupils use what the authors call a "key word strategy." An example is when textbook tasks contain either the word "more" or the word "less" and the student is supposed to use either the addition algorithm or the subtraction algorithm: "Tom has 4 marbles and Lisa has 3 *more* than Tom. How many marbles does Lisa have?" If a student's choice between the two algorithms is solely based on what word that appear in a particular task, the student is using the key word strategy. The strategy is very useful for solving many context tasks containing the word "more" without the pupil having to consider the mathematical meaning of the task formulation. Not all "more"-tasks would of course be possible to solve with this algorithm. If the task was formulated "Mary bought 17 sodas. If Mary bought 5 sodas *more* than Mike, how many did Mike buy?" the method would result in an incorrect answer.

The reasoning in a task solution attempt will be called ***familiar MR/AR*** if:

- The strategy choice is founded on identifying the task as being of a familiar type, in the sense that it belongs to a familiar set of tasks that all can be solved by recalling a complete answer or by the same known algorithm.
- The strategy implementation consists of recalling and writing down the answer (MR) or implementing the algorithm (AR).

Familiar memorised and algorithmic reasoning are abbreviated FMR and FAR respectively.

Delimiting algorithmic reasoning

If the task at hand is not completely familiar to the student, and the student have access to too many algorithms to try all of them, one possibility is to delimit the set of possible algorithms. This type of imitative reasoning is called *delimiting algorithmic reasoning.*

The reasoning in a task solution attempt will be called ***delimiting AR*** if:
• The strategy choice is founded on choosing an algorithm from a set which is delimited by the reasoner through the included algorithms' surface property relations to the task.
• The strategy implementation consists of implementing the algorithm. No verificative argumentation is required. If the algorithm does not lead to a (to the reasoner) reasonable conclusion, the implementation is not evaluated but simply terminated and a new algorithm is chosen.

One example, mentioned in Lithner (2006) and described in detail on pages 10–15 in Bergqvist et al. (2003), consists of a description, an interpretation, and an analysis of the reasoning of an upper secondary school student. Sally is introduced to the task "Find the largest and smallest values of the function $y = 7 + 3x - x^2$ on the interval $[-1, 5]$" and is asked to solve the task while describing what she is doing. Sally starts solving the task by differentiating the function and finding the x-value, $x = 1.5$, for which the derivative is zero. She evaluates $y(1.5) = 9.25$ and then stops. Sally says that she should have gotten two values, and that she does not know what she has done wrong. After some hesitation on how to continue, she leaves the task for 20 minutes (while she works with two other tasks) and then returns. On her second attempt Sally decides to use the graphing calculator. She draws the graph and tries, unsuccessfully, to use the built in minimum-function. When this does not work, Sally tries another method: the calculator's table-function. Since the calculator is set for integer steps, the function's largest value in the table is 9 and the smallest value is -3. The contradiction between this result and the result 9.25 that Sally found during her first attempt makes her believe that the table cannot be used. Sally now tries to find another acceptable algorithm to use for her solution, and she decides to solve the equation $7 + 3x - x^2 = 0$. She successfully uses the standard algorithm for solving quadratic equations and ends up with two approximate values of x: 4.54 and -1.54. Sally believes that these two values are the answers to the task, but cannot answer a direct question on why this method would result in the largest and the smallest value.

Sally makes several strategy choices: differentiating the function and solving when zero, using the calculator's minimum-function, using the calculator's table-function, and solving the quadratic equation. Every choice is based on the different algorithms having surface connections with the task at hand. The algorithms are not chosen completely at random, but randomly from a delimited set of possible algorithms. She does not evaluate or analyse the outcome, and as soon as she finds that an algorithm does not work, she chooses a new one.

Delimiting algorithmic reasoning is abbreviated DAR.

Guided algorithmic reasoning

Different types of *guided AR* (GAR) are also presented in the framework. A student that uses *text-guided AR* selects an algorithm through surface similarities between the given task and an occurrence in an available text source, e.g. an example in the textbook. The student then simply copies the procedure in the identified occurrence. During *person-guided AR* the strategy choices are controlled and performed by someone other than the reasoner, e.g. the teacher or a fellow student. The algorithm is implemented by following the other person's guidance and by performing the remaining routine calculations.

3.1.2 The NCTM Principles and Standards for School Mathematics

A framework presented by the National Council of Teachers of Mathematics (NCTM) was chosen as the general framework used in an analysis in Article IV (NCTM, 2000). The framework is abbreviated *the Standards* and consists of six principles (equity, curriculum, teaching, learning, assessment, and technology), five content standards (number & operation, algebra, geometry, measurement, and data analysis & probability), and five process standards (problem solving, reasoning & proof, communication, connections, and representation). The Standards is primarily developed for school mathematics and not for university studies, and neither the principles nor the content standards are used in the analysis performed in the study.

One of the analyses of the interviews in Article IV is based on the highly relevant, but partly rather vague, *process standards* presented in the Standards. Article IV focus on university teachers' views on the required reasoning in calculus exams. The analysis of the teachers' descriptions of how they construct exam tasks and what standards they aim at testing, is based on the process standards. The purpose is to describe to what extent they focus on different processes during exam construction.

The process standards are:

Problem solving—engaging in a task for which the solution method is not known in advance. The students should be able to *build new mathematical knowledge through problem solving*, *solve problems that arise in mathematics and in* *other contexts*, *apply and adapt a variety of appropriate strategies to solve problems*, and *monitor and reflect on the process of mathematical problem solving.*

Reasoning & proof—developing ideas, exploring phenomena, justifying results, and using mathematical conjectures in all content areas. The students should be able to *recognize reasoning and proof as fundamental aspects of mathematics*, *make and investigate mathematical conjectures*, *develop and evaluate mathematical arguments and proofs*, and *select and use various types of reasoning and methods of proof.*

Communication—sharing ideas and clarifying understanding. The students should be able to *organize and consolidate their mathematical thinking through communication*, *communicate their mathematical thinking coherently and clearly to peers, teachers, and others*, *analyze and evaluate the mathematical thinking and strategies of others*, and *use the language of mathematics to express mathematical ideas precisely.*

Connections—connect mathematical ideas. The students should be able to *recognize and use connections among mathematical ideas*, *understand how mathematical ideas interconnect and build on one another to produce a coherent whole*, and *recognize and apply mathematics in contexts outside of mathematics.*

Representation—the act of capturing a mathematical concept or relationship in some form and the form itself. The students should be able to *create and use representations to organize, record, and communicate mathematical ideas*, *select, apply, and translate among mathematical representations to solve problems*, and *use representations to model and interpret physical, social, and mathematical phenomena.*

It is important to note that the concept of *reasoning* is not explicitly defined in this framework. The context in which the word is used in NCTM (2000) indicates that it denotes some kind of high-quality analytical thinking.

> "People who reason and think analytically tend to note patterns, structure, or regularities in both real-world situations and symbolic objects; they ask if those patterns are accidental or if they occur for a reason; and they conjecture and prove. (...) Being able to reason is essential to understanding mathematics. By developing ideas, exploring phenomena, justifying results, and using mathematical conjectures in all content areas and—with different expectations of sophistication—at all grade levels, students should see and expect that mathematics makes sense" (p. 55).

3.1.3 Schoenfeld's framework for characterisation of mathematical problem-solving performance

Schoenfeld (1985) have introduced a framework for adequate characterisation of mathematical problem-solving performance. The framework describes knowledge and behaviour necessary to do this characterisation according to four categories: *resources*, *heuristics*, *control*, and *belief systems*. This framework is appropriate when e.g. discussing the possible consequences of different types of teaching or student activities on the students problem solving competence (see Section 3.2.4). This competence is often one of the selected competences for the mathematics students in major theoretical frameworks (NCTM, 2000; Niss and Jensen, 2002). Below is a relatively short presentation of the four categories of knowledge and behaviour in Schoenfeld's framework.

Resources is described as mathematical knowledge possessed by the individual that can be brought to bear on the problem at hand. The knowledge is exemplified as intuitions and informal knowledge regarding the domain, facts, algorithmic procedures, "routine" non-algorithmic procedures, and understandings (propositional knowledge) about the agreed-upon rules for working in the domain.

Schoenfeld use *Heuristics* to denote strategies and techniques for making progress on unfamiliar or nonstandard problems; rules of thumb for effective problem solving. This includes drawing figures, introducing suitable notation, exploiting related problems, reformulating problems, working backwards, and testing and verification procedures.

With *Control* Schoenfeld means global decisions regarding the selection and implementation of resources and strategies. Relevant examples are planning, monitoring and assessment, decision-making, and conscious meta-cognitive acts.

The concept *Belief Systems* deals with one's "mathematical world view," the set of (not necessarily conscious) determinants of an individual's behaviour, e.g. about self, about the environment, about the topic, and about mathematics.

3.1.4 Vinner's conceptual framework for analytic and pseudo-analytic thought processes

Vinner (1997) presents a framework for the characterisation of different types of reasoning or thought processes. It is described as "a conceptual framework within which some common behaviors in mathematics learning can be described and analyzed" (p. 1). The author introduces the concepts *pseudo-analytical* and *pseudo-conceptual* and discusses examples from mathematical classrooms, homework assignments, and exams in relation to these concepts. The author does not define the concept of *analytic*. Vinner (1997) writes:

> "Most of the situations in mathematics education are problem solving situations, usually, routine problems.[3] In these situations a problem is posed to the students, and they are supposed to choose the solution procedure suitable for the given problem. The focus is not on *why* a certain procedure does what it is supposed to do. The focus is on *which* procedure should be chosen in order to solve the problem, and on *how* to carry out that procedure. The question *why* is usually irrelevant to these situations. The intellectual challenge lies in the correct selection of the solution procedure. The student is expected to be *analytical*" (p. 110).

In this short presentation of the framework, the focus will be on different types of behaviour. The concept of *pseudo-analytical* behaviour is introduced in contrast to analytical behaviour. Vinner suggests the following diagrammatic models:

Analytic A student who is supposed to solve a routine problem taken from a mathematical problem repertoire has to have the following:
• A pool of solution procedures.
• Mental schemes by means of which the type of a given mathematical problem and its particular structure can be determined.
• Mental schemes by means of which a solution procedure can be assigned to a given mathematical problem whose type and structure were previously determined.

Pseudo-analytic A student who solves problems using pseudo-analytical behaviour has to have the following:
• A pool of typical questions and their solution procedures.
• Mental schemes by means of which a similarity of a given question to one of the questions in the pool can be determined.

Two important features of the pseudo-analytical model are described by Vinner. Firstly, one of the most important characteristics of pseudo-analytical behaviour is the lack of control procedures. Vinner continues: "The person is responding to his or her spontaneous associations without a conscious attempt to examine them. The moment a result is obtained there are no additional procedures which are supposed to check the correctness of the answer" (p. 114). This aspect is similar to the control category introduced by Schoenfeld (1985). The description is connected to Vinner's suggestion that a student using analytical behaviour focuses on whether an answer to a task is correct, while a student

[3]Vinner uses a different definition of the concept "problem" and lets the word denote a mathematical task of any kind.

using pseudo-analytical behaviour focuses on whether the answer will be credited by the teacher. Secondly, Vinner claims that another characteristic of the pseudo-analytical model is simply that it is easier and shorter than the analytical.

Vinner argues that the most important goal for the students is to achieve an answer, preferably using as little effort as possible. He calls this the *minimum effort principle*. The principle implies according to the author that the pseudo-analytical behaviour, being shorter and lacking control procedures, is more effective in relation to the students' primary goal. Vinner believes that the pseudo-analytical behaviour is very common and this belief is connected to the minimum effort principle.

3.2 Context for the studies

In several studies of situations where students solve mathematical problems, the students end up in a situation where they are unable to carry on. Analyses of their work imply that they are hindered by their own inability to reason—or inability to somehow choose to reason—creatively, i.e. to consider the intrinsic mathematical objects of the task and their properties, and then try a solution attempt based on these objects and properties. Instead, they often try to apply different familiar algorithms—unsuccessfully since the task is nonstandard and cannot be solved with such algorithms—until they seem to find all options exhausted. This single-minded focus on imitative reasoning might be one of the main reasons for students' difficulties in mathematics. Empirical studies performed by members of the research group at Umeå university, several within the "Meaningful mathematics" project, indicate that students primarily choose imitative reasoning in task solving situations (Lithner, 2000a,b; Palm, 2002; Lithner, 2003; Bergqvist et al., 2003).

Several examples of imitative reasoning are described in the reasoning framework (Lithner, 2006) and a couple of these examples are mentioned in Section 3.1.1: the pupil using the key word strategy when solving "more"-tasks (see page 20) and Sally using delimiting algorithmic reasoning to solve a maximisation task (see page 21). The two examples might seem very different, but actually have a lot in common. Both Sally and the pupil apply familiar algorithms to solve a mathematical task. Sally and the pupil do not consider the intrinsic mathematical properties of the objects in the tasks in order to produce a solution—correct or incorrect—but simply choose between one or several familiar algorithms. Sally's reasoning and similar examples are presented in a study by Bergqvist et al. (2003). The study focuses on the students' grounds for different strategy choices and implementations. The results indicate that mathematically well-founded considerations were rare and that different types of imitative reasoning were dominating.

The reasoning framework (Lithner, 2006) is aligned with studies and theories regarding students' reasoning all over the world (Skemp, 1978; Schoenfeld, 1991; Tall, 1996; Vinner, 1997; Verschaffel et al., 2000; Palm, 2002). Skemp (1978) introduced the concepts "instrumental understanding"—knowing how—and "relational understanding"—knowing how *and* why—where relational understanding might be seen as necessary in order to perform creative reasoning. Skemp argues that instrumental understanding of a mathematical procedure can, to a spectator, appear to be a display of 'true' understanding, but is in fact only a skilled use of algorithms without any insight into why they work. Similarly, as presented in Section 3.1.4, Vinner (1997) discusses a "pseudo-analytical" behaviour that results in shorter and less demanding solutions than an analytic behaviour. He argues that since the students primarily search an answer to the task, they will choose pseudo-analytical behaviour over analytical.

Several important questions follow these empirical observations and theoretical explanations. Some of them will be addressed in the following sections: Why do the students[4] choose this type of reasoning instead of considering the intrinsic mathematics in the tasks? To what extent is it possible for the students to be successful in their mathematics studies using only imitative reasoning? What are the consequences for the students' mathematical development if only imitative reasoning is demanded from the students?

The first section will however focus on the role of tests and their influence on students' reasoning.

3.2.1 The influence of tests on students' reasoning

Exams are a part of the students' learning experience and several studies show that assessment in general influence the way students study (Kane et al., 1999). An exam provides an occasion when the students engage in solving mathematical tasks, just as they do when they solve textbook exercises. The time they spend solving these exam tasks is admittedly much shorter than the time they spend solving textbook exercises, but exams are often used to determine the students grades and might therefore be regarded as more important by the students.[5] Thus, the students probably pay attention to what type of reasoning that is generally required to solve exam tasks. The exams are also sources of information for the students when it comes to the concept of mathematics. The students know that the exams are constructed by professional mathematicians, and the content and design is judged by these mathematicians to be suitable for testing the students' knowledge on a specific subject. It is therefore reasonable to believe that

[4]The word "student" will from now on be used to denote both university students and school pupils.

[5]This is especially true for students at university level in Sweden since it is common that passing the exam is equal to passing the course. The Swedish system and its consequences are discussed in more detail in Section 3.3.

the students' beliefs concerning, and conceptions of, mathematics are affected by the exams and their design.

3.2.2 Why do the students choose imitative reasoning when it does not work?

The word "choose" in the headline does in this context not mean that the students necessarily make a conscious choice or a well-considered selection, but rather that they have a subconscious preference for certain types of procedures. An important question is why students are engaged in imitative reasoning, especially in the situations when it is not working.

One reason that students might choose imitative reasoning in these cases is that the didactical contract (Brosseau, 1997) might allow them to not always succeed (Lithner, 2000b). As established by Lithner, there is an important difference between school tasks and the professional use of mathematics. In school:

> "One is allowed to guess, to take chances, and use ideas and reasoning that are not completely firmly founded. Even in exams, it is acceptable to have only 50 % of the answers correct and, if you do not, you will get another chance later. But it is absurd if the mathematician, the engineer, and the economist are correct only in 50 % of the cases they claim to be true. This implies that it is allowed, and perhaps even encouraged, within school task solving to use forms of mathematical reasoning with considerably reduced requirements on logical rigour" (p. 166).

At a university course in mathematics in Sweden, it is very common that the only thing determining whether a student is approved or not, is a written exam. It is also common that a student has to correctly solve approximately half of the tasks in the exams to pass. This is in alignment with the contents of the quote above.

Another reason that the students choose imitative reasoning could be that it usually is a successful method. This is suggested in a study containing a general overview of four first-year undergraduate Swedish mathematics students' main difficulties while working with two mathematical tasks (Lithner, 2000a). The author concludes in the summary:

> "The four students often focus mainly on what they can remember and what is familiar within limited concept images. This focus is so dominating that it prevents other approaches from being initiated and implemented. There are several situations where the students could have made considerable progress by applying (sometimes relative elementary) mathematical reasoning. [...] Maybe the behaviour described above has it origins in that this usually is the best way for

> students to work with their studies? At first sight it might be most efficient when entering a task to (perhaps without understanding) superficially identify the type of task, somewhat randomly choose one from the library of standard methods, apply the familiar algorithms and procedures, and finally check with the solutions section" (p. 93–94).

As described in Section 3.1.4 the concept of imitative reasoning bear some resemblance with Vinner's definition of *pseudo-analytic* behaviour (Vinner, 1997). Vinner describes a pseudo-analytical model that is much shorter than an analytic model designed to solve the same type of tasks. He argues that the preferable procedure to achieve a goal usually is *the minimal effort procedure*, and that the most important thing for the students is to achieve an answer. A shorter process equals a smaller effort, and thus the students choose the pseudo-analytical behaviour.

A third reason for the students' choices of methods is indicated in a recent article by Bergqvist and Lithner (2005). The study focuses on how creative reasoning can be simulated in demonstrations by the teachers, e.g. by explicit references to mathematical properties and components while demonstrating task solutions. The studied teachers sometimes do this, but in rather limited and modest ways, and instead focus on presenting algorithmic methods.

Assume that students prefer imitative reasoning because such reasoning demands a smaller effort than creative reasoning. Assume also that the students encounter very little creative reasoning and further that the school and university environments do not *demand* creative reasoning to such an extent that the students need it to e.g. pass exams or to get passing grades. Perhaps these circumstances are enough to support the students' persistent choice of imitative reasoning (even when it does not work, because it works often enough). Research concerning the degree of veracity in these assumptions is important when it comes to understanding what really is going on. The assumption that the students encounter very little creative reasoning is discussed in the following section.

3.2.3 To what extent do the students encounter creative reasoning?

Hiebert (2003) discusses the concept *opportunity to learn*. He argues that what the students learn is connected to the activities and processes they are engaged in. The teachers give the students the opportunity to learn a certain competence when they provide the students with a good chance for practising the specific processes involved in that competence. This means that students that never are engaged in practising creative reasoning during class, are not given the opportunity to learn creative reasoning. It is therefore crucial to examine to what extent the students encounter creative reasoning in textbooks and teaching. It is also relevant to study if and how creative reasoning is *demanded* of the students e.g. in tests, exams, and when the students' work is graded.

Within the "Meaningful mathematics" project, one of the goals has been to examine these questions through studies of e.g. textbooks (Lithner, 2003, 2004), teachers' pracice (Bergqvist and Lithner, 2005), and tests (Boesen et al., 2005; Palm et al., 2005). The results from these and other international studies are presented below. The two articles presented in this thesis are specifically focused on exams at university level (Article III) and university teachers' views on the required reasoning in exams (Article IV).

Textbooks

There are several reasons to believe that the textbooks have a major influence on the students learning of mathematics. The Swedish students—both at upper secondary school and at the university—seem to spend a large part of the time they study mathematics on solving textbook exercises. This is indicated by several local and unpublished surveys but also by a report published by the Swedish National Agency for Education (Johansson and Emanuelsson, 1997). The Swedish report on the international comparative study TIMSS 2003 (Swedish National Agency of Education, 2004) also shows that Swedish teachers seem to use the textbook as main foundation for lessons to a larger extent than teachers from other countries. The same study notes that Swedish students, especially when compared to students from other countries, work independently (often with the textbook) during a large part of the lessons.[6] At university level in Sweden it is common that the students have access to the teachers, via lectures or lessons, no more than 25 % of the time that they are expected to spend studying a course. All these results and circumstances imply that the textbooks play a prominent role in the students learning environment.

The latter of the textbook studies mentioned above indicates that it is possible to solve about 70 % of the exercises in a common calculus textbook with text-guided algorithmic reasoning (Lithner, 2004).

Teachers' practice

The teachers' practice, especially what they do during lectures, is another factor that affects the students' learning. As was described in the previous section (3.2.2), a study of how creative reasoning was demonstrated via simulation by the teachers showed that this was done only to a small extent (Bergqvist and Lithner, 2005). Teachers also often argue that relational instruction is more time-consuming than instrumental instruction (Hiebert and Carpenter, 1992; Skemp,

[6] The TIMSS 2003 study referred to in this case treat students, and teachers of students, in grade 8 (students approximately 14 years old).

1978). There are however empirical studies that challenge this assumption, e.g. Pesek and Kirshner (2000).

Vinner (1997) argues that teachers may encourage students to use analytical behaviour by letting them encounter tasks that are not solvable through pseudo-analytical behaviour. This is similar to giving the students the opportunity to learn CR by trying to solve CR tasks. Vinner comments, however, that giving such tasks in regular exams will often lead to students raising the 'fairness issue,' which teachers try to avoid as much as possible. He concludes that this limits the possible situations in which students can be compelled to use analytical behaviour.

Tests and assessment

As mentioned, there are many reasons that tests are important to study, and several studies concerning tests and exams have been carried out within the research group at Umeå university. Palm et al. (2005) examined teacher-made tests and Swedish national tests for upper secondary school. For students at a mathematics course at this level in Sweden, the national test is one of many tests that the students participate in, the other tests are teacher made. The focus of the study was to classify the test tasks according to what kind of reasoning that is required of the students in order to solve the tasks. The analysis showed that the national tests require the students to use creative reasoning to a much higher extent (around 50 %) than the teacher made tests (between 7 and 24 % depending on study programme and course). The results from these studies indicate that upper secondary school students are not required to perform creative reasoning to any crucial extent.

This result is in alignment with other studies indicating that teacher-made tests mostly seem to assess some kind of low level thinking. An example is an analysis of 8800 teacher-made test questions, showing that 80 % of the tasks were at the "knowledge-level" (Flemming and Chambers, 1983). Senk et al. (1997) classify a task as *skill* if the "solution requires applying a well-known algorithm" and the task does not require translation between different representations. This definition of skill has many obvious similarities with Lithner's definition of algorithmic reasoning. Senk et al. (1997) report that the emphasis on skill varied significantly across the analysed tests—from 4 % to 80 % with a mean of 36 %. The authors also classified items as requiring *reasoning* if they required "justification, explanation, or proof." Their analysis showed that, in average, 5 % of the test items demanded reasoning (varying from 0 % to 15 %). Senk et al. (1997) also report that most of the analysed test items tested *low level thinking*. This means that they either tested the students' previous knowledge, or tested new knowledge possible to answer in one or two steps. In the introduction, the authors mention a number of research results of interest in the context of this thesis. An example is a number of American reports indicating that standardised tests focus on low-level thinking. The authors refer to some scholars arguing that the teachers that aim at preparing their students for these tests, cannot prioritise the type of instruction the students would need to develop the type of competences described in the Standards (NCTM, 2000). Similarly, the authors mention a couple of studies of textbook publishers' tests, showing that the tests neither measured the content nor the process standards defined in NCTM (1989). Senk et al. (1997) further mention four American studies of high school teachers' practice that indicate that short-answer tasks is the basis for much of the teachers' assessment.

3.2.4 What might be the consequences of the students' persistent choice of imitative reasoning?

Algorithms or not?—the Math Wars

The discussion on whether teaching algorithms is good or bad, or maybe even harmful, for the students and their mathematical development is not new and has in the United States sometimes been called "the Math Wars" (Schoenfeld, 2004). Several researchers within mathematics education have shown examples of how students that work with algorithms seem to focus solely on remembering the steps, and have argued that this weakens the students' understanding of the underlying mathematics (Leinwand, 1994; Burns, 1994; McNeal, 1995; Kamii and Dominick, 1997; Pesek and Kirshner, 2000; Ebby, 2005). This research can be seen as aligned with NCTM (1989), where 'memorising rules and algorithms' was one of the topics recommended decreased attention. On the other side of the debate resides mostly mathematicians and teachers that argue e.g. that alternative approaches of teaching algorithms should be considered before banishing them completely from the classroom (Wu, 1999) and that standard algorithms of arithmetics play a critical role in the students' mathematical development (Ocken, 2001).

Rote learning might hinder later conceptual learning

My intention with this thesis is to examine a part of the Swedish university system from the standpoint that *students that only focus algorithmic solutions of mathematics tasks will eventually limit their resources when it comes to other parts of mathematics e.g. problem solving.* This standpoint is to some extent supported by the empirical research articles mentioned above (as arguing against the teaching of algorithms). One example is the study by Pesek and Kirshner (2000) where it was shown that a group of students who received only relational instruction outperformed a group of students who received instrumental instruction prior to relational instruction. The authors conclude that "initial rote learning of a concept can create interference to later meaningful learning." Similar results was presented in a case study by Ebby (2005) where the author followed a student during three school years, from grade two to grade four of elementary school. Ebby concludes from her findings that "for some children, learning to use the algorithm procedurally actually *prevents* them from learning more powerful mathematical concepts" (p. 85).

Limitation of problem-solving competence

Another argument that supports this standpoint is presented by Lithner (2004) who discusses how a single-minded use of a type of imitative reasoning called *text-guided AR* might have negative consequences on students' problem solving competence. Lithner presents his argument according to the framework on problem solving provided by Schoenfeld (1985) (see also Section 3.1.3). Lithner states that since the intrinsic mathematical properties of a task do not have to be considered when using text-guided AR, the *resources* that may be developed are "restricted to surface mathematical areas." He further argues that this type of imitative reasoning does not satisfactory support the students development of *heuristic strategies* and *control*, since text-guided AR do not really demand either of those. Even the students' *beliefs* concerning the nature of mathematics might be affected, since a persistent use of only imitative reasoning might strengthen "the common belief that mathematics is about following procedures developed by others." Such a belief may hinder the student from even attempting own constructions of solutions to mathematical problems, a major element in problem solving.

Affects on students' beliefs

Similar conclusions concerning students' beliefs were obtained in a case study by McNeal (1995). She followed an elementary school student, Jamey, when he participated first in a second grade inquiry-based mathematics class and then in a textbook-based class in third grade. In second grade most of the children invented their own procedures for adding two-digit numbers and through interviews and analyses of the Jamey's reasoning, it could be shown that he believed that "addition should make sense in terms of actions on mathematical objects." After only eight weeks in third grade, Jamey's beliefs about mathematics—or about his obligations in school—had changed. He appeared at this time to instead think of mathematics as "a set of arbitrary rules to be applied to the digits in arithmetic problems without having a meaning that connected to his own understanding of numbers." McNeal also states that Jamey was "confused by the obligation to recall steps of a procedure without the opportunity to clear up misunderstandings." The presented results show that Jamey's beliefs concerning mathematics changed due to "the repetition over time of actions that did not fit with his prior beliefs, and the subtle confirmation of his new beliefs and expectations as his responses were praised, ignored, or dismissed." No single lesson will change a student's beliefs, but a constant focus on imitative reasoning—from teachers, textbooks, and tests—might narrow the student's idea of what mathematics really is.

3.3 The Swedish system

3.3.1 Practical issues

Sweden has approximately 9 million inhabitants. University studies in Sweden are free for the students and the they often support themselves via government funded post secondary student aid (available to everyone). There are a little more than 40 university colleges of which 12 are universities (in Sweden a university college is called a university only if the government have granted the right to issue degrees at graduate level) that provide courses in mathematics. The mathematics teachers at the universities have either a PhD in mathematics or mathematics courses approximately corresponding to a masters degree. They usually have a PhD if they teach graduate courses.

It is quite common that the students go through one course at a time and perhaps four courses each semester. In this case each course lasts 4–5 weeks and often contains daily lectures and a couple of lessons providing time for questions per week. At the end of each course there is often a final examination that lasts about six hours and usually consists of 6-10 exam tasks (each task may be divided into two or three parts). The result from the final examination very often determines the students grades completely, but is sometimes combined with results from smaller examinations, focusing on parts of the course content, offered to the students during the course. A student that passes the exam (or exams) is given credits for the course.

The students performance is often graded: *failed*, *passed*, or *passed with distinction*. The grade is based on the students' results, i.e. the number of points, on the final examination. The maximum number of points is often between 20 and 40. The grade *passed* often corresponds to requiring approximately half of the maximum points and the grade *passed with distinction* often corresponds to 75 % of the points. It is common that the lecturer grades the exams him-/herself, usually together with any other teachers involved in the course.

This setting naturally varies between universities, courses, and teachers. There are e.g. parallel courses (i.e. two courses lasting for 8-10 weeks at half the pace), other types of examinations, computer workshops affecting the grades, or other types of grades (e.g. the scale *failed, 3, 4*, and *5*).

3.3.2 Consequences for Swedish students and teachers

In the Swedish system, the knowledge and skills that a student (that passed) will be equipped with after a course is directly related to what is tested in the exam, and how. This "threshold"-function strengthens the assumption that the design of the exams affect both teachers and students. In Sweden, passing the mathematics exams at university level is also important for the students from an economic perspective. They need to pass the exams to be granted continued post secondary student aid. This situation adds a social and economical weight to the exams importance for the students. There is another economical perspective from the teachers point of view. The percentage of students that pass the exam is relevant when it comes to each department's funding. A high passing rate might be important to the teachers' own futures, at least in the long run.

3.4 Some aspects concerning the methods used in the articles

3.4.1 The classification tool

Since students seem to follow the minimum effort principle (Vinner, 1997) and use imitative reasoning (IR) to a great extent, it is important to determine to what degree this method may be successful. Article III examines the reasoning that Swedish university students in mathematics are expected to perform in order to pass exams, see Section 3.5.1. This is done through the analysis of over 200 introductory calculus exam tasks. Each task was classified according to the type of reasoning it required the students to perform. The analysis was performed using a theoretical classification tool designed to classify upper secondary school test tasks (Palm et al., 2005).

The following subsections contain a discussion of the validity of the classification tool and a presentation of the classification procedure.

The rationale behind the classification tool

The reasoning framework Lithner (2006) was designed for the analysis of different types of reasoning, e.g. the reasoning a student uses when solving an exam task. Determining what type of reasoning an exam task demands from the students is quite different, especially since there is no actual reasoning to analyse. When classifying tasks it is therefore necessary to consider the relation between the task and the student that the task is designed for, cf. (Schoenfeld, 1985).

The basic idea of the classification tool is to determine if it is possible for the students to solve a specific task using IR, or if creative reasoning (CR) is required. The classification tool uses information concerning the students' learning experiences in order to determine how *familiar* a task is to the students. Since it is not possible to determine the students' *complete* learning experience—each exercise solved, each lecture attended, each piece of information offered, etc.—it is necessary to simplify the actual situation. The analysis is therefore mainly based on the content of the course textbook. The classification procedure includes a comparison between each task (including its solution) and the textbook: the theoretical content, the exercises recommended by the teacher, and any lists of important definitions, theorems, and proofs that the teacher handed out during the course. Using the textbook as the basis for the analysis is reasonable since the Swedish students—both in upper secondary school and at the university—seem to spend a large part of their time solving textbook exercises, see Section 3.2.3. The textbooks' importance is also implicated by the university students' limited access to the teacher, and that they often are expected to study independently more than 50 % of a typical 'workday.' The classification procedure does not take old exams into consideration, even though they often are a source of information for the students. The reason is simply that it is very difficult to find out what exams were available for a particular group of students. An inclusion of the relevant previous exams in the classification procedure would result in more tasks being judged as familiar to the students. The classification results will therefore at least not exaggerate the percentage of tasks solvable through IR.

The validity of the classification tool is supported by a follow-up study carried out by Boesen et al. (2005) where the reasoning requirements established by the classification tool were compared to students' actual reasoning. This was done with Swedish national tests at upper secondary school and the results showed that the students' reasoning followed the reasoning requirements established by the classification tool to a high extent. In 74 % of the analysed solution attempts, the students actually used the type of reasoning previously determined by the classification tool. In 18 % of the attempts a less creative type of reasoning was used, and in 8 % of the cases a type of reasoning with more CR was used. These results indicate that the classification tool, in spite of the simplifications discussed above, point at what type of reasoning that is demanded of the students.

The basic idea of the classification procedure

As mentioned, the goal of the classification was to determine to what extent it is possible for the students to successfully use IR when solving exam tasks. As a consequence, the tasks that required creative reasoning (CR) were sorted into two different classes depending on how much creative reasoning they required. If a task was almost completely solvable using IR, and required CR only in a local modification of the algorithm, the task was said to require *local creative reasoning* (LCR). If the task had no solution globally based on IR and therefore demanded CR all the way through, it was said to require *global creative reasoning*

(GCR). Parts of the solution to a GCR task could be based on local IR. Since LCR tasks are, by definition, possible to partly solve using IR, the separation between LCR and GCR was necessary in order to capture to what extent CR is in fact demanded in the exams.

A task is classified as an imitative reasoning task, MR or AR, if circumstances indicate that the students are familiar with a task's formulation and an algorithm that solves the task. A task is instead classified as a creative reasoning task, LCR or GCR, if the task is judged not to be solvable with IR by the students and there is a reasonable CR solution. The task is classified as LCR or GCR depending on to what extent CR is needed. The classification tool consists of a step-by-step procedure of systematical examination of the situation, in order to determine the students' familiarity to each task. The procedure identifies any *occurrences* of the task, i.e. very similar tasks and solutions in the theory text, as examples, or as a recommended exercises.

A task was judged to be ***familiar*** to the students if it:
• asked the student to state a fact or a theory item, e.g. a definition or a proof, that the students had been informed during the course might be asked for at the exam (solvable with MR) *or*
• the textbook contained at least three (3) occurrences of the task, and each of these occurrences shared enough characteristics with the task to make it possible for the students to identify the applicable algorithm (solvable with AR).

Many students probably need to encounter a type of task more than once before they are able to use familiar AR (or delimiting AR) to solve such a task during an exam. That the minimum number of necessary occurrences in the classification tool was set to be three, is obviously an arbitrary choice. The choice is however supported by the validation of the tool presented in the previous section. The students can sometimes use text-guided AR to solve an exam task if the exam conditions include a textbook or some other source of solved examples or formulas. This was not the case for any of the exams in Article III.

The similarity between a task and each possible occurrence was examined by the use of so called *task variables*, e.g. the number of words in a task or the task content domain. The task variables together describe the distinguishing features of the task in a systematic way. The variables are then used to determine the number of occurrences of the task in the textbook, and to argue whether the task was familiar to the students or not. The step-by-step classification procedure is outlined and exemplified below.

The step-by-step procedure of the classification tool:

1. **Analysis of the exam task**
   The task variables for the exam task are determined.
   (a) *A solution*—an answer or an algorithm that solves the task
   (b) *The context*—e.g. the real-life situation
   (c) *Explicit information about the situation*—e.g. a mathematical function in the task
   (d) *Other key features*—e.g. key words, syntactic features, hints
2. **Analysis of the textbook occurrences**
   The task variables for the textbook occurrences are compared to the task variables for the exam task.
   (a) *Occurrences in examples and exercises*—including differences and similarities between the task and the occurrence (based on the four task variables)
   (b) *Occurrences in the theory text*—including differences and similarities between the task and the occurrence (based on the four task variables)
3. **Argumentation and conclusion**
   (a) *Argument for the requirement of reasoning*—The similarities and differences between the task and the occurrences are discussed, counted, and used to argue for whether the task is familiar or not. The examples and exercises that are similar to the task, but not similar enough, are also discussed.
   (b) *Conclusion on reasoning demand*—The task is classified as MR, AR, LCR, or GCR depending on the previous argumentation.
4. **Comment**
   The task and its classification is commented upon if there is anything relevant to the classification that is not included in the previous steps.

Example of the classification of an MR task: 'Prove a theorem'

1. **Analysis of the exam task**
   "Prove the Mean-value theorem. Rolle's theorem does not need to be proved."
   (a) *A solution*
   Proof: Suppose f satisfies the conditions of the Mean-Value Theorem. Let $g(x) = f(x) - (f(a) + \frac{f(b)-f(a)}{b-a}(x-a))$. For $a \leq x \leq b, g(x)$ is the vertical displacement between the curve $y = f(x)$ and the chord line $y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ joining $(a, f(a))$ and $(b, f(b))$. The function g is also continuous on $[a, b]$ and differentiable on (a, b) because f has these properties. In addition, $g(a) = g(b) = 0$. Since $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$, it follows that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

(b) *The context*
None

(c) *Explicit information about the situation*
None

(d) *Other key features*
None

2. **Analysis of the textbook occurrences**

   (a) *Occurrences in examples and exercises*
   None

   (b) *Occurrences in the theory text*
   The theorem and the proof is stated in the textbook.

3. **Argumentation and conclusion**

   (a) *Argument for the requirement of reasoning*
   The proof of the theorem is on a list handed out by the teacher and the students were told that the list contained the theorems and proofs that might be asked for during the exam. The task is therefore familiar to the students, and possible to solve using memorised reasoning.

   (b) *Conclusion on reasoning demand*
   Memorised reasoning (MR)

4. **Comment**
   It might be possible for the students to use creative reasoning (CR) to deduce some of the, maybe forgotten, steps of the proof. It the proof was shorter and the theorem less complex, it could perhaps have been possible for the students to use global creative reasoning to solve the task (except for the stating of the theorem itself). For two main reasons, the task is still classified as an MR task. Firstly, as discussed in Section 3.1.1 in relation to memorised reasoning, the students seem to memorise long and complex proofs without being able to explain even quite simple parts of the proof. Secondly, the handed out list makes it *possible* for the students to use MR to solve the task.

Example of the classification of a GCR task: 'Give an example'

1. **Analysis of the exam task**
   "Give an example of a function that is left continuous but not continuous at $x = 1$. Draw the graph of the function."

   (a) *A solution*
   Choose a value for the function in $x = a$ where a is the given point. Choose a continuous function g that attain the chosen value at the given point and a continuous function h that attain another value at the given point. Define the answer f as

$$f(x) = \begin{cases} g(x) & x \leq a \\ h(x) & x > a \end{cases}$$

(b) *The context*
None

(c) *Explicit information about the situation*
None

(d) *Other key features*
"left continuous", "continuous", "draw the graph"

2. **Analysis of the textbook occurrences**

(a) *Occurrences in examples and exercises*
None. Some exercises ask the students to determine whether a given function has a certain continuity property or not, but those exercises have other solutions.

(b) *Occurrences in the theory text*
None

3. **Argumentation and conclusion**

(a) *Argument for the requirement of reasoning*
There are a lot of functions in the theory text, in examples, and in exercises, that have properties similar to those asked for in this task. These occurrences usually demonstrate how to, or ask the students to, determine whether a given function has a certain continuity property or not. There are no exercises in the textbook that ask the student to construct a function with given properties, at least not when the properties concern continuity in any form. The students have not encountered an algorithm or procedure that is possible to use to solve this task, and there is nothing on the list of theory items corresponding to the task's assignment. Therefore, the task is *not familiar* and not possible to solve using IR. Since there is no familiar algorithm that the solution can be based on, the task requires global creative reasoning.

(b) *Conclusion on reasoning demand*
Global creative mathematically founded reasoning (GCR)

4. **Comment**
According to the analysis of the textbook, several exercises ask the students to determine whether a given function have some continuity property or not. In this task, the student is instead asked to construct a function with a given continuity property. For a student capable of using (and trying to use) the intrinsic mathematical properties of the objects in this task, it might be quite straight-forward to understand what is asked for. Constructing the solution might still be complicated, since the task is formulated "backwards" compared to what the student is used to. The task is not familiar to the students and may also, but does not have to, be difficult for them to solve.

3.4.2 Choice of frameworks

Choice of general framework

There are in the literature several general frameworks describing mathematical competences from different perspectives, e.g. NCTM (2000); Niss and Jensen (2002); Mullis et al. (2003). All of these frameworks describe mathematics as a subject consisting of different types of knowledge and competences and are in many ways very similar. In choosing a framework as basis for discussions about and/or analyses of competences within mathematics, there is in fact several possible choices that would all suffice. Article IV contains an analysis focusing what types of competences that the teachers describe or mention during the interviews. As a basis for this analysis, the NCTM framework NCTM (2000), abbreviated "the Standards," was chosen. The Standards is a comprehensive, research-based, and thoroughly worked-out framework that, even though it is not primarily designed as a research tool, is highly appropriate for the kind of analysis taking place in this case. As indicated, the Standards was chosen as one of many possible and appropriate ones, all compatible with the chosen reasoning framework.

Choice of reasoning framework

The general comprehensive frameworks are usually not intended primarily as research tools, and they are often too vague to be used as the basis for more detailed analyses of mathematical reasoning. There exist several frameworks that can be used to categorise the understanding that lies behind a student's reasoning, e.g. Skemp (1978) and Sfard (1991), but these do not aim at categorising the reasoning specifically. The framework proposed by Vinner (1997) focuses reasoning, and also to a large extent the types of reasoning that are treated within the articles in this thesis. As is clear from previous sections, Vinner's arguments are very much aligned with what is discussed in Lithner's framework.

However, Vinner's framework is not as useful as Lithner's framework for a couple of reasons. Firstly, Vinner does not give a characterisation or definition of the concept *analytical* behaviour. As a consequence, it is difficult to use the framework to separate between analytical and pseudo-analytical behaviour when analysing students' reasoning and exam tasks. Secondly, the vocabulary is quite extensive and partly confusing. The framework focuses on *thought processes* and *behaviour*, but also use expressions like *model, line of thought, mode of thinking, mode, strategies,* and *abilities.* The presentation is therefore partly difficult to follow and it makes the framework less useful in concrete classification situations.

The reasoning framework proposed by Lithner (2006) is less vague and presents explicit characteristics for both creative and imitative reasoning. It focuses only on the performed reasoning and not the underlying thought processes, which makes it more concrete. Lithner's reasoning framework is also more explicitly founded on empirical studies and describes key features of students' reasoning according to these studies. All in all, the framework suggested by Lithner is more adequate for the studies performed within this thesis.

3.5 Summaries of the articles in Part B

3.5.1 Article III

Title: Types of Reasoning Required in University Exams in Mathematics

The purpose of the study was to examine the reasoning required by university students and to examine if is it possible for students to solely use imitative reasoning (IR) and still pass the exams. To this end, all introductory calculus exams produced during the academic year 2003/2004 was collected from the author's university and from three randomly selected other Swedish universities. More than 200 tasks from 16 exams were then analysed according to a theoretical reasoning framework. The tasks were sorted into four task classes—memorised reasoning (MR), algorithmic reasoning (AR), local creative reasoning (LCR), and global creative reasoning (GCR)—using a theoretical classification tool based on the framework. For each class the types and solutions of the tasks were further examined in order to determine in what ways the different reasoning types were demanded.

The results show that about 70 % of the tasks were solvable by IR, i.e. MR or AR. It was further shown that approximately 90 % of the tasks were classified as MR, AR, or LCR. Since all LCR solutions were based on global AR, the students could solve most of the tasks using solutions based on IR. The results also show that 15 of 16 exams were possible to *pass* using IR. Another result was that 15 of 16 exams were possible to *pass with distinction* using IR and LCR, i.e. using solutions based on IR.

The MR tasks belonged all to one of three types: tasks having solutions being *definitions*, *theorems*, or *proofs*. The AR tasks could be divided into four different categories: tasks with solutions consisting of *basic algorithms*, *complex algorithms*, *choice-dependent algorithms*, or *proving algorithms*. All LCR tasks had, as mentioned, solutions based on global algorithms that were familiar to the students, but needed some small adjustment using CR. The analysis of the GRC tasks' solutions resulted in three task subcategories: tasks with solutions consisting of *a construction of an example*, *the proof of something new*, or *modelling*.

3.5.2 Article IV

Title: University Mathematics Teachers' Views on the Required Reasoning in Calculus Exams

The purpose of the study was to examine the teachers' views and opinions on the reasoning required by students taking introductory calculus courses at Swedish universities. Six of the teachers that constructed the exams classified in Article III were interviewed. The questions concerned how they construct exams, whether required reasoning is an aspect that they consider during exam construction, and how they regard the results from Article III. The interviews were transcribed and analysed.

The analysis show that the teachers are quite content with the present situation. During exam construction, they primarily consider the tasks' content domain and their degree of difficulty. The teachers express the opinion that the degree of difficulty is closely connected to the task type, and that tasks demanding creative reasoning (CR) are more difficult than tasks solvable with imitative reasoning (IR). The teachers therefore use the required reasoning as a tool to regulate the tasks' degree of difficulty, and not as a task property of its own. The teachers think that the students today have less previous knowledge than they used to. In combination with the teachers' lack of time during exam construction and teaching, the teachers feel that it is unreasonable to demand CR to any higher extent.

A possible answer to the general question of why the exams are designed the way they are, with respect to required reasoning is therefore: *the exams demand mostly imitative reasoning since the teachers believe that they otherwise would, under the current circumstances, be too difficult and lead to too low passing rates.*

3.6 Conclusions and Discussion

3.6.1 Conclusions

Article III

- Approximately 70 % of the exam tasks were possible to solve using only imitative reasoning
- Approximately 90 % of the exam tasks were possible to solve completely (the MR and AR tasks) or partly (the LCR tasks) using imitative reasoning
- It was possible for the students to *pass* 15 of 16 exams using only imitative reasoning
- It was possible for the students to *pass* 15 of 16 exams *with distinction* without using global creative reasoning

Article IV

- The teachers consider the required reasoning of a task to be closely related to the task's degree of difficulty
- The teachers use the tasks' required reasoning to regulate the degree of difficulty of the tasks
- The teachers say that the exams would be too difficult if creative reasoning was required to a higher extent

The results from Article III support results from earlier research: the students encounter creative reasoning only to a limited extent, and do usually not have to perform creative reasoning to pass exams. The teachers' views on the required reasoning presented in Article IV give a reasonable explanation to the required reasoning in exams at introductory calculus courses. The teachers are of the opinion that the exams would be too difficult if creative reasoning was required to a higher extent.

3.6.2 Discussion

There are a lot of interesting issues related to these results and some of them are discussed in the articles, e.g. if tasks demanding creative reasoning really have to be more difficult than tasks solvable with imitative reasoning (Article IV). A couple of other issues related to possible future research are discussed below.

The teachers' views of learning and assessment

An interesting discussion on the gap between modern theories of learning and how assessment today is designed is presented by Shepard (2000). The author gives a historical perspective on the development of learning theories and argues that traditional testing practices have their roots in outdated behaviourist ideas of learning. According to Shepard, testing and assessment have not embraced the development of the cognitive and constructivist learning theories, which include new thinking about assessment. The author describes a few key features of the behaviourist learning theories, such as learning occurs by accumulating atomised bits of knowledge, learning is sequential and hierarchical, and transfer is limited to situations with a high degree of similarity. Shepard argues that this type of principles, that were dominating within education for decades, is still dominating within assessment. She says: "For example, when teachers check on learning by using problems and formats identical to those used for initial instruction, they are operating from the low-inference and limited transfer assumptions of behaviourism," p.11 in Shepard (2000). The first key feature, that learning occurs by accumulating atomised bits of knowledge, Shepard illustrates with a quotation from Skinner (1954).

> "The whole process of becoming competent in any field must be divided into a very large number of very small steps, and reinforcement must be contingent upon the accomplishment of each step." (p. 94)

During the interviews of Article IV, one of the teachers mentioned that today's design of the exam tasks is a reflection of the need to "make it a bit more standardised and possible for the students in today's situation to handle. They need to get credit for knowing small pieces." This statement is aligned with the mentioned behaviourist views on learning and assessment. The teacher did admittedly not advocate this type of exam design, but acknowledged that it was common and perhaps inevitable. Since the university mathematics teachers' general view of learning and assessing is not known, it is difficult to determine if this statement is representative for the whole group, for this specific teacher, or not at all. This unclear situation supports a suggestion by Shepard (2000), that studies of the professional development of teachers is an important part of research on assessment.

It is still possible to discuss the teachers practical situation and its relation to the gap described by Shepard. In Sweden it is very common that the teachers at university level have not studied pedagogy, and many of them have been fostered into teaching in the academical environment simply through their own experiences. Whether learning theories and pedagogical methods are seen as relevant at a mathematics department or not, often depends on the staff's interest and experience, rather than a managemental commitment to pedagogical development. It is possible that there is no gap between the teachers' underlying view on learning in relation to the exams' design, but that their views do not mirror the modern ideas of learning and assessing. Another possibility is that teachers in general are well aware of modern learning theories, but that circumstances hinder them from putting the theories into practice. In Article IV, all teachers mention for example that time is a factor when they construct exams.

Further studies concerning university mathematics teachers' professional development and their views on learning and assessment are highly relevant and would be interesting to perform.

Teachers' awareness

In an interview study presented in a doctoral thesis (Boesen, 2006), the idea of the teachers' *awareness* is introduced. One of the author's aims with the study is to determine the Swedish upper secondary school mathematics teachers' awareness of creative reasoning requirements in mathematics tests. The teachers were introduced to six test items, four classified as algorithmic reasoning and two as creative reasoning, from the Swedish national course tests, and were asked to identify the most significant difference between the different tasks. They were also asked whether the tasks did require different forms of reasoning. The teachers descriptions of the differences between the tasks could be sorted into four different "levels of awareness", ranging from 0 to 3. Each level is connected to one or several of the teachers' statements and with *Level 0* corresponding to a teacher separating between the tasks only by mathematical content (task content domain), *Level 1* corresponding to a teacher viewing the tasks only as having different degrees of difficulty, *Level 2* corresponding to a teacher identifying the AR tasks as "routine" or "standard" tasks and the CR tasks as e.g. interesting or fun, and *Level 3* corresponding to a teacher stating that the CR tasks demand that the students understand the mathematics, in contrary to the AR tasks that are routine or standard. As Boesen (2006) states: "Although a specific teacher seldom fits exactly into one of these levels (...) An attempt to pair the eight teachers to one of these levels would give overweight at levels 0-2" (p. 136).

If these levels of awareness are applied to the teachers in Article IV, the result is interestingly enough different. All of the teachers in the study express an awareness corresponding to at least level 2, and one or two perhaps at level 3. This might of course to some extent depend on differences in method between the two studies, but it is possible that the difference is connected to the differences between the two groups of teachers. For example, the teachers in Boesen's study are upper secondary school mathematics teachers and have studied mathematics at university level, but probably not as much as the university teachers interviewed in Article IV. Most of the university teachers have a PhD in mathematics. These teachers, that have continued to study mathematics at graduate level, have encountered a lot more tasks, but they may also have encountered a lot more different types of tasks. If they have met a greater variety of tasks, this could have given them a higher awareness of the different reasoning types. One of the teachers in the interview study actually stated that at graduate level "it is exclusively global creative reasoning." There are several other differences between the two groups, e.g. the level of the mathematics they teach, the curricula, practical circumstances, and so on, that might affect their level of awareness to some extent.

The differences between the two groups of teachers are rather vague, and the reasons for the difference are also unclear. It would be interesting to examine possible explanations for the differences in teachers' awareness of required reasoning more closely.

3.7 Future research

The two aspects considered in the discussion, the teachers' awareness of required reasoning and their views on learning and assessment, give rise to ideas for future research. There are other loose ends that inspire to further studies.

As mentioned in Article III, three different task classes, memorised reasoning, algorithmic reasoning, and global creative reasoning, contain tasks requesting that the students prove something. The solution methods to the tasks in the different task classes are obviously different, but the wording of the tasks are often very similar. In the Swedish national tests for upper secondary school mathematics, the tasks with this wording probably have to be solved using creative reasoning. This claim is based on the assumption that the upper secondary school students are not informed in advance of any theory that might be asked for during the test, and they usually do not practise applicable algorithms. That tasks with this wording can be solved with AR or MR is a new situation to the students at introductory calculus if this assumption is true. It is difficult to say how and if this possible new situation affects the students. It would be interesting to study the assumption, the claim, and the possible consequences.

The situation described in the discussion on the teachers' awareness of required reasoning is interesting and complex. The teachers in the two groups seem to have different mathematical background, but there may also be factors connected to the contents of the curricula. The upper secondary school teachers are supposed to teach problem solving and reasoning according to the curricula, but have probably not encountered as much creative reasoning (CR) tasks as the university teachers. The university teachers have probably more experience of CR tasks and seem more aware of the aspect. However, they do not seem compelled to focus on CR, at least not by the curricula, and their exams do not contain a significantly larger percentage of CR tasks. The situation is therefore not only complex but also partly contradictory. Further research of course curricula in relation to the teachers' mathematical experience and task construction would be interesting.

Another idea for future research is connected to the validity of the classification tool. The validity was examined in a study concerning upper secondary school test tasks (Boesen et al., 2005). Since the university students partly are in a different situation than the upper secondary school students, it would be interesting to carry out a similar study at university level.

As mentioned earlier, one of the teachers expressed in Article IV the opinion that the courses at graduate level contain almost only global creative reasoning. It would be interesting to study the required reasoning in exams connected to courses on higher levels than introductory calculus, but also to examine the students' views on different types of reasoning. To what extent do the students perceive the differences between tasks solvable with imitative reasoning and tasks demanding creative reasoning? How do the students view or handle the different types of tasks when they study? Studies of this types of questions could enhance the understanding of the impact of the required reasoning in textbooks and exams.

3.8 Epilogue

The results of Article III and IV show that university undergraduate students are not required to perform creative reasoning to pass introductory level calculus exams. I suspect that classifications of exam tasks from other undergraduate courses would give similar results. Moreover, I believe that the students are not given many opportunities to learn creative reasoning during their undergraduate studies.

As mentioned in Section 3.2.3, teachers often argue that relational instruction is more time-consuming than instrumental instruction. I think that teachers similarly believe that teaching CR is more time-consuming than teaching IR. My guess is that this belief is connected to what one of the teachers in Article IV said: that "if there are algorithms to learn, the students often have an ability to learn quite difficult things." This is an interesting observation with possibly important consequences. That the students can use IR to solve tasks that are based on advanced mathematical concepts may give the impression that the students have learnt advanced mathematics. But since it is possible to use IR without considering the intrinsic mathematical properties of the tasks, this impression is not reliable. One of the major differences between CR and IR, is that CR is mathematically founded. This imply that IR tasks can include more advanced mathematical concepts than CR tasks and still be possible to solve for the students. A possible consequence of this is that if the courses should give the students more opportunities to learn CR, the mathematical content may have to be reduced. The students would not be able to handle as advanced mathematical concepts as the teachers are used to, but they could probably have a better conceptual understanding of the concepts that they did handle.

It is possible that it is more time-consuming to teach CR than to teach IR, at least if the same mathematical concepts should be included. The question is what we want the students to learn. How important it is that they learn a large repertoire of advanced algorithms? How important it is that they develop their conceptual understanding? There is probably no simple answer, but I believe that the situation today is too imbalanced. The students focus almost completely on learning to implement algorithms, and have neither the opportunity nor the incentive to reason creatively.

There is a lot to gain by letting the students become more familiar with solving unfamiliar tasks and I think that a more balanced situation is worth striving for.

Bibliography

Bergqvist, T. and Lithner, J. (2005). Simulating creative reasoning in mathematics teaching. *Department of Mathematics and Mathematical statistics, Umeå University, Research Reports in Mathematics Education*, no. 2.

Bergqvist, T., Lithner, J., and Sumpter, L. (2003). Reasoning characteristics in upper secondary school students' task solving. *Department of Mathematics and Mathematical statistics, Umeå University, Research Reports in Mathematics Education*, no. 1.

Boesen, J. (2006). *Assessing Mathematical Creativity—Comparing national and teacher-made tests, explaining differences and examining impact.* PhD thesis, Umeå Universitet.

Boesen, J., Lithner, J., and Palm, T. (2005). The relation between test task requirements and the reasoning used by students. *Department of Mathematics and Mathematical statistics, Umeå University, Research Reports in Mathematics Education*, no. 4.

Brosseau, G. (1997). *Theory of didactical situations in mathematics.* Dordrecht: Kluwer.

Bu, S. and Schachermayer, W. (1992). Approximation of Jensen measures by image measures under holomorphic functions and applications. *Transactions of the American Mathematical Society*, 331:585–608.

Burns, M. (1994). Arithmetic: The last holdout. *Phi Delta Kappan*, 75:471–476.

Duval, J. and Sibony, N. (1995). Polynomial convexity, rational convexity, and currents. *Duke Mathematical Journal*, 79:487–513.

Ebby, C. B. (2005). The powers and pitfalls of algorithmic knowledge: a case study. *The Journal of Mathematical Behaviour*, 24(1):73–87.

Flemming, M. and Chambers, B. (1983). Teacher-made tests: Windows on the classroom. In Hathaway, W., editor, *Testing in the schools: New directions for testing and measurement.* Jossey-Bass, San Francisco.

Gamelin, T. W. (1969). *Uniform Algebras.* Prentice-Hall, Inc., Englewood Cliffs, N. J.

Haylock, D. (1997). Recognising mathematical creativity in schoolchildren. *Zentralblatt fuer Didaktik der Mathematik*, 29(3):68–74.

Hegarty, M., Mayer, R., and Monk, C. (1995). Comprehension of arithmetic word problems: A comparison of successful and unsuccessful problem solvers. *Journal of Educational Psychology*, 87(1):18–32.

Hiebert, J. (2003). What research says about the NCTM standards. In Kilpatrick, J., Martin, G., and Schifter, D., editors, *A Research Companion to Principles and Standards for School Mathematics*, pages 5–26. Reston, Va.: National Council of Teachers of Mathematics.

Hiebert, J. and Carpenter, T. (1992). Learning and teaching with understanding. In Grouws, D., editor, *Handbook for Research on Mathematics Teaching and Learning*, pages 65–97. New York: Macmillan.

Johansson, B. and Emanuelsson, J. (1997). *Utvärdering i naturkunskap och matematik: lärare i grundskolan berättar: slutrapport från UNO-LÄR-projektet (In Swedish).* The Swedish National Agency for Education: Liber distribution, Stockholm.

Kamii, C. and Dominick, A. (1997). To teach or not to teach algorithms. *Journal of Mathematical Behavior*, 16(1):51–61.

Kane, M., Crooks, T., and Cohen, A. (1999). Validating measures of performance. *Eductional Measurement: Issues and Practice*, 18(2):5–17.

Leinwand, S. (1994). It's time to abandon computational algorithms. *Education Week*, 9:36.

Lithner, J. (2000a). Mathematical reasoning and familiar procedures. *International Journal of Mathematical Education in Science and Technology*, 31:83–95.

Lithner, J. (2000b). Mathematical reasoning in task solving. *Educational Studies in Mathematics*, 41:165–190.

Lithner, J. (2003). Students' mathematical reasoning in university textbook exercises. *Educational studies in mathematics*, 52:29–55.

Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *Journal of Mathematical Behaviour*, 23:405–427.

Lithner, J. (2006). A framework for analysing creative and imitative mathematical reasoning. *Department of Mathematics and Mathematical statistics, Umeå University, Research Reports in Mathematics Education*, no. 2.

McNeal, B. (1995). Learning not to think in a textbook-based mathematics class. *Journal of Mathematical Behavior*, 14:18–32.

Mullis, I., Martin, M., Smith, T., Garden, R., Gregory, K., Gonzalez, E., Chrostowski, S., and O'Connor, K. (2003). *TIMSS Assessment Frameworks and Specifications 2003 (2nd Edition).* Chestnut Hill, MA: Boston College.

NCTM (1989). *Curriculum and evaluation standards for school mathematics.* National Council of Teachers of Mathematics, Reston, Va.

NCTM (2000). *Principles and Standards for School Mathematics.* National Council of Teachers of Mathematics, Reston, Va.

Niss, M. and Jensen, T., editors (2002). *Kompetencer og matematiklæring: Ideer og inspiration til udvikling af matematikundervisning i Danmark.* Uddannelsestyrelsens temahaefteserie nr. 18, Undervisningsministeriet.

Ocken, S. (2001). Algorithms, algebra, and access. http://www.nychold.com, september 2006.

Oka, K. (1937). Sur les fonctions analytiques de plusierus variables. ii. domaines d'holomorphie. *J. Sci. Hiroshima Univ. Ser. A*, 7:115–130.

Palm, T. (2002). *The Realism of Mathematical School Tasks—Features and Consequences.* PhD thesis, Umeå university.

Palm, T., Boesen, J., and Lithner, J. (2005). The requirements of mathematical reasoning in upper secondary level assessments. *Department of Mathematics and Mathematical statistics, Umeå University, Research Reports in Mathematics Education*, no. 5.

Pesek, D. and Kirshner, D. (2000). Interference of instrumental instruction in subsequent relational learning. *Journal for Research in Mathematics Education*, 31(5):524–540.

Pólya, G. (1945). *How to Solve It.* Princeton NJ: Princeton University Press.

Pólya, G. (1954). *Mathematics and Plausible Reasoning, Vol. I and II.* Princeton NJ: Princeton University Press.

Rossi, H. (1960). The local maximum modulus principle. *Annals of Mathematics*, 72:1–11.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving.* Orlando, FL: Academic Press.

Schoenfeld, A. H. (1991). On mathematics as sense-making: An informal attack on the unfortunate divorce of formal and informal mathematics. In Voss, J., Perkins, D., and Segal, J., editors, *Informal Reasoning and Education*, pages 311–344. Lawrence Erlbaum Associates, Hillsdale, NJ.

Schoenfeld, A. H. (2004). The math wars. *Educational Policy*, 18(1):253–286.

Senk, S. L., Beckmann, C. E., and Thompson, D. R. (1997). Assessment and grading in high school mathematics classrooms. *Journal for Research in Mathematics Education*, 28(2):187–215.

Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflection on processes and objects as different sides of the same coin. *Educational studies in mathematics*, 22:1–36.

Shepard, L. A. (2000). The role of classroom assessment in teaching and learning. *CSE Technical Report 517.* www.cse.ucla.edu/CRESST/pages/reports.htm.

Silver, E. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *Zentralblatt fuer Didaktik der Mathematik*, 29(3):75–80.

Skemp, R. (1978). Relational understanding and instrumental understanding. *Arithmetic Teacher*, 26(3):9–15.

Skinner, B. F. (1954). The science of learning and the art of teaching. *Harvard Educational Review*, 24:86–97.

Sriraman, B. (2004). The characteristics of mathematical creativity. *The Math-*

ematics Educator, 14(1):19–34.

Stolzenberg, G. (1963a). A hull with no analytic structure. *Journal of Mathematics and Mechanics*, 12:103–111.

Stolzenberg, G. (1963b). Polynomially and rationally convex sets. *Acta Mathematica*, 109:259–289.

Swedish National Agency of Education (2004). TIMSS 2003: Swedish pupils knowledge in mathematics and science in school year 8 in a national and international perspective (in swedish). *Swedish National Agency of Education reports*, 255.

Tall, D. (1996). Functions and calculus. In Bishop, A., Clements, K., Keitel, C., Kilpatrick, J., and Laborde, C., editors, *International Handbook of Mathematics Education*, pages 289–325. Kluwer, Dordrecht.

Verschaffel, L., Greer, B., and DeCorte, E. (2000). *Making sense of word problems*. Swets and Zeitlinger Publishers, Lisse, The Netherlands.

Vinner, S. (1997). The pseudo-conceptual and the pseudo-analytical thought processes in mathematics. *Educational Studies in Mathematics*, 34:97–129.

Wermer, J. (1982). Polynomially convex hulls and analyticity. *Arkiv för Matematik*, 20:129–135.

Wu, H. (1999). Basic skills versus conceptual understanding—a bogus dichotomy in mathematics education. *American Educator, American Federation of Teachers*.