PROSPECTIVE MATHEMATICS TEACHERS' INFORMAL AND FORMAL REASONING ABOUT THE CONCEPTS OF DERIVATIVE AND DIFFERENTIABILITY

Antti Viholainen
Department of Mathematics and Statistics
University of Jyväskylä
Finland

Email: anvihola@maths.jyu.fi

Abstract

Antti Viholainen

Prospective mathematics teachers' informal and formal reasoning about the concepts of derivative and differentiability

Department of mathematics and statistics
University of Jyväskylä
Finland

The scientific nature of mathematics is extremely exact, detailed and abstract. This is important, above all, in order to preserve its unambiguousness. However, holistic and concrete interpretations are very important in creative mathematical thinking. They are as well important in mathematical understanding, because mathematical knowledge presented in a formal form is usually not very explanatory and thus does not underpin understanding. Several classifications regarding mathematical thinking and mathematical understanding presented in the literature are based on this dichotomy. The classification of mathematical reasoning into informal and formal types used in this work is also based on it. The informal reasoning is based on visual or physical interpretations of mathematical concepts, as against the formal reasoning means exact reasoning based on axioms, definitions and previously proven theorems. The formal reasoning concerning a certain concept is usually based on the definition of the concept at issue.

This study examined informal and formal understanding of the concepts of derivative and differentiability and the use of informal and formal reasoning in problem solving situation where these concepts were needed. The subjects of the study were mathematics education students in the middle or in the final phase of their studies. The data are based on a written test given at six Finnish universities and on some oral interviews of the participants of the test.

The study showed that connecting informal and formal reasoning was often difficult for the students. In particular, the students seemed to have a tendency to avoid using the definition of the derivative in problem solving situations. This considerably hindered problem solving processes and in some cases led to erroneous conclusions. Inability to use the definition does not explain this tendency, as several students were able to use the definition when they were asked to do so. Corresponding tendencies to avoid using the definitions have also been observed in several previous studies. On the other hand, there are several studies showing that in certain cases, students' reasoning may be too heavily based on the definition. Both of these results indicate that crossing the line between informal and formal representation systems is difficult for many students. This inability often restricts students' reasoning. In order to improve this, the teaching of mathematics should support the development of the coherence of students' knowledge structure. Among other things, it should strengthen the understanding of connections between informal and formal representations.

Key words: argumentation, coherence of the concept image, concept image, definition, derivative, informal and formal, informal interpretation, mathematical reasoning, modelling, representation, visualization.

Tiivistelmä

Antti Viholainen

Tulevien matematiikan opettajien informaali ja formaali päättely derivaatan ja derivoituvuuden käsitteiden yhteydessä

Matematiikan ja tilastotieteen laitos
Jyväskylän yliopisto
Suomi

Matematiikka on tieteelliseltä luonteeltaan periaatteessa äärimmäisen täsmällistä, yksityiskohtaista ja abstraktia. Tämä on tärkeää ennen kaikkea yksiselitteisyyden säilyttämiseksi. Kuitenkin luovassa matemaattisessa ajattelussa kokonaisvaltaiset ja konkreettiset tulkinnat ovat erittäin tärkeitä. Ne ovat tärkeitä myös matematiikan ymmärtämisen kannalta, sillä matemaattinen tieto formaalissa muodossa esitettynä ei yleensä ole kovin selittävää eikä siten ymmärtämistä tukevaa. Useat kirjallisuudessa esitetyt luokittelut matemaattiselle ajattelulle ja matemaattiselle ymmärtämiselle perustuvat tähän dikotomiaan. Siihen perustuu myös tässä työssä käytetty jaottelu informaaliin ja formaaliin päättelyyn. Informaalilla päättelyllä tarkoitetaan matemaattisten käsitteiden visuaalisiin tai fysikaalisiin tulkintoihin perustuvaa päättelyä, kun taas formaali päättely tarkoittaa aksioomiin, määritelmiin ja aikaisemmin todistettuihin teoreemoihin perustuvaa aukotonta päättelyä. Tiettyä käsitettä koskeva formaali päättely perustuu yleensä käsitteen määritelmään.

Tämän tutkimuksen tavoitteena oli selvittää matematiikan aineenopettajaopiskelijoiden derivaatan ja derivoituvuuden käsitteiden informaalia ja formaalia ymmärtämistä ja informaalin ja formaalin päättelyn käyttöä näihin käsitteisiin liittyvissä ongelmanratkaisutilanteissa. Tutkimukseen osallistuneet opiskelijat olivat opintojensa keski- tai loppuvaiheessa. Tutkimuksen aineisto perustuu yhteensä kuudessa suomalaisessa yliopistossa suoritettuun kirjalliseen testiin ja joidenkin testin osanottajien suullisiin haastatteluihin.

Useissa kohdin tutkimuksessa tuli esiin, että informaalin ja formaalin päättelyn yhdistäminen oli opiskelijoillle vaikeaa. Erityisesti derivaatan kohdalla opiskelijoilla näytti olevan taipumus välttää määritelmän käyttöä ongelmanratkaisutilanteissa. Tämä vaikeutti ongelmanratkaisuprosessia huomattavasti ja joissakin tapauksissa johti virheellisiin johtopäätöksiin. Heikot taidot käyttää määritelmää eivät selitä tätä taipumusta, sillä useat opiskelijat kuitenkin osasivat käyttää määritelmää tehtävissä, joissa määritelmän käyttöä nimenomaisesti pyydettiin. Määritelmän välttelytaipumuksia on havaittu myös monissa aikaisemmissa tutkimuksissa, mutta toisaalta joissakin tutkimuksissa taas ongelmaksi on todettu liika määritelmäsidonnaisuus. Molemmanlaiset tulokset kertovat kuitenkin samasta ilmiöstä: Informaalien ja formaalien representaatiojärjestelmien välisten rajapintojen ylittäminen on opiskelijoille vaikeaa ja tästä johtuen heidän päättelynsä on rajoittunutta. Tilanteen parantamiseksi matematiikan opetuksen tulisi olla informaalien ja formaalien representaatioiden välisten yhteyksien ymmärtämistä ja muutoinkin tietorakenteen koherenttisuuden kehittymistä tukevaa.

Asiasanat: argumentointi, derivaatta, informaali ja formaali, informaali tulkinta, käsitekuva, käsitekuvan koherenssi, mallintaminen, matemaattinen päättely, määritelmä, representaatio, visualisointi.

Author: Antti Viholainen
Department of Mathematics and Statistics
University of Jyväskylä
Finland

Email:anvihola@maths.jyu.fi

Supervisors: Professor Pekka Koskela
Department of Mathematics and Statistics
University of Jyväskylä
Finland

Lecturer Petri Juutinen
Department of Mathematics and Statistics
University of Jyväskylä
Finland

Senior Researcher, Docent Pekka Kupari
Institute for Educational Research
University of Jyväskylä
Finland

Reviewers: Professor Elena Nardi
School of Education and Lifelong Learning
University of East Anglia
Great Britain

Senior Researcher, Docent Kaarina Merenluoto
Department of Teacher Education
University of Turku
Finland

Opponent: Lecturer, Docent Juha Oikkonen
Department of Mathematics and Statistics
University of Helsinki
Finland

Acknowledgement

In Autumn 2002, when I was just completing my master's degree, I seriously planned to begin a career as a teacher of mathematics. I was very interested in questions on the learning of mathematics –a subject which students very often find difficult. Then Professor Pekka Koskela told me about the possibility to do research about mathematics education students' understanding of mathematics. However, I did not know much about the conventions of educational research or about the present trends in mathematics education research, and, in addition, nobody at our department was working in this area full time. I therefore thought this challenge was quite impossible for me.

Even though the situation seemed hopeless at first, there have been kind and patient people around me who have crucially helped me to get this doctoral thesis project to move ahead. First, I would like to thank my supervisors, Pekka Koskela, Pekka Kupari and Petri Juutinen. I have had dozens of long discussions during the last six years both together with them all and privately with each of them. Second, I want to thank my long-term colleague and friend Markus Hähkiöniemi, who at the same time with me with his heart and soul also studied the learning of derivative and prepared his doctoral thesis. The thorough discussions with these people have been very useful and inspiring. In fact, I think they have often been necessary for the progress of my work.

Several institutions have had an important impact on my work. First, of course, I would like to thank the Department of Mathematics and Statistics for all the financial, technical and spiritual support for the study. Another important institution has been The Finnish Graduate School of Mathematics, Physics and Chemistry Education, which arranged semiannual conferences, where it was possible to meet national and international researchers, learn research methodology and get feedback of your own research. Especially, I would like to thank Professor Erkki Pehkonen, the leader of the graduate school. He also gave me many useful comments and advice. As well, in 2004, The Nordic Graduate School in Mathematics Education sponsored in 2004 my participating to The Second YERME Summer School in Poděbrady Czech Republic, and, in 2005, arranged The First Nordic Summer School in Mathematics Education in Jyväskylä. Special thanks go to Professor Barbro Grevholm, the director of the Nordic Graduate School. The summer school in Poděbrady was my first real international research contact. In addition, in 2006 I had the possibility to take part also in The Third YERME Summer School in Jyväskylä. In addition to these summer schools, I have taken part in three high-level conferences: CERME 4, PME 30 and CERME 5. These all together have helped me, at least a little, to understand what is going on in the area of mathematics education research internationally. Professor Guershon Harel and Professor Tommy Dreyfus, my group leaders in the summer schools, deserve my special thanks for very competent personal comments.

I would like to thank the Magnus Ehrnrooth Foundation for the financial support for participating in PME 30 and CERME 5 conferences. Furthermore, I am grateful to this foundation also for the financial support for the data collection in Autumn 2004 and in Spring 2005. Thanks are also due to the following departments and contact persons (in brackets), who made the data collection possible: Department of Teacher Education opf the University of Jyväskylä (Sinikka Kaartinen), Department of Educational Sciences and Teacher Education of the University of Oulu (Lasse Jalonen), Department of Teacher Education of the University of Turku (Aija Ahtineva), Department of Teacher Education of the University of Tampere (Päivi Portaankorva-Koivisto), Department of Mathematics and Statistics of the University of Helsinki (Ritva Hurri-Syrjänen), Department of Mathematics [1] of the University of Joensuu (Lasse Eronen) and Department of Mathematics and Mathematical Statistics of the University of Umeå (Nina Rudälv). All the students who in a devoted way took part in my study deserve my gratitude. In addition, I would like to thank the Emil Aaltonen Foundation, the Ellen and Artturi Nyyssönen Foundation and the University of Jyväskylä for financial support.

In addition to personal contacts, many high-level scientific papers have had an important impact on my thinking. For example, David Tall's writings about concept image and concept definition and about the three worlds of mathematics, as well as Manya Raman's work about private and public arguments and about the key idea helped me in a crucial way to direct the theoretical framework of the study.

Thanks go also to Professor Jouni Viiri for personal discussions and for the interesting seminars which he started to lead here in Jyväskylä when I was finishing my study. Thanks are also due to all the members of this seminar group.

Moreover, I would like to thank all the reviewers of my journal and conference papers for well-thought-out comments. Thanks are also due to Doctor Elena Nardi and Senior Researcher Kaarina Merenluoto who worked as reviewers of the whole study. They gave also comments which will be very important for me if I sometimes continue the research in this area. I am grateful also to Tuula Blåfield for her enormous work in checking the language of all my papers and the final thesis.

Last, I would like to thank my parents and all my friends for their support. Especially, I would like to thank my fiancée Anne for all her love.

On the 1st of April this year, I did two things: First, I finished my doctoral thesis and sent it to the review process. Second, I put an engagement ring on Anne's finger. Neither of these was an April fool's trick.

Jyväskylä, the end of July 2008,

Antti Viholainen

[1]Nowadays, Department of Physics and Mathematics

Contents

List of the included articles

A. Viholainen, A.: 2006a, 'Relationships between informal and formal reasoning in the subject of derivative', in Bosch, M. (ch. ed.): *Proceedings of The Fourth Congress of the European Society for Research in Mathematics Education*, Sant Feliu de Guíxols, Spain, 17.-21. February 2005, FUNDEMI IQS, Universitat Ramon Llull, pp. 1811-1820.

B. Viholainen, A.: 2006b, 'Why is a discontinuous function differentiable?', in Novotná, J., Moraová, H., Krátká, M. and Stehliková, N. (Eds.): *Proceedings of The 30th Conference of the International Group for the Psychology of Mathematics Education*, Charles University, Faculty of Education, Prague, Czech Rebublic, Vol 5, pp. 329-336.

C. Viholainen, A.: 2007, 'Students' choices between informal and formal reasoning in a task concerning differentiability', in Pitta-Pantazi, D. and Philippou, G. (Eds.): *Proceedings of The Fifth Congress of the European Society for Research in Mathematics Education*, Department of Education, University of Cyprus, pp. 2369-2378.

D. Viholainen, A.: 2008a, 'Incoherence of a concept image and erroneous conclusions in the case of differentiability', *The Montana Mathematics Enthusiast*, 5(2&3), 231-248.

E. Viholainen, A.: 2008b, 'Finnish mathematics teacher students' informal and formal arguing skills in the case of derivative', *Nordic Studies in Mathematics Education*, 13(2), 71-92.

1 Introduction

What is derivative? According to its formal definition, derivative is the limit of the difference quotient. In the visual sense, derivative describes the steepness of the graph in question. Strictly speaking, it is the slope of the tangent line. Furthermore, derivative can be said to be an instantaneous rate of change. At least in some sense, all these interpretations of derivative usually come out in the teaching of mathematics wherever basic analysis is taught. Indeed, emphases may differ between institutions: For example, Bingolbali (2005) and Bingolbali et al. (2006) found that the "rate of change" aspect was emphasized in mechanical engineering education, whereas the tangent aspect was more stressed in mathematics courses at university. In addition, the concept of derivative is often associated to differentiation rules and other calculation algorithms, and, in fact, in teaching these are usually emphasized more than any of the above-mentioned interpretations. But do students learn to use different interpretations of the concept of derivative correctly and effectively in their reasoning? Are they able to construct a coherent view of this concept or do the different interpretations stay as separate pieces of knowledge?

Very often theoretical mathematical knowledge does not appear to have any connection to the empirical world in which we live. The definitions of the concepts and theorems with their proofs may be fully abstract, at least in the form in which they are presented in textbooks, in lectures or in research papers. It may be difficult, even for an experienced mathematician, to recognise how these definitions and theorems could be connected to any phenomena in real life. In addition, in mathematics reasoning and the use of language is required to be more rigorous than in everyday life. Abstractness and rigour are thus unavoidable features of mathematics, but they present a notable challenge for mathematics education, because they often make students feel that mathematics is odd and difficult.

However, the historical origin of most mathematics is in real life. Several mathematical concepts have been created by defining some physical phenomenon generally and rigorously, in other words, by interpreting a physical phenomenon mathematically. For example, the concept of derivative can be considered as a mathematical interpretation of the instantaneous rate of change. Furthermore, mathematical concepts often have interpretations which are based on visualization, such as the interpretation of derivative as the slope of the tangent line. Physical and visual interpretations are important in inventing new ideas and in applying mathematics. They help to create a connection between mathematics and the empirical world, and thus they make holistic views and the effective use of intuition, associations and mental images possible in mathematical reasoning.

This thesis concerns the dual features of mathematics presented above. The study deals with Finnish university students' mastery of the *informal* and *formal*[2] sides of mathematics in the case of the concept of derivative. The majority of the students in question were prospective teachers at the middle

[2] The terms "informal" and "formal" are defined more precisely in Section 2.7.

or at the final phase of their studies. We will study how the students used informal and formal reasoning when they solved mathematical problems. In addition, we will analyse some erroneous conclusions made by the students and study the role of the *coherence of the concept image* (see Section 3.3) in reasoning.

2 Informal and formal features of mathematics

In this section some views about the informal and formal sides of mathematics are reviewed. A short historical review about the development of mathematics as a discipline is then presented. After that, several ways to classify different sides of mathematics, mathematical thinking and mathematical arguments presented in the literature are reviewed. Related to that, the role of models in working with mathematics is considered. Especially, the role of the visual models and visualization is emphasized. Finally, I present my own classification for mathematical arguments which I assume to be applicable in this study. This classification is illustrated by some examples concerning the concept of derivative. In addition, I express some ontological assumptions about mathematics on which this classification as well this study as a whole is based.

2.1 The historical evolution of formal mathematics

Almost throughout its history, mathematics has been considered as a deductive science. Nevertheless, the building of the axiomatic system of mathematics has proved to be difficult. Circa 300 BC Euclid made a very notable attempt. In his famous book *Elements* he presented definitions, axioms, postulates and theorems with their proofs. For centuries mathematicians considered his work to be ideal. Before Euclid, Aristotle had emphasized that in mathematics it was necessary to start with simple, unquestionable truths and carefully prove all other truths from them. Now Euclid seemed to have achieved this goal.

However, by the end of the 19th century, many deficiencies with respect to rigour had been noticed in Euclid's work. The problem was that the basis of mathematics had been bound to the natural reality, and so almost all mathematics rested on an empirical and pragmatic basis (Kline, 1972). Due to that, mathematical proofs had logical gaps which could not be fixed. David Hilbert (1862-1943) realized that the axiomatic system of mathematics had to be separated from empirical reality. He built a system which includes some undefined concepts which do not have a definition. All the other concepts are defined by using previously defined concepts or listed undefined concepts. All concepts are thus absolutely determined independently from their empirical origin. The axioms are, as well, separated from the empirical reality in Hilbert's system: When in Euclid's system the axioms are self-evident truths which nobody can deny, in Hilbert's system axioms are arbitrary claims which are assumed to be true in the system. They determine the relationships between the undefined concepts. All other relationships have to be verifiable by using only previously proven theorems, axioms, definitions and the rules of proving. Definitions, axioms, theorems and their proofs are presented by using logical language. Therefore, Hilbert's axiomatic system of mathematics consists of seven parts: logical language, rules of proof, undefined concepts, axioms, definitions, theorems and proofs of the theorems (Sibley, 1998).

Before the middle of the 19th century, mathematical concepts usually had physical interpretations and mathematics was understood as a tool to model and idealize realities in nature. But after that, the number of concepts without any immediate physical interpretation increased. This changed the view about mathematics: Mathematics had to be understood as an arbitrary creation of a human, not as a body of truths about nature (Kline, 1972).

2.2 Formal, algorithmic and intuitive components of a mathematical activity

According to Fischbein (1994), mathematics should be considered from two points of view: On one hand, mathematics is a formal, deductive and rigorous body of knowledge, and, on the other hand, mathematics can be regarded as an activity of a human. Fischbein divides mathematics as a human activity into *formal*, *algorithmic* and *intuitive* components. The formal component of mathematical activity refers to activities with the formal axiomatic system. By the algorithmic component Fischbein means capabilities to use mathematical solving procedures. The understanding of the formal side of mathematics is not enough to acquire these skills, but practical training is needed. However, a deep and comprehensive understanding requires that also the formal justifications of solving procedures are understood. In this way, the formal and algorithmic components are connected, and this connection is a basis for a symbiosis between meanings and skills, which, furthermore, is a basic condition for efficient mathematical reasoning.

Fischbein defines *an intuitive cognition* as the kind of cognition that is accepted directly without the feeling that any kind of justification is required. The intuitive component of mathematical activity is based on intuitive cognitions. Sometimes intuitive cognitions may be in accordance with the formal results, but sometimes they may contradict them, and so they may become epistemological obstacles in learning, solving or invention processes. On the other hand, they may play a facilitating role in these processes. Fischbein admits that the relationship between the formal and the intuitive aspects of mathematical reasoning in learning, understanding and solving processes is very complex and usually not easily identified and understood.

2.3 Three worlds of mathematics

Tall's theory about three worlds of mathematics describes the cognitive development of mathematical knowledge. The three worlds are *conceptual-embodied world*, *proceptual-symbolic world* and *formal-axiomatic world* (Tall, 2005). According to Tall (1995), the learning of mathematics is on one hand based on *perceptions of*, and, on the other hand, *actions on*, objects in the environment. The perceptions of representations of mathematical objects form a basis for the conceptual-embodied world, and the actions on mathematical symbols are a basis for the proceptual-symbolic world. The formal-axiomatic world refers to the formal axiomatic system of mathematics as described above.

In the perception-based learning process, the building of knowledge proceeds from primitive perception to more refined conceptions. At first, concrete objects or their visuo-spatial representations are perceived, then properties of the objects are analysed and described verbally, and, finally, the objects are classified, which leads to collections and hierarchies of objects. At the early stages of this development, the focus of attention is on the specific details of perceptions, but later it moves to underlying regularities. Also conceptions about mathematical objects change and become more abstract and more ideal: For example, at the early stages the concepts of point and line are understood as physical objects, but later they are understood as ideal concepts so that the point does not have size and that the line does not have thickness.

The proceptual-symbolic world is based on Gray's and Tall's (1994, 2001) theory about symbols acting dually as processes and concepts. The concepts can be seen, in one sense, as products of the corresponding processes. For example, the symbol "3/4" stands for both the process of dividing 3 by 4 or the concept of the fraction three fourths. The character string "$3x + 2$" stands both for the process "add three times x and two" or the product of this process. It may also stand for the binomial "$3x + 2$" or the corresponding function. Gray and Tall use their own word "procept" (*pro*cess + con*cept* = *procept*) in order to refer to an amalgam of three components: a process, a product of this process (a concept), and a symbol which represents either the process or the concept. By *proceptual thinking* Gray and Tall mean the ability to consider symbols flexibly both as processes and concepts. This ability means that an individual is able to compress a process to an entity, but, he/she is also able, when needed, to decompose concepts as processes that produce the concepts.

An action-based learning process begins by making some actions on the objects. At first, a sequence of actions, a *procedure*, is performed by using a step-by-step algorithm. After several repetitions the procedure is automatised, and an individual is able to see it as an entity so that he/she can consider it without referring to the single steps. Then the process is encapsulated as a mental object. The APOS theory (Asiala et al., 1996; Dubinsky and McDonald, 2001) and Sfard's (1991) theory about the *reification* describe the cognitive development of processes into objects.

According to Tall, the perception-based learning in the conceptual-embodied world and the action-based learning in the proceptual-symbolic world develop parallel. The overall features in this development are the *compression* of knowledge into *thinkable concepts* that can be held in the focus of attention and the creating of *connections* between these concepts. After a reflection of the *embodied concepts* (See Tall, 2005) and procepts, an individual can understand the fundamental ideas of the formal axiomatic system. The *generalization* and the *abstraction* are important factors in this process. The learning in the conceptual-embodied and proceptual-symbolic worlds happens by exploring qualities of existing objects by perceptions and actions, and by compressing appropriately the information achieved in this way. Instead, in the formal-axiomatic world, the elements of the formal-axiomatic system are starting points for all the reasoning. The arguments in each of the three worlds are based on different foundations: In the conceptual-embodied world something is true if it is seen, -or if it can be imagined to be seen- to be true, in the proceptual-symbolic world symbolic manipulations are warrants for truths, and in the formal-axiomatic world something is true if it is either assumed by an axiom or a definition or if it is deductively proved from them (Tall, 2003b; 2004).

In university mathematics, learning may begin from the formal-axiomatic world. This, for example, happens when a formal definition is given to a student without any introduction to the concept. If needed, the student has to conclude by himself/herself the meaning of the concept on the basis of the definition. Creating visual or physical interpretations for the definition may be an effective method to do this (See Section 2.7). However, in some situations the formal definition can be used in the raw in the formal context. Pinto (1998) and Pinto and Tall (2001) noticed that, in situations requiring the use of formal definitions and a formal theory, some students tended to create concrete interpretations, whereas others attempted to base their reasoning only on the formalism.

Hähkiöniemi (2006b) has applied the theory about the three worlds of mathematics in the case of the concept of derivative. He has developed a hypothetical learning path to derivative. In this model, the learning in the conceptual-embodied world means perceiving the rate of change by using different representations concerning the tangent, local straightness and increase, steepness and horizontalness of a function. Furthermore, the rate of change can be explored by moving a hand along a graph or by placing a pencil as a tangent. On the other hand, students may get acquainted with the idea of derivative by calculating average rates of changes over different intervals. This corresponds to the learning in the proceptual-symbolic world. According to Hähkiöniemi, these operations introduce students to the concepts of instantaneous velocity and to the problems in determining it. This creates a natural need for the limiting process of the difference quotient and thus for the formal definition of derivative. Hähkiöniemi also shows that derivative can be considered as an object already at the early stages of the learning process (Hähkiöniemi, 2006a). This result is neither in accordance with the APOS theory nor with Sfard's reification theory, which claim that in learning the process phase should come before the object phase.

Tall's theory about the cognitive development of mathematical knowledge is in many senses comparable with the historical development of mathematics as a discipline: Also the history of mathematics has begun from perceptions and elementary actions, later the amount of knowledge has increased, and a need to compress, generalize, abstract and finally formalize the existing knowledge has emerged.

2.4 The role of models in mathematics

Both Fishbein's and Tall's theories bring out several views about mathematics: Mathematics is, on one hand, an abstract formal system, in which only fully exact deductive arguments are acceptable warrants for the truth. On the other hand, mastery of different kind of procedural skills is essential in working with mathematics. Moreover, it is important to comprehensively understand the meaning of mathematical concepts, theorems, and so on.[3] Different kind of representations are effective tools contributing to this understanding. For example, Goldin (1998) presents a wide spectrum of representations that are used in mathematical thinking. Also *modelling* is based on the use of representations. Modelling often emerges from the need to get a comprehensive view about a situation. Even though mathematical knowledge has in principle been separated from the empirical reality, from a cognitive point of view it is difficult to treat objects which do not have any empirical interpretation. If these interpretations are not given, they have to be created, and models are useful tools for this. The *concrete models* are useful in creating appropriate *mental models* about the situation. In fact, they act as *external representations* (see Section 3.2) of the mental models.

The direction of modelling may be either from abstract to concrete or vice versa: In learning mathematics and in inventing new mathematical ideas, an abstract mathematical occasion can be modelled by using concrete representations. Instead, when mathematics is applied in other disciplines, a nonmathematical occasion has to be modelled mathematically. Both of these processes require a thorough understanding of the connections between the original occasion and its model.

Fishbein (2001) defines the term "model" in the following way:

"Considering two systems, A and B, B is defined as a model of A if it is possible to translate properties of A in terms of B so as to produce consistent descriptions of A in terms of B, or to solve problems - originally formulated in terms of A - by resorting to a translation in terms of B." (p. 312)

Fishbein reminds that a model is partially different from the original system, and thus its relevance is limited. A model may also include properties which are not relevant for the original system. Thus, a reasoning based only on a single model may lead to erroneous conclusions with respect to the original system.

In metamathematics, a model of an axiomatic system is defined as "...a set of objects together with interpretations of all the undefined terms of the axiomatic system such that all the axioms are true in the set using the interpretations" (Sibley, 1998; 29). It is possible to prove that all theorems which are true in an axiomatic system are also true in a model of the system, but all statements which are true in the model are not necessarily true in the system. So it is not possible to find out the truths of the system on the basis of models. In practice this means that argumentation based only on models cannot

[3] In this connection the term "meaning" refers to a personal way of understanding. I assume that mathematical knowledge does not have an objective meaning with respect to the empirical world (see Section 2.9).

be substituted for formal proofs. However, different models are very important in creative mathematical thinking, and in practice their usage is based on the fact that claims that can be reasoned on the basis of a model are very often verifiable also in the axiomatic system.

According to Fishbein, models may be based on the use of analogies, prototypes or diagrams. They may also be abstract, if the direction of the modelling is from concrete to abstract: For example, a formula describing a physical event is an abstract model for a concrete occasion. Usually models are explicit and purposefully created, but, according to Fishbein, sometimes the models may be tacit. The tacit models may influence reasoning process without the individual being aware of their origin and their effect. Therefore, they may cause conflicts. For example, an individual may consider geometrical objects only as figural models even if he/she is aware of the abstract nature of them. According to Fishbein, people often tend to consider a line as a straight queue of small spots which have equal size. With this model in mind, it is difficult to accept that line segments of different lengths, considered as abstract geometrical objects, are equal in the number of points. Also mathematicians may be misled by tacit models. For example, up to the 19th century mathematicians believed that every continuous function is differentiable at least somewhere. This conclusion was evident, because the function was interpreted as an ink trace on a sheet of paper, which represented the graph of the function, and it was impossible to draw a trace which is continuous but does not have a tangent at any point. However, the model was inadequate: Riemann presented an example of a function which was continuous but nowhere differentiable.

Presmeg (2006a) defines that a *sign* is an *interpreted relationship* between a *sign vehicle* and an object that the sign vehicle represents or stands for in some way.[4] On the basis of this definition, modelling can be interpreted to be based on the use of *signs*. According to Presmeg, all mathematical objects are more general than their particular instantiations in sign vehicles: For example, a thin segment of a straight line drawn on a sheet of paper (a sign vehicle) may represent the ideal concept of a straight line (an object). The limitations of the sign vehicles have to be known when treating mathematical objects by their sign vehicles: One has to be aware which qualities of the sign vehicle are characteristic of the corresponding object and which are characteristic only of the used vehicle. If these restrictions are not known or taken into account, conflicts like above may happen. On the model level, an individual has to be aware of the relationships between the model and the original occasion that it represents.

In the following section we study the roles of visual models and visualization in mathematical thinking in more detail.

[4]The terms like "sign" and "representation" have been widely used in the discipline of mathematics education, but their meanings have been ambiguous to some extent.

2.5 Usage of visual models

2.5.1 What is visualization?

According to Presmeg (2006a), a mental imagery may occur in various modalities, such as sight, hearing, smell, taste or touch. Goldin's (1998) classification for representations brings out a similar variety of modalities. However, Presmeg considers the visual modality the most prevalent one in mathematical thinking. The term "visual model" may mean a concrete picture drawn on a paper, but it can also refer to mental images. In fact, concrete pictures can be considered as external representations of the mental images. Gutiérrez (1996) defines visualization in mathematics "as the kind of reasoning activity based on the use of visual or spatial elements, either mental or physical, performed to solve problems or prove properties" (p.9). He presents the following main elements for visualization:

- *Mental images* are cognitive representations of mathematical concepts or properties by means of visual or spatial elements. In this connection, the term "mental image" means only mental images depicting visual or spatial information.[5] According to Presmeg's (1986b) classification, observed (visual) mental images may be concrete pictorial images, pattern images representing abstract mathematical relationships in a visual way, images of written formulae or kinaesthetic images based on physical movements. In addition, mental images can be static or dynamic.

- *External representations* are verbal or graphical representations of concepts or properties.

- *Processes of visualization* are mental or physical actions in which mental images are involved. The action may mean either creating of mental images by transforming non-figural data into visual form or interpretation of existing mental images to generate information from them. The latter action may mean "observation and analysis of mental images, transformation of mental images into other mental images" or "transformation of mental images into other kinds of information" (p.10).

- *Abilities needed in visualization:* For example, the ability to identify a specific figure by isolating it out of a background, ability to recognise the independence of some properties from size, position, orientation etc., the ability to produce dynamic mental images and the ability to identify similarities and differences between several objects, pictures and mental images.

Zazkis et al. (1996) define visualization as an act in which connections between mental constructs and external objects perceived through senses are established. According to them, this can mean construction of mental processes or mental objects which are associated with external objects, or it can

[5] My personal comment: "Visual image" or "visual mental image" could be a better term in this connection.

mean construction of external objects which are identified with mental constructs. These two directions for the act of visualization are analogous with the processes of visualization presented by Gutiérrez. However, Zazkis and colleagues view the visualization clearly as an interaction between the cognitive structure of an individual and the external reality. According to my interpretation, Zazkis and colleagues consider only external constructs visual, whereas Gutiérrez does not at all emphasize the distinction between internal (cognitive) and external constructs. In his paper, visual constructs are primarily mental ones. This distinction shows that the term "visual" is ambiguous among researchers of mathematics education. In the following, visual constructs are considered as mental ones, which can be expressed externally.

2.5.2 Visualization as modelling

The process of using visualization as a thinking tool in working with the theory of mathematics can be understood also as modelling. In this process abstract mathematical concepts are interpreted visually. In the sense of Fishbein's definition for a model, this means that the definitions of the concepts in the formal axiomatic system of mathematics get visual counterparts, and these constitute a visual system representing the original abstract system. As emphasized in the previous section, it is very important to be aware of the restrictions of the model. It has to be noticed that the visual interpretations are only *prototypes* of the abstract concepts. Thus, it is important that visual interpretations are used in a controllable way such that their restrictions and deficiencies are taken into account. An uncontrollable use of visual images may lead to conflicts and erroneous conclusions (Aspinwall et al., 1997).

2.5.3 Visual vs. analytical reasoning

Analytic reasoning based on the use of symbolic representations and construction of logical inference chains is sometimes considered as an opposite thinking mode to the reasoning based on visualization. However, visual reasoning cannot be considered purely perceptual, but analytic argumentation is an essential factor in it (Dreyfus, 1994). Analytic reasoning is often very exact and detailed, but wider trends of the whole process may be bypassed in it. In contrast, visualization can reveal holistic features of the problem situation by simplifying and concretizing it, and thus it can operate as a map showing a direction for the reasoning process. This, however, requires that an individual deeply understands the connections between visual and symbolic/formal representations. According to Zazkis et al. (1996), "perhaps the most harmful, yet quite common difficulty with visualization is that students have shown a lack of ability to connect a diagram with its symbolic representation, a process some authors consider to be an essential companion to visualization" (p. 437).

At least in some situations, tendencies to visualization in mathematical problem solving seem to differ between individuals. For example, Krutetskii (1976) divided schoolchildren into analytic thinkers, geometric thinkers and harmonic thinkers. According to him, analytic thinkers had a very strong verbal-logical component in their reasoning, and it predominated over a weak visual-pictorial component. In addition, analytic thinkers neither had an ability nor felt a need to use visual supports in problem solving. In the case of the geometric thinkers everything was vice versa. Harmonic thinkers had a strong ability to use both verbal-logical and visual-pictorial components in their reasoning, but their preferences could vary. Also Clements (1982) proposed a corresponding division into visualizers, verbalizers and mixers. According to some older studies (Krutetskii, 1976; Presmeg 1985; 1986a), in which these kind of classifications have been applied at the school level, most of the successful students were non-visualizers. These kind of results created an impression that the use of visualization is in some way connected to weak success in mathematics. However, according to Presmeg (1986a), a notable reason for these results was that nonvisual methods were strongly emphasized by the curriculum, requirements, textbooks and practices of teaching, and this was not optimal for students who had a tendency to think visually.

Researchers have later paid particular attention to students' preferences to use visualization and other methods in their reasoning. Students' reluctance to use visualization has been considered as a problem. For example, Vinner (1989) designed a calculus course in which visual considerations were strongly emphasized, but despite that students seemed to have a tendency to avoid visual considerations. However, the reluctance to use visualization does not depend on the abilities of visual thinking: It has been shown that also students who are well able to think visually may be reluctant to do so (Eisenberg and Dreyfus, 1991). According to Eisenberg and Dreyfus, one notable reason for why students prefer algorithmic to visual reasoning is that visual reasoning is often cognitively yet more demanding. In addition, students may consider visual arguments non-mathematical and thus they believe them to be unacceptable.

Later, the divisions into visualizers and non-visualizers have been criticized and they have been considered useless, because both analytic and visual strategies have been found to be important in the rich understanding of mathematical concepts and in effective reasoning in problem solving (Zazkis et al., 1996). It is also notable that the line between visual and non-visual reasoning is unclear: According to Presmeg (1992), imagistic processing, which is often considered as a characteristic feature of visual thinking, may in an abstract form have a central role also in thinking modes which are conventionally considered non-visual. On the other hand, the nature of visual reasoning may in some cases be very analytic and formal (Arcavi, 2003). Therefore, instead of classifying individuals into visualizers and non-visualizers, it has been considered more important to study how people having different mathematical skills combine visual and analytic elements in their reasoning. One model describing the interaction between analytic and visual modes of reasoning is the V/A (visualization/analysis) model, which has been generated by Zazkis et al. (1996) and refined by Stylianou (2002). According to this model, visual approaches benefit from analytic thinking, and, respectively, analytic approaches can be enriched by visualization.

2.5.4 The role of visualization in mathematics

The ability to utilize visualization in an effective way in mathematical reasoning seems to create a substantial distinction between experts and novices. Stylianou (2002) noticed that mathematicians use visualization in a very systematic way, so that in their reasoning the visual and analytic steps are very closely connected and they interact with each other. Instead, many students have difficulties in analysing visual representations, and, therefore, they cannot utilize visualization in problem solving (Stylianou and Dubinsky, 1999). Students considered visual representations useful mostly in geometric problems, whereas mathematicians saw a wider variety of problems where visualization could be used (Stylianou and Silver, 2004). Also Raman (2002; 2003) and Merenluoto (2001) have compared students' and professional mathematicians' reasoning. Raman found that an essential difference was that mathematicians considered visual and formal arguments closely connected so that the visual arguments in an essential way contributed to inventing ideas in constructing a formal proof. Instead, students could not recognise connections between visual and formal arguments. Merenluoto has studied mathematicians' and high school students' conceptions about real numbers. She found that mathematicians' formal conceptions were strictly connected to their rich informal models, which were abstracted from the central characteristics of the concept at issue. Instead, high school students' conceptions about the real numbers seemed to be fragmented.

Even though mathematicians frequently use visualization when they construct arguments, they usually do not bring that out in their reports or presentations. Also in lectures given by mathematicians, arguments are often presented in their final formal forms and descriptions of the processes of inventing ideas may be omitted. At least to some extent, this convention has transferred also to school mathematics. According to Eisenberg and Dreyfus (1991) and Dreyfus (1994), this is a notable reason as to why students so often consider visual reasoning as a non-mathematical and unacceptable method. Furthermore, as told above, this is a notable reason for students to be reluctant to visualize in mathematical reasoning. In order to change the situation, Presmeg (1997) suggests that "if more teachers were aware of both the power and the possible pitfalls of visualization in mathematics, more students would be encouraged to overcome the disadvantages and benefit from the considerable strengths of using their imagery more fully in mathematics" (p. 310). This issue gives a challenge also for teacher education.

The position of visualization in mathematics has occasionally raised vivid discussion among the researchers of mathematics education (Presmeg, 2006b). Due to the development of cognitive psychology, for example, the wide potential of visualization in teaching and learning of mathematics has been recognised during the last decades. In addition to that it has been recognised that visualization is a very important tool which promotes understanding and assists in problem solving, the appreciation of the visual reasoning and visual arguments as such has increased among mathematics educators. Several researchers (Arcavi, 2003; Dreyfus, 1994; Rodd, 2000) have proposed that visual arguments should not be considered only as tools of thinking but they could be considered as an intended and accepted form of the final arguments. Due to the development of information technology, the use of computer graphics has become an important tool in mathematical problem solving. This has, furthermore, increased the role of visualization both in school mathematics, in the research of mathematics and in the fields where mathematics is applied. At the beginning of the 1990s, Zimmermann and Cunningham (1991) anticipated that "if present trends are any indication, it seems that mathematics will evolve in a direction which will make visualization even more important in the future than it is now" (p. 7). It seems that so far this anticipation has been fulfilled fairly well.

2.6 The dual nature of mathematics and mathematical argumentation

As explained above, there exist two tendencies which are typical of mathematics: tendency towards generalization, abstraction and rigour, and, on the other hand, tendency towards intuitive and holistic understanding. Even though these tendencies may seem to be opposite, in advanced mathematical reasoning both are present and they interact with each other. In the following, three classifications concerning mathematical argumentation are reviewed. They all are essentially based on the distinction between the two tendencies mentioned.

2.6.1 Privat and public arguments

Raman (2002; 2003) defines that *a private argument* is an argument which engenders understanding and that *a public argument* is an argument which is sufficiently rigorous for a particular mathematical community. Representatives of the "particular mathematical community" are, for example, reviewers of journals or teachers at school who judge the arguments. Private arguments have an essential role in facilitating the conceptual and holistic understanding of relationships between concepts. They are very often based on visualization. In contrast, a public argument should reveal step-by-step the progress of inference and justifications for each step. The construction of a public argument often requires procedural skills to carry out calculations and other procedures. However, the structure of a public argument does not often reveal broad trends and central ideas of the argument, which are very important in the construction process of the argument.

Raman defines three types of ideas needed in the process of proof production: *a heuristic idea* related to the private argument, *a procedural idea* related to the public argument and *a key idea* related to a connection between the private and public arguments. She defines these ideas in the following way:

"The first type of idea used in proof production is called a heuristic idea. This is an idea based on informal understanding, e.g. grounded in empirical data or represented by a picture, which may be suggestive but does not necessarily lead directly to a formal proof. A heuristic idea gives a sense of

understanding, but not conviction." (Raman, 2003; p. 322)

"The second type of idea used in proof production is called a procedural idea. This is an idea based on logic and formal manipulations which leads to a formal proof without connection to informal understanding. A procedural idea gives a sense of conviction, but not understanding." (ibid.; pp. 322-323)

"Finally, the third type of idea that can lead to proof production is called a key idea. A key idea is an heuristic idea which one can map to a formal proof with appropriate sense of rigor. It links together the public and private domains, and in doing so gives a sense of understanding and conviction." (ibid.; p. 323)

In summary, the heuristic idea *gives a sense that* something ought to be true, a procedural idea *demonstrates that* a particular claim is true, and the key idea *shows why* it is true.

As mentioned in the previous section, professional mathematicians, according to the results of Raman's study, are able to see connections between the private and public arguments, but inexperienced students consider them separately. Raman infers that proof productions of mathematicians are essentially based on key ideas, but students often do not have key ideas at all. This means that when a mathematician recognises a heuristic idea, he/she usually has no problems in translating it into a formal proof, and, respectively, a mathematician often easily recognise heuristic ideas through formal proofs. Due to that, mathematicians can use and construct heuristic and procedural ideas simultaneously so that both ideas clarify each other. Instead, students are often unable to utilize heuristic ideas in producing proofs. It is also possible that students are able to understand or even produce particular steps of a proof, but at a heuristic level they still cannot see why the claim is true. In addition, it is possible that students do not realise the fundamental role of proofs in mathematics.

2.6.2 Justifications and warrants

A corresponding classification of mathematical arguments is presented by Rodd (2000). The criterion in this classification is the effect of an argument on an individual. According to Rodd's terminology, *a justification* is an argument which has an effect on intuitive beliefs. It gives a reason to believe that a claim ought to be true. Instead, *a warrant* is an argument which convinces one that the claim is undoubtedly true. It exhibits a logical inference chain which shows the truth of the claim. At the intuitive level, a warrant is not necessarily very explanatory, and thereby justifications are also needed. A formal proof is the most usual type of a mathematical warrant. However, it has to have personal effect on an individual: If the individual does not understand the steps of the inference chain in the proof, the proof is not a warrant for him/her. On the other hand, Rodd shows that a warrant does not necessarily have to be a formal proof, but it can be based, for example, on visual reasoning.

Even though foundations of Raman's and Rodd's classifications are in some extent different, they quite far correspond with each other: Very often, private arguments in Raman's classification are justifications in Rodd's classification, and public arguments correspond with warrants. However, Raman's classification emphasizes the connection between different argument types, but in Rodd's classification different types of arguments do not necessarily have any connections between them. Justification may be based, for example, on experimental inductive reasoning, which is against the nature of mathematics, but which can yet have an effect on personal beliefs.

2.6.3 Syntactic and semantic proof productions

Weber and Alcock (2004) distinguish between *syntactic* and *semantic proof productions*. They define the syntactic proof production as an occurrence in which "the prover draws inferences by manipulating symbolic formulae in a logically permissible way" and the semantic proof production as an occurrence in which "the prover uses instantiations [6] of mathematical concepts to guide the formal inferences that he or she draws" (p. 209). In the following, the terms *"syntactic knowledge"* and *"semantic knowledge"* refer to knowledge and abilities needed in these types of proof production.

Weber and Alcock regard the semantic knowledge especially as a guide for the syntactic proof production: It acts as a map which guides in choosing proper facts and theorems to apply. By using semantic knowledge, an individual can in a meaningful way make sense of the claim to be proven, get suggestions about inferences that could be drawn and become convinced at an intuitive level about the truth of the claim. These points emerged also in Iannone's and Nardi's study (2007), in which mathematicians were interviewed and their views about the importance of semantic and syntactic knowledge were inquired. The mathematicians also believed that the semantic knowledge contributes to the thorough understanding of mathematical concepts and to the flexible use of the concepts in different situations. About the syntactic knowledge the mathematicians thought that it would help in defining and clarifying concepts without ambiguity and act as a shared language of mathematics making communication easier, as well as a "checking device" for the reasoning which is based on intuition or on the semantic knowledge. In addition, the syntactic knowledge makes the manipulation of formal statements possible, which is often an effective tool in mathematical reasoning.

In their study, Weber and Alcock (2004) compared problem solving processes in proof production between undergraduates and doctoral students, and they concluded that doctoral students utilized semantic proof production more than undergraduates. Doctoral students used various instantiations about concepts in their reasoning, whereas undergraduates mainly used only formal definitions in their reasoning. In addition, Weber (2001) showed that deficiencies in the strategic (semantic) knowledge often prevent undergraduates from solving a problem even though they possess all knowledge and procedural skills

[6] By an instantiation Weber and Alcock mean "a systematically repeatable way that an individual thinks about a mathematical object, which is internally meaningful to that individual" (p. 210). The instantiations may be considered either through their physical representations or through mental images.

needed in the problem. Therefore, the semantic knowledge has a crucial role in effective mathematical reasoning.

Creating and using of instantiations make semantic proof production complex and challenging. In semantic proof production one has to be able to create instantiations which are "rich enough that they suggest inferences that one can draw" for abstract mathematical concepts (Weber and Alcock, 2004; p. 229). However, the instantiations should not suggest inferences which are in contradiction with the formal theory. Therefore, it is important that the instantiations are created and used in a controllable way. They should be controlled by using the formal definitions of the corresponding concepts. Thus, it is important that one thoroughly understands the connections between the instantiations and the formal definitions.

2.6.4 Other analogous classifications

Weber's and Alcock's classification is originally intended for an analysis of the process of proof production. However, it is analogous with the two classifications presented above. Private arguments based on heuristic ideas, justifications influencing personal beliefs and semantic proof production are all connected to intuitive and holistic argumentation. In contrast, public arguments, warrants and syntactic proof production refer to general and rigorous, but yet procedural argumentation. In addition to these three classifications, Davis and Hersh (1981) and Hanna (1990), among others, have also distinguished between arguments which convince (remove all doubt about the truth of a claim) and arguments which explain. Weber and Alcock draw an analogy between their own classification and Skemp's (2006) classification, in which a distinction between *instrumental understanding* (understanding *what* to do) and *relational understanding* (understanding also *why* to do) is made. Furthermore, an analogy can be drawn between any of the above mentioned classifications and the classification of mathematical knowledge into *procedural* and *conceptual knowledge* (Hiebert and Lefevre, 1986: Haapasalo and Kadijevich, 2000). This classification, as well as that of Skemp, concerns not only mathematical argumentation but mathematical understanding in a broader sense.

All the classifications presented above are summarized in Table 1. Even though they are all to some extent analogous, they cannot be considered identical, but each of them has been created for their own purposes. In the following section I introduce my own classification for the purposes of this thesis. Also this one is analogous to the classifications presented above.

Table 1: Different classifications of mathematical understanding into procedural/formal and intuitive/holistic components.

Author	Procedural/Formal component	Intuitive/Holistic component
Raman	Public argument / Procedural idea	Private argument / Heuristic idea
Rodd	Warrant	Justification
Weber and Alcock	Syntactic understanding	Semantic understanding
Skemp	Instrumental understanding	Relational understanding
Hiebert and Lefevre / Haapasalo and Kadijevich	Procedural knowledge	Conceptual knowledge

2.7 Definitions of the informal and formal arguments

According to Toulmin's (2003) model of argumentation, an argument has always three main elements: *The data* is the information concerning the initial state, *the conclusion* is the claim which is argued, and *the warrant* is an explanation for why the data necessitate the conclusion. Often a conclusion drawn from a certain data is possible to be argued by using different warrants, that is to say, a claim can have several arguments.

This thesis, like many studies reviewed above, deals with students' preferences and conviction regarding different arguments in reasoning, but, in addition, I have had as a central goal to study students' abilities to produce different kind of arguments for mathematical claims. Unlike in the studies reviewed above, in some tasks of this study students were explicitly asked to produce, for given claims, several arguments whose final forms were required to be fundamentally different. Therefore, in the classification of arguments which is used in this study, the arguments have been classified according to the foundations of their warrants.

I define that an argument is *formal* if its warrant is purely based on the elements of the formal axiomatic system of mathematics. In other words, an argument is formal if it explicitly shows how the conclusion logically follows from definitions, axioms and previously proven theorems and from the given data. A formal argument has to be also systematic and rigorous.

Furthermore, I define that an argument is *informal* if its warrants are based on the use of *informal interpretations* of concepts or situations which the argument is concerned with. By informal interpretations in this study I mean *visual* or *physical* interpretations. In the informal reasoning [7] the aim is to find appropriate concrete counterparts for abstract mathematical concepts and through them to interpret the abstract problem field in question. These counterparts are often constructed through visualization, but they can also

[7] I understand that *a reasoning* is a mental construction process, and *an argument* is a product of this process.

be found from physical reality. The informal interpretations concerning the problem field should form a model of the abstract mathematical situation (see Section 2.4), and the informal arguments are based on this model. In Section 2.8 some examples of visual and physical interpretations and of their use concerning derivative are presented.

In Toulmin's model, the term "warrant" does not have the same meaning as it has in Rodd's terminology (see Section 2.6). As Inglis et al. (2007) interpret, Rodd regards a warrant as an element which *removes* the uncertainty about the truth of the conclusion, whereas Toulmin sees it as an element which *reduces* this uncertainty. In the above definitions the warrant has in principle the same meaning that it has in Toulmin's model. However, if a formal argument is exact and detailed enough, it should remove the uncertainty about the truth, that is, it should be a warrant also according to Rodd's terminology.

In my classification both types of arguments are considered independent, whereas in all three classifications presented above, the different arguments are considered as elements of mathematical reasoning or proving processes. Thus, my classification to some extent offers possibilities to classify arguments on the basis of their final forms without analysing their producing processes or their effects on an individual's understanding. In this sense, my classification can be considered less relative than the other three classifications presented above. However, this classification, like the other classifications, cannot be considered absolute, because in the cases of some arguments it may depend on personal interpretations whether the argument is regarded to be bound to the formal axiomatic system or whether it is regarded to be based on other interpretations about the meanings of the used concepts and their qualities. Especially, it may be controversial whether the connections to the formal axiomatic system are shown explicitly enough. In practice, all phases of the inference chain of a formal argument are almost never shown explicitly, but the acceptance of the argument is dependent on whether the argument convinces an individual (*individual acceptance*) or a society (*social acceptance*) of the fact that the conclusion is compatible with the formal axiomatic system. Some interpretations which include visual or physical components may yet have an abstract nature and they may be very exactly inferred from the elements of the formal axiomatic system. In addition, arguments based on these interpretations may also be very exact and general. This also makes it difficult to draw a precise line between informal and formal arguments. It is also notable that, in some cases, a warrant of an argument may include both informal and formal elements. For example, in long proofs some details may be argued informally even though the structure of the proof is formal otherwise.

The reviews above concerning the dual nature of mathematics and mathematical thinking (Sections 2.1 - 2.6) indicate that it is important for teachers of mathematics to be able to use both informal and formal arguments effectively in their own reasoning and that they are aware of pedagogical potentials of arguments of both types. The benefits of the informal argumentation are the same that Raman presents for heuristic ideas and private arguments, Rodd for justifications and Weber and Alcock for semantic understanding. The formal argumentation also in several ways corresponds with procedural ideas, public arguments, warrants and syntactic understanding. In fact, heuristic ideas and semantic understanding are necessary resources which are needed in constructing informal arguments, and procedural ideas as well as syntactic understanding are needed in formal argumentation processes.

A distinction between the reasoning based on a concept definition and the reasoning based on a concept image have been made in several studies (Tall and Vinner, 1981; Vinner, 1982; 1991; Vinner and Dreyfus, 1989; Rösken and Rolka, 2007). This way to classify mathematical reasoning is essentially the same as my classification into informal and formal arguments. However, I regard the concept definition, or, in fact, its personal interpretation, as an essential part of the concept image (see Section 3.1), and thus I do not want to use these terms in the connections relating the classification of reasoning.

Because derivative is a central concept in this thesis, some examples about informal and formal interpretations concerning this concept and arguments based on these interpretations are presented in the next section.

2.8 Informal and formal interpretations in the case of the concept of derivative

Formally, derivative in the case of a real-valued function of a single variable is defined as a limit of the difference quotient in the following way:

A function $f : \mathbb{R} \to \mathbb{R}$ is called *differentiable at a point* $x_0 \in \mathbb{R}$ if the limit

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. The value of the *derivative* at the point x_0 equals to this limit. Alternatively, the limit of the difference quotient can be presented in the following form:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Furthermore, a function is called *differentiable in an open set* A if the set A is a subset of the domain of the function and the function is differentiable at all points of the set A. Furthermore, a function is called *differentiable* if it is differentiable at all points in its domain.

Visually, derivative describes the *steepness* of the graph of a function. The sign of the derivative reveals whether the graph is going up or down, and the absolute value describes how steep the uphill or the downhill in the graph is.[8] If we want to use a more exact visual interpretation, we can say that the derivative at a given point is the slope of the tangent line drawn to the graph at this point. The derivative can also be illustrated by sliding a pencil along the graph from left to right (Hähkiöniemi, 2006a). In that case the

[8]In this connection, it is assumed that the graph has been drawn to the ordinary Cartesian co-ordinates in which the values of the variable are situated in the horizontal axis so that their values increase from left to right and the values of the function lie in the vertical axis increasing from down to up.

pencil always lies over the tangent line, and the nib of the pencil points in the direction to which the graph is going at the point in question. The derivative is proportional to the angle between the pen and the x-axis. In fact, it is the tangent of this angle. The "sliding pen" interpretation brings out dynamics to the visualization of derivative.

The limiting process presented in the definition of derivative can be interpreted visually, for example, in the following way: First, the difference quotient is interpreted as the slope of the secant line which goes through points $(x_0, f(x_0))$ and $(x_0+h, f(x_0+h))$. Then, during the limiting process, the point $(x_0 + h, f(x_0 + h))$ slides along the graph toward the point $(x_0, f(x_0))$, and at the same time the secant line turns in the direction of the tangent line which goes through the point $(x_0, f(x_0))$. The limiting process itself is dynamic, and so is also its visual interpretation.

The differentiability at a point visually means the possibility to draw an unambiguous tangent line to the graph at the point in question. In the interpretation presented above for the limiting process, the secant line has to approach the same tangent line independently whether the point $(x_0 + h, f(x_0 + h))$ approaches the point $(x_0, f(x_0))$ from the left side or from the right side (that is, whether h is positive or negative). This indicates that the graph of a differentiable function has to look smooth: It cannot, for example, include corners or jumps. So, the smoothness of the graph is one visual interpretation for differentiability.

An instantaneous rate of change of some quantity can be regarded as *a physical interpretation* of derivative. For example, instantaneous speed is the derivative of the distance passed, instantaneous acceleration is the derivative of the speed, the electric current is the derivative of the flowing electric charge through a surface, and so on. Also these interpretations can be used as thinking tools when engaging with mathematical problems dealing with derivative.

The limiting process in the definition of derivative can be interpreted physically in the following way: First, the difference quotient can be interpreted as an average rate of change between two physical states. Then, in the limiting process the distance between these states becomes smaller and smaller, and thus the average rate of change approaches the instantaneous rate of change. A physical interpretation for differentiability could be stepless changes of rate. According to the fundamental theorem of calculus, every continuous function has an integral function, which is naturally differentiable. Therefore, if the changes in a physical event are stepless, then the function describing this event is differentiable. This "stepless changes of rate" interpretation for differentiability is not mathematically complete: Continuity is not a necessary condition for integrability.

For example, let us take the theorem stating that the derivative of a constant function is zero. Formally, this can be proven by a minor calculation starting from the definition of derivative (see Task 3b in Appendix 3). Visually, the graph of a constant function presented in Cartesian co-ordinates is a horizontal straight line. The theorem can thus be argued by explaining that the graph does not include any uphills or any downhills, it is nowhere going up or down, or that the tangent drawn to the graph is everywhere a horizontal straight line. Physically, one can explain that if a quantity is constant, its value does not change, and thus the rate of change is everywhere zero.

2.9 The basic assumptions of the study

This study is based on the assumption according to which the truth values of mathematical statements are determined by a formal axiomatic system, which is in accordance with Hilbert's ideal (see Section 2.1). All elements of mathematical knowledge are assumed to belong to some system of this kind. If the formal axiomatic system were assumed to be global and unchangeable, this foundation could be considered objectivistic. However, I consider both the construction and the learning of this system as a socio-cultural process, and thus this standpoint is well in accordance with the assumptions of the relativistic paradigms, such as socio-cultural or socio-constructivistic paradigms. Generally speaking, it is not necessary to assume anything about the wideness in which the system is socially shared. However, in the case of the concepts of derivative, differentiability and continuity, the definitions presented in the test form (see Appendix 1) are very widely shared among mathematicians and mathematics educators. Furthermore, these definitions are essentially based on the definition of limit, which is widely shared as well. Therefore, it is justified to assume that there exists a widely shared formal axiomatic system which determines what is true and what is false regarding these concepts. There is a wide consensus among mathematicians and mathematics educators of the fact that the teaching of analysis should be based on this system, that is, the views about these concepts presented in teaching should be in accordance with this system. In this study, the foundation mentioned above infers that the views and reasoning of the subjects are evaluated with respect to this system. Views which are (according to the reviewer's interpretation) in contradiction with the given definitions are considered erroneous independently of whether they are internally coherent or not.

I also assume that even though the definitions of mathematical concepts belong to a formal axiomatic system which may be very widely shared, the concepts themselves do not carry any objective or global meanings concerning empirical reality. Therefore, in the learning process the goal is to create meanings for concepts at a personal level, and this creating process is restricted by the definitions. Respectively, the whole formal axiomatic system can be thought to be meaningless as such. These foundations of mathematical knowledge and learning are very well in accordance with the constructivist world view. They are also in accordance with Hilbert's ideal to make mathematical concepts fully abstract and thus independent of empirical reality.

3 Concept image and its coherence

3.1 What is a concept image?

In 1981 Tall and Vinner published a famous paper (Tall and Vinner, 1981), in which they introduced the term *"concept image"*. According to them, a concept image "consists of all the cognitive structure in the individual's mind that is associated with a given concept" (p. 151). The cognitive structures may, for instance, be mental images or interpretations based on different kind of representations about properties or processes concerning the concept. Therefore, the concept image is essentially connected to an individual's personal way to understand a concept. Tall and Vinner quite strongly emphasize the distinction between the concept image and the formal theory in their paper. In fact, it seems that the concept image was originally created as a tool to analyse the distinction between the personal way of understanding and the formal theory. This distinction comes out more strongly in Vinner (1991) where the reasoning based on the concept image has been explicitly separated from the reasoning based on the *concept definition*. The central question in this paper was whether students' reasoning was based on the concept definition or on the concept image. Both in that paper and in Tall's and Vinner's original paper the concept image is rather seen as a collection of vague conceptions concerning the concept which do not have any connections to the concept definition. Instead, Tall has later considered the concept definition as a part of the concept image (Tall, 2003a; 2005).

In this study, I fully accept the definition of the concept image as stated above. However, I do not wish to create any confrontation between the concept image and the formal theory but I want to emphasize that the formal theory can have -and it should have- an essential effect on the concept image. According to the ontological assumptions presented in Section 2.9, it can be assumed that there exists a formal axiomatic system which includes the formal concept definition. This definition is in the raw only a queue of symbols, and thus an individual has to create a meaning for it, that is, he/she has to interpret the definition by means of his/her existing knowledge structure. I call this interpretation a *personal interpretation of the formal concept definition* and regard it as a part of the concept image (see Figure 2 in Article D). The personal interpretation of the concept definition may be dependent on the context in which the definition appears or in which it is used. In some situations, the definition may be required to be interpreted informally, whereas in a situation where only formal reasoning is used a formal interpretation may be sufficient.

3.2 What does the concept image include?

The concept image includes all the *conceptions* that an individual has about the concept. For example, in the case of derivative, an individual may have conceptions about the differentiation rules, about the definition of derivative, about the visual or physical meaning of derivative, about the relationships between derivative and some other mathematical concepts, and so on. The conceptions can be regarded as basic elements of the concept image. They form a basis for the *knowledge structure* concerning the concept. Such structures as *mental images* and internal *representations* about the concept can be considered either as conceptions or as structures which consist of several conceptions. The role of mental images in visualization has been considered in Section 2.5, and, in the following, some views about the representations and about their role in mathematical thinking are reviewed. In the literature, several definitions have been presented for the terms "conception", "mental image" and "representation", and they have often been used without deliberating their exact meaning. Sometimes it may be difficult to recognise the difference between their meanings.

A simple, traditional definition for a representation is that it is a configuration which, as a whole or part by part, represents something else (Goldin and Kaput, 1996). Representing in this case means correspondence, association, standing for, symbolizing or other interaction between configurations. According to Goldin and Kaput, *internal representations* are mental configurations of individuals, whereas *external representations* are physically embodied, observable configurations, such as words, pictures, symbols, and so on. Only external representations can be perceived, but, by observing an individual's actions with the external representation some conclusions about the internal representations can be made. The internal representations concerning a certain concept can be regarded as elements of the concept image. It has been seen convenient to consider representations as representational systems, because their inner structures and relationships to other representations are complicated, and, thus, it is difficult to isolate representations from each other (Goldin and Kaput, 1996; Goldin, 1998). Internal representation may be verbal, symbolic, visual, kinesthetic or formal, but in a broader sense, also affects and factors concerning strategic planning, monitoring or decision making may be regarded as representations. This broad variety of representations comes out in Goldin's classification for representation systems. However, in this study the representations are primarily regarded as cognitive structures.

Many researchers have aimed to focus the research about representations on students' use and construction of representations, and, therefore, they have regarded the representations rather as *tools for thinking* (Hähkiöniemi, 2006a; 2006b). They have emphasized that representations as such are not important, but that it is more crucial what possibilities the representations offer to an individual's thinking. Therefore, this view emphasizes more the active role of the learner, and the distinction between external and internal representations is considered irrelevant. Instead, the representation is understood to consists of "external and internal sides which are equally important and do not necessarily stand for each other but are inseparable" (Hähkiöniemi, 2006a, p. 56).

Connections between different kind of representations enable flexible mathematical reasoning. Goldin and Kaput have defined that a connection between two external representations is *weak* if an individual is able to predict, identify or produce one representation from the other and that the connection is *strong* if an individual is from a given action upon another representation able to predict, identify or produce the results of the corresponding action on the other representation. Even though representations in these definitions are external, the connections between them are internal, that is, they are part of the knowledge structure. Following the view of representations as thinking tools, Hähkiöniemi has defined that a person makes an *associative connection* between two representations if he/she changes from one representation to another and that a person makes a *reflective connection* if he/she uses one representation to explain the other. In Article D in Section 2.2 some examples of different kind of connections between representations in the case of derivative are presented. The classifications of connections mentioned above have been established for representations, but it might be possible to classify all kinds of conceptions correspondingly.

Conceptions, mental images, representations and connections depend on the context, and thus it is impossible to do an absolute mapping about them. Using a context-specific analysis, it is yet possible to study which of these are used in each context. If some conceptions, mental images, representations or connections come out in several, different kind of contexts concerning a certain concept, it can be said that they probably in general have a significant role in the concept image of this concept.

3.3 The coherence of the concept image

The concept image is built up, evolved and changed through experience. In order to offer potential for creative mathematical thinking, the concept image has to be wide and multifaceted. It is, however, important that the concept image is well organized. In this study, I have used the term *"coherence of a concept image"* to refer to the level of organization of the concept image. The following list includes some properties of a highly coherent concept image.

1. An individual whose concept image is considered has a clear personal conception about the concept.

2. Conceptions, cognitive representations and mental images concerning the concept are well connected to each other.

3. The concept image does not include internal contradictions.

4. The concept image does not include conceptions which are in contradiction with the formal axiomatic system of mathematics.

Essentially, this list is the same as presented in Article D in Section 2.2. Only minor linguistic changes have been made.

According to the ontological assumptions presented in Section 2.9, mathematical concepts are meaningless as such, but individuals may have different kind of conceptions about them. The criterion 1 in the above list indicates that an individual has a clear *personal view* what the concept means *for him/her*. This personal view may depend on the context so that the individual sees the meaning of the concept differently in different contexts. Certainly, he/she may have different views about the meaning of the concept also in the same context. In any case, it is important for the coherence of the concept image that the different conceptions are not in contradiction with each other or with the formal theory of mathematics (cf. the criteria 3 and 4).

The criterion 2 refers to mental connections between the elements of the concept image. Both the amount and the nature of the connections are important. In Section 3.2, Goldin's and Kaput's classification for weak and strong connections and Hähkiöniemi's classification for associative and reflective connections between representations were reviewed. On the basis of the definitions of these connection types, it is justifiable to consider strong connections better indications than weak connections about the coherence of the concept image, and, respectively, reflective connections can be considered better indications than associative connections about the coherence of the concept image. However, with respect to the coherence, weak and associative connections are better than erroneous or missing connections.

During a mental process, only a certain part of the concept image is activated. Tall and Vinner (1981) call this part an *evoked concept image*. It is possible that factors of the concept image which are activated at different times are contradictory. If the contradictory factors are activated, for instance, always at a different time or in different contexts, the contradiction may remain unconscious. However, these factors would cause a cognitive conflict if they were activated at the same time. Thus, Tall and Vinner call them *potential conflict factors*. Due to a cognitive conflict, a contradiction becomes conscious, and an individual may try to resolve it. The revealing and resolving of contradictions usually increase the coherence of a concept image, but, instead, conscious but unsolved contradictions often cause confusion and feelings of insecurity also in a broader context. The criterion 3 in the above list refers both to conscious and unconscious contradictions in the concept image.

The criterion 4 refers to conceptions which are erroneous with respect to the formal theory of mathematics. Tall and Vinner (1981) write about that: "A more serious type of potential conflict factor is one in the concept image which is at variance not with another part of the concept image but with the formal concept definition itself..." "...students having such a potential factor in their concept image may be secure in their own interpretations of the notions concerned and simply regard the formal theory as inoperative and superfluous" (p. 4). It is possible that a concept image is in a broad extent internally coherent, but it includes elements which are erroneous with respect to the formal axiomatic system. The case study presented in Article D offers an illustrative example of that. However, according to the basic assumptions of this study (see Section 2.9), the truth values of all conceptions are determined by the formal axiomatic system. Therefore, mathematical reasoning should be tied up with this system. In practice, this requires that the formal theory has a central role in the reasoning and that the formal theory is on the personal level interpreted so that it is possible to draw conclusions which are in accordance with the formal theory.

Evidently, the criteria mentioned in the above list are dependent. For example, if the concept image includes internal contradictions (the criterion 3), it can be considered as an indication of inadequate connections between the elements of the concept image (the criterion 2), and, furthermore, internal contradictions almost surely indicate also contradictions with respect to the formal theory (the criterion 4).

When exploring mental structures, it is important to take into account that these structures may change all the time. An individual may reflect and change his/her conceptions, create new connections between the elements of the knowledge structure, forget some details which are not in active use, and so on. Therefore, conclusions about conceptions, mental images, representations, connections or contradictions can actually concern only the moment when they are observed. By analysing an interview session entirely, it may be possible to conclude when the changes in the mental structure happen and what consequences these changes may have. It is usually not appropriate to attempt to determine the overall level of the coherence of the concept image. Instead, the criteria for the coherence of the concept image offer a framework to analyse single observations about reasoning: The single observations can be interpreted either as indications of the coherence or indications of the incoherence of the concept image. In the studies presented in Articles B and D, this kind of analysis has been used in order to explain reasons for observed erroneous conclusions.

The term *conceptual knowledge* refers partially to the same issues as the term coherence of the concept image. An essential difference between these terms is that the coherence of the concept image refers to the understanding of a single concept, whereas the term conceptual knowledge refers to the thorough understanding of the relationships between different concepts. These terms are compared in more detail in Article D in Section 2.2. In the same article in Section 2.3, some previous studies about mathematics students' reasoning in the area of calculus or basic analysis are reviewed. These studies reveal several cases where the coherence of the concept image seems to be inadequate.

4 The research project

4.1 Goals of the study and research questions

The main goal of the present study was to explore how students understand the informal and formal interpretations in the case of derivative, how they use these interpretations in problem solving, and how they understand the connections between these interpretations. The students were subject teacher students who were at the middle or at the final phase of their studies. Thus, this study can be placed among studies about the subject matter knowledge of prospective mathematics teachers. However, the teaching viewpoint is not strongly emphasized, and thus this study can rather be considered as an investigation about students' reasoning in the tertiary-level mathematics.

With respect to the mathematical content, research is restricted to the concepts of derivative and differentiability of real-valued functions of a single variable. To some extent, the concepts of continuity and the limit are also considered. These concepts were chosen, because they are important both in mathematics at upper secondary school and in elementary courses at university. In addition, both visual, physical and formal interpretations and several computational algorithms are widely used in connection with these concepts, and thus these concepts seemed to be appropriate for a study about different kind of interpretations and connections between them.

Originally, the research questions were formulated in the following way:

1. How well do the students understand the most important visual interpretations of the concepts of derivative and differentiability?
2. How well do the students understand the connections between these interpretations and formal definitions of these concepts?
3. How well are the students able to use the visual interpretations in reasoning and in argumentation?
4. How well are the students able to use the formal definitions in reasoning and in argumentation?
5. What is the role of the visual interpretations and what is the role of the formal definitions in students' reasoning and argumentation?

The "most important visual interpretations" meant the interpretation of derivative as a slope of a tangent line and as a measure of the steepness of the graph (see Section 2.8). Later, the more general term "informal" was chosen instead of the term "visual". As stated in Section 2.7, the term "informal", in addition to visual interpretations, refers to all kinds of interpretations which are not formal.

During the analysis the research question 5 proved to be particularly interesting. Therefore, in this thesis, this question is emphasized more than the other ones. Moreover, the issues concerning erroneous conclusions and coherence of the concept image emerged during the analysis, and it turned out that some parts of the collected data were appropriate also for studies about these issues. Because these issues proved interesting both in practical and in theoretical sense, their treatment was included in this thesis.

4.2 Methodology

4.2.1 The design of the study

In the present study, both quantitative and qualitative methods have been applied. The quantitative data were collected by a written test which was arranged at six Finnish universities and at one Swedish university between October 2004 and March 2005. In total, 160 Finnish and 20 Swedish mathematics subject teacher students took part in it. Many of these students were majoring in mathematics, while others had mathematics as a minor subject (see Table 2 in Appendix 5). The qualitative data consist of the individual interviews of 21 participants of the written test. The interviews were conducted only at three Finnish universities. At these universities, the test was arranged first, and the answers were quick-analysed in order to select the interviewees. The interviews were usually held a couple of days after the test. Thus the students probably remembered well the situation in which they answered the test, and at the interviews it was possible to discuss their reasoning and feelings in this situation.

It was required that all participants had passed at least 20 Finnish credits (about 35 ECTS credits) in mathematics. The reason for this was that the goal was to study the reasoning of students who had experience about university mathematics and who probably would work as mathematics teachers. Those who had passed this amount of studies in mathematics had probably also passed the basic courses in which derivative was considered. Answers of 14 Finnish students had to be excluded from the study, because they did not have enough credits. In addition, because the teaching practices and the degree requirements at Swedish universities differ from those at Finnish universities, the Swedish students were not included. Therefore, the final number of participants was 146. This is an extensive sample of Finnish subject teacher students in mathematics, because in Finland about 150-250 mathematics subject teacher students graduate yearly, and the sample consisted of students from all but one university providing mathematics teacher education in Finland. However, it cannot be considered as a representative sample, because no sampling method was applied in the selection of the participants of the test and because nothing is thus known about the students who did not participate. At most universities, the test was arranged in a lecture of educational studies. During each lecture, all those present also took part in the test. In most cases, the lecture in question was included in the study programme of the course, and thus, in these cases, participation in the test was part of the completion of the course. Indeed, the success in the test did not have any effect on the completion of the course.

The main criterion in the selection of the interviewees was the goal to select different kind of students by their success in the test: The aim was to select students who had succeeded well in the tasks requiring informal reasoning but poorly in the tasks requiring the use of formal definitions, students who had succeeded well in formal reasoning but poorly in the informal reasoning, students who had succeeded well both in the informal and in the formal tasks and students who had succeeded poorly in both kind of tasks. Students' possibilities and willingness to take part in the interviews were also taken into account. It was not particularly important to what extent the group of the interviewees constituted a representative sample of all the participants of the test.

Some statistics about the participants is presented in Appendix 5. Distributions about gender, main subject, number of passed credits in mathematics, duration of studies, study success in mathematics and aims for the future are presented first for all participants and then separately for the interviewees. There were fewer females among the interviewees than in the whole sample, and the mathematics majors were overrepresented among the interviewees. Especially, the difference in the proportions of gender was not intentional.

Each interview might have offered data for a case study. However, only a few of them were analysed thoroughly in the studies presented in this thesis. Article D is based on an analysis of one interview and Articles A and B on two interviews. In addition, it was possible to analyse interviews by coding them quantitatively. This kind of analysis has been used in the study presented in Article C.

Methodologically, the design of this study can be considered as a *mixed method design*, and according to Creswell's and Plano Clark's (2007) terminology, particularly as an *explanatory mixed method design.* In the mixed method design in general, quantitative and qualitative data concerning the same research problem are mixed in some way in order to achieve a better understanding of the problem than either of the data alone could offer. In the explanatory mixed method design, the quantitative data are first collected and analysed, and then some interesting results or cases are selected to a more profound qualitative study. Because the test answers had an essential role in the selection of the interviewees and the discussions at the interviews partially considered the interviewee's test answers, the study in its entirety can be considered as an explanatory mixed method study. In addition, at the analysis phase, some results of the quantitative study raised questions which could be studied by analysing the interviews. For example, the starting point of the study about erroneous conclusions, which is considered in Articles B and D, was the result according to which more than one fourth of the participants considered the function presented in the question 2d in the test form discontinuous but yet differentiable (see Appendix 4).

Several parts of the study are based either only on the quantitative data or only on the qualitative data: The study presented in Article E is based only on the quantitative data, whereas studies in Articles A, B and D consider only the interviews, except that the starting point of the studies in Articles B and D was the surprising result of the quantitative study mentioned above. These studies can be regarded as independent studies, which have their own goals. Therefore, to some extent, this thesis can be considered as a collection of separate studies.

Some features of *triangulation mixed method design* (Creswell and Plano Clark, 2007) are also included in this study: In the triangulation mixed method design both the quantitative and qualitative data are collected and analysed first independently. After that, mixing can be made by comparing results of both studies or, for example, transforming qualitative data into quantitative data and then comparing two quantitative data sets. To some extent, the whole study can be considered as a triangulation mixed method study: It consists of separate quantitative and qualitative studies which are mixed in the discussion section. The study presented in Article C is a clear triangulation mixed method study in which transformation has been made. In this study, some parts of the qualitative data were coded into a quantitative form and the conclusions were drawn by mixing this transformed data with the quantitative data on the test results.

4.2.2 The test form

The form used in the written test is presented in Appendix 1. The answering time was 90 minutes, and the test was arranged under supervision. The test form was delivered to all participants of the lecture in which the test was arranged, and the participants were required to spend at least 30 minutes for answering the test. No additional tools besides pen, answering paper and the test form were allowed.

On the first page of the test form, some background knowledge was inquired. Due to the possible call into an interview, the answers had to be identified, and for that reason the names of the participants were inquired. However, some participants refused to give their name, and in these cases the use of a pseudonym was accepted. However, the use of pseudonyms was one reason why it was not possible to get data about participants' study history from study registers of the universities. Thus, questions about their main subject, secondary subjects, number of passed credits in mathematics and the starting year of university studies were included in the test form. The participants were also asked to describe their study success verbally. Alternatively, the average grades of passed studies could have been inquired, but this question might have been more difficult for the students, and thus the resulting answers might have been more unreliable. In order to motivate participants to answering, they were promised feedback about their answers by email. Thus they could regard the test situation also as a learning opportunity.

In Appendix 5 some statistics about the participants' answers to the background questions is presented.

On the second page of the test form, the definitions of continuity, derivative and differentiability were given. The interest of the study was rather to explore how the students can use the formal definitions than to explore how well they remember them. In addition, some participants probably had not worked with these concepts for a long time. That is why it was reasonable to give the definitions at the test situation. Instead of the so-called ϵ-δ -definition, the definition of continuity was given in the form which was based on the limit.

This definition was sufficient for the tasks in this test, and it was assumed to be easier for the students to understand and use. At upper secondary school, the definition of continuity is usually given in this form.

The goal of Tasks 1 and 2 was to measure the visual understanding of derivative, differentiability and continuity. These tasks were thus established on the basis of research question 1. In Tasks 3 and 4, a claim was asked to be argued both by explaining informally (visually) and by proving formally. The claim in Task 3 was assumed to be very simple, and the claim in Task 4 should be more complicated. The informal explanations are connected to research question 3 and the formal proofs to research question 4. Task 5 was established in order to study how the students understand the connections between informal interpretations and formal definitions (research question 2). In this task students were on the basis of the definition asked to argue why derivative can be interpreted as a slope of a tangent line and why differentiability can be interpreted to mean smoothness (cornerlessness) of the graph. In both questions in Task 5, it was important to understand visually the limiting process in the definition of derivative. Task 6 was planned to be a more complicated problem, in whose solution both informal and formal reasoning would have been needed.

4.2.3 The interviews

The tasks presented in Appendix 2 formed a suggestive basis for the interviews. However, the progress of each interview was essentiallly dependent on the interviewee's views, which came out during the discussion, and on his/her success in solving the problems of the test and of the interview. The design of the interviews can thus be considered semi-structured: According to Kvale (1996), a *semi-structured interview...*

"...has a sequence of themes to be covered, as well as suggested questions. Yet at the same time there is an openness to changes of sequence and forms of questions in order to follow up the answers given and the stories told by the subjects". (p. 124)

The interviews were partially planned personally for each interviewee on the basis of the his/her background (see the questions about the background in the test form, Appendix 1) and test answers. For example, with students who seemed to have poor skills in the area of derivative, we discussed only some basic issues concerning the interpretations of the concepts and passed the more demanding tasks (Tasks 3 and 4, see Appendix 2) partially or entirely. With some students we discussed also test answers which were erroneous or otherwise seemed to reveal interesting views needing more explanation.

The same definitions as in the test were on view during the interviews. The interviewees could also see their test answers. The interviews were videotaped so that the camera was focused on the answer sheet. Thus the videotaped data made it possible to analyse the process in which the traces on the sheets were created. One interview took 40-60 minutes depending on its structure planned beforehand and on the progress of the discussion.

In most cases, the interviewee was first asked to explain the visual interpretation of the limiting process in the definition of derivative. (See Task 1 in Appendix 2.) This task was essentially the same as Task 5a in the test. After that the interviewee's views about the relationship between continuity and differentiability were inquired: The interviewer (the author) asked if the interviewee believed all continuous functions to be differentiable, all differentiable functions to be continuous or whether he/she believed that these concepts did not have any dependence between them. In Task 2 the interviewee was asked to explain why a function whose graph has jumps cannot be differentiable. Also in this task it was important to have a visual view about the limiting process of the difference quotient. In this task we considered both a case in which a function has "one jump" at a point, like the function in Task 2b of the test, and a case in which it has "two jumps", like the function in Task 2d. Therefore, the interviewee had to argue his/her answers to these test-questions. Especially with the students whose knowledge about the theme seemed to be poor, we discussed also the other functions in Task 2, the interviewee's test answers concerning their continuity and differentiability and, on the general level, the visual interpretation of continuity and differentiability. We could also take up some additional functions. The main goal of interview tasks 1 and 2 and the discussions described above was to explore how the students understood the visual meaning of continuity and differentiability of a function and how they were able to apply the formal definition of derivative in a visual context and to see connections to the definition in the visual context. Therefore, discussions about Tasks 1 and 2 and the additional discussions were especially connected to research questions 1 and 2 (see Section 4.1), but to some extent they offered useful data also about students' abilities to use informal interpretations and the formal definition in reasoning. Thus they were partially also connected to research questions 3 and 4.

In Tasks 3 and 4, a problem whose solution was asked to be found and argued was given. The interviewees were free to choose their method to approach the problem: They were allowed to use informal or formal reasoning or both in their solution. In some cases, a second solution, which was required to be based on a different type of approach than the first one, was also asked. In Task 3, nothing in the form in which the problem was given referred to an informal or to a formal solution. In Task 4, the definitions of an even and an odd function were given by describing visual interpretations for these concepts. This might contribute to choosing a visual approach at first. However, the problem itself in Task 4 was presented quite formally, and thus it was reasonable to present these definitions in an informal form in order to keep both ways to approach the problem attractive at least to some extent. The main goal of Tasks 3 and 4 was to study the roles of informal and formal reasoning in problem solving processes. Thus these tasks were designed on the basis of research question 5. However, using them at the interviews to some extent offered useful data also about research questions 3 and 4.

4.3 A preliminary study

In order to test the form of the written test and the interview questions, a small-scale preliminary study was arranged before collecting the actual data. Two students who were majoring in mathematics and who were at the final phase of their studies were asked to answer the questions of the test form, and after that they were interviewed. Article A is based on the interviews of this preliminary study. The interview questions and the structure of the interviews were essentially the same as in the actual study.

5 Results of the study

5.1 About students' overall success in the test

The overall results of how the students succeeded in the written test are collectively presented in Appendix 4. They are also reported in Viholainen (2007). Detailed criteria which were used in the evaluation of test answers are presented in Appendix 3. The test results of Tasks 3 and 4 are analysed more deeply in Article E, and the results concerning the interviewees' answers to these tasks are considered also in Article C. The rest of the results supplement the picture about the overall test success, and they also offer interesting data for further studies. Indeed, it has to be noted that valid conclusions regarding students' absolute skills cannot be drawn on the basis of the test results, because it is not known how intensively the students answered to the test. The students were aware that the test results did not have any concrete effect on their future life. Therefore, external reasons which could motivate the students to do their best in the test were quite weak. This is a problem especially in the cases of Tasks 3-6, in which thoroughly constructed arguments were required. So, for these tasks, it was more appropriate to examine the relative success between the tasks than the absolute points, because it could be assumed that the average answering intensity varied less between the tasks than it varied between the subjects. Instead, in Task 2 the effect of the answering intensity was not as significant, because this task consisted of multiple-choice questions. Indeed, one of the most interesting test result concerns Task 2: It turned out that 26% of the Finnish participants thought the function in Task 2d to be discontinuous but differentiable. In Articles B and D, reasons for these kind of erroneous conclusions are proposed.

In the following, the results presented in the included articles are summarized.

5.2 Students' informal and formal arguing skills

Article E is based on a statistical analysis about the results of Tasks 3 and 4 in the test and students' estimates about their amount of passed studies in mathematics and about their study success. As told above, each one of Tasks 3a, 3b, 4a and 4b was graded by using the points 0, 1 and 2. The points obtained from Tasks 3a and 4a, and, correspondingly, points obtained from Tasks 3b and 4b were added together. The variables created in this way tell about students' abilities to argue the claims in Tasks 3 and 4 informally and formally. In Article E, Tasks 3 and 4 are for technical reasons marked with numbers 1 and 2.

The following findings were observed in the study:

- For both tasks, the means of the points obtained from the formal tasks were a little higher than those of the informal tasks. However, the differences were not statistically significant. The difference between the total informal points and the total formal points was not statistically significant either.

- Almost three fourths (74.6 %) of the students obtained at least one point both from Task 3a and from Task 3b. Instead, almost half (48.6 %) of the students obtained zero points both from Task 4a and from Task 4b.

- The correlation between the informal points and the formal points was statistically very significant. On the basis of crosstabulation between the total informal points and the total formal points, it seemed that poor success in the formal tasks implied poor success also in the informal tasks, but, instead, the success in the formal tasks could be good even if the success in the informal tasks was poor.

- Succeeding both in the informal tasks and in the formal tasks appeared to be dependent on the amount of passed studies in mathematics and on the success in these studies. However, some results of the analysis refer that these factors could have a stronger effect on succeeding in the formal tasks than on succeeding in the informal tasks.

- The sample included a subgroup consisting of 17 students whose good success in the informal tasks cannot be explained by a good study success in mathematics. The difference between the numbers of passed credits in this group and the numbers of passed credits in the whole sample was not significant, and thus good success in the informal tasks cannot be explained either by the large amount of passed credits. This suggests that the number of passed credits and the level of the study success are not sufficient explanatory factors for succeeding in the informal tasks.

- The sample included also a subgroup consisting of 28 students whose good success in the formal tasks cannot be explained by a good study success in mathematics. The average number of passed credits was in this group significantly higher than in the whole sample, and thus the large number of passed credits can be considered as a possible explanatory factor for the good success in the formal tasks. This suggests that the amount of passed studies may have an important role in the development of formal proving skills.

The last two observations have been mentioned in Article E, but the details of the analysis which conveyed to them have not been described. However, a statistical analysis of the relationships between the points obtained from Tasks 3 and 4, the number of passed credits in mathematics and the study success is presented in Appendix 6. These two observations are results of this analysis. The details of this analysis were omitted from Article E, because the analysis was quite complicated and describing it would have thus required a lot of space. In addition, the results of this analysis can be considered only suggestive, because the data about the number of passed credits and the study success were based on students' own estimates, and thus it was not very reliable (see Section 3.3 in Article E or Appendix 5).

5.3 Informal and formal reasoning in problem solving

Articles A and C mainly consider the roles of informal and formal reasoning in students' problem solving processes. Students' informal and formal reasoning skills are also analysed. In both of these studies the qualitative paradigm is dominant, even though also quantitative methods have been applied in Article C.

In Article A interviews of two students, Anna and Ben, are described and analysed. These interviews were held at the preliminary phase of the actual study (see Section 4.3). Both interviewees were majoring in mathematics, and they were at the final phase of their studies. They both had succeeded quite well in their studies. The analysis of the interviews led to the following findings:

- Both students seemed to have a very clear conception that derivative visually means steepness of the tangent line. They were also able to use this interpretation successfully in their reasoning. For example, both of them solved the interview task 4 visually by using this interpretation.

- Both students also seemed to understand that the difference quotient visually means the slope of the secant line. In the interview task 1, they managed to interpret visually both the beginning state (the secant state) and the end state (the tangent state) of the limiting process in the definition of derivative. However, when they were asked, by using the definition of derivative, to explain why the function in the test task 2a is not differentiable, both of them seemed to have problems in understanding the dynamics of the limiting process visually. For example, Ben believed the limit of the difference quotient to be equal to the limit of the slope of the graph. (In Task 2a, both parts of the graph of the function were straight lines.)

- A clear difference between the students' tendencies to choose between informal and formal reasoning was found: Anna often chose a visual approach, whereas Ben preferred to use formal reasoning. Another difference between their ways to use informal and formal reasoning was that Anna often utilized informal and formal reasoning simultaneously, whereas Ben kept them more separate. For example, Anna tried to apply visualization when she constructed a formal proof. Instead, Ben usually carried out a reasoning from the beginning to the end either entirely formally or entirely informally. He, certainly, could present both informal and formal arguments for results, but these pieces of reasoning were not connected at least explicitly.

In Article C, students' performances in the interview task 3 have been analysed. It was possible to approach the problem of this task either informally by sketching the graph or formally by using the definition of derivative. It was expected that neither of these alternatives was overattractive. The central goal was to study the students' tendencies to choose between informal and formal methods when they tried to solve this problem. The relationship between these choices and students' potential abilities to consider derivative/differentiability informally and formally was also explored. The conclusions regarding potential abilities were made on the basis of the results of Tasks 3 and 4 of the written test and on the basis of the data about the number of passed credits and the success in studies in mathematics.

Task 3 was considered in 18 out of 21 interviews, and it turned out that these 18 interviews could be divided into four classes, based on the choices between informal and formal methods. The classification formed was the following:

- Three students solved the problem formally. All these solutions were correct, and the authors were also convinced of them. During the solving process the students did not have any serious problems.

- Three students used only informal reasoning in this task. One of them succeeded to solve the problem correctly, but the other two students failed.

- Six students tried an informal reasoning first, but since they did not succeed in it, they changed to formal reasoning. Two of them managed to find the correct solution by using the formal reasoning.

- Six students studied differentiability by erroneous methods. For example, three of them thought that continuity was a sufficient condition for differentiability. The solving processes of these six students were not further classified in this study.

Any other combination of informal and formal reasoning events did not appear. For example, none of the students began formally and then changed to informal reasoning. The above observations also reveal the difficulty of the informal solution with respect to the formal solution in the case of this task: Altogether, nine students attempted to solve the problem informally, but only one of them succeeded in it. Also, nine students attempted a formal solution, and five of these attempts were successful.

The students in the first class had passed a lot of mathematics with good success, and they also did well in the test, both in the informal tasks (Tasks 3a and 4a) and in the formal tasks (Tasks 3b and 4b). Therefore, they probably had a strong potential ability to consider derivative both informally and formally. Instead, the students in the second and the third class had passed fewer courses in mathematics and/or their study success had been weaker. Their success in the informal tasks of the test varied, but five of them had obtained the highest scores from the formal tasks.

Overall, the results of the study presented in Article C can be summarized in the following way: The interview task 3 proved to be more difficult to solve informally than formally. Despite that, many students seemed to have a strong tendency to try on informal solution, even though, according to the results of the written test, many of them would probably have had the potential ability also for a formal solution. Only the students in the first class, who had a wide and successful experience about the mathematics taught at university, started by a formal method.

5.4 Erroneous conclusions and the coherence of a concept image

The discussions at the interviews about continuity and differentiability (see Section 4.2.3) offered interesting data about students' conceptions and reasoning strategies. The result that more than one fourth of the students in the test answered that the function in Task 2d is discontinuous but differentiable (see Appendix 4) raised the question about the reasons for such erroneous conclusions. Therefore, the discussions of the interviews were analysed from this point of view. In Article B, two discussions which seemed to offer the most interesting data for this issue have been analysed. The analysis of the coherence of a concept image proved to be useful in understanding the emergence of erroneous conclusions. Especially, in one discussion the interaction between the insufficiencies in the coherence of the concept image and erroneous conclusions proved to be very clear. The study presented in Article D considers the deep analysis of this interaction.

The interviewees of the study presented in Article B, Mark and Theresa, were majoring in mathematics. Both of them were at the final phase of their studies. According to the self-estimates (see Section 4.2.2 and Appendix 3), Mark's study success had been average and Theresa's study success good. In the test both students had answered that function f in Task 2d is discontinuous but differentiable.

At the beginning of the interview, Mark had the conception that continuity is a necessary condition for differentiability. Also Theresa remembered that there existed some theorem about the relationship between these concepts, but she did not remember exactly what it stated. However, both of them had to change their conceptions, due to conclusions which they made during the interviews: At the end of the interview, Mark concluded that a discontinuous function can be differentiable, whereas Theresa concluded that no theorem exists concerning this relationship. Both of these conclusions are erroneous with respect to the formal theory, and the analysis showed that the incoherence of a concept image was a central factor contributing to their emergence.

During the interviews students were asked to study differentiability of the functions in Tasks 2a, 2b and 2d of the test. Moreover, the following function was used:

$$f_4(x) = \begin{cases} x, & x < 1, \\ x+1, & x \geqslant 1. \end{cases}$$

In their reasoning both students mainly used methods which were not fully compatible with the formal theory. Using the differentiation rules, Mark first differentiated both expressions used in the definition of the piecewise defined function. Then he checked if both expressions for the derivative obtained an equal value at the point in which the expression was changed. Mark seemed to have a strong confidence in this differentiation method. He did not seem to have any doubts about it, and it seemed to work very well for the functions presented in the test (the functions in Tasks 2a, 2b and 2d), and this probably further strengthened his confidence. However, with function f_4, using this method led to a conflict with the conception according to which a differentiable function has to be continuous. In order to resolve the conflict, Mark made use of the definition of derivative, but he applied it incorrectly so that the differentiation method seemed to be compatible with the definition (see Figure 6 in Article D). Also, for the function of Task 2a, the interviewer asked Mark to calculate the left-hand limit of the difference quotient at the point $x = 1$. Mark performed also this calculation erroneously (see Figure 5 in Article D) so that its result was in accordance with the reasoning based on the differentiation method. Finally, trusting the differentiation method, Mark resolved the above-mentioned conflict by changing his view about the relationship between continuity and differentiability.

Theresa studied differentiability on the basis of the graphs: She studied visually whether it was possible to draw a unique tangent line to all points of the graph. However, her conception about the tangent was not compatible with the formal theory: In the case of the function in Task 2d, she drew a tangent line to the point where the jump happened (the point (4,3)), as if there were no hole in the parabola. Thus she concluded that this function was differentiable. As regards the function of Task 2a and function f_4, Theresa used also the same differentiation method as Mark. In the latter case, the tangent method and the differentiation method led to contradictory results. However, in this conflict, Theresa trusted the tangent method without any doubts.

In addition to the fact that the definition was applied incorrectly, an essential problem seemed to be that the definition had a minor role in the reasoning. The students hardly ever used the definition unless the interviewer explicitly asked them to do it. Thus their arguments did not have proper connections to the formal theory. The strong confidence in the methods used could be one important reason why the students did not see a need to use the definition. By trusting these methods, the students built structures which were quite well internally coherent, but which were in contradiction with the formal theory. It also turned out that the students thought it was difficult to use the definition, and this was another probable reason for their tendencies to avoid using it.

6 Discussion

The usable data for this thesis were very large, and only some of it have been employed in the studies included in this thesis. The data offer potential for several additional studies regarding students' informal and formal reasoning or the coherence of the concept image. In future, I may analyse the data further. However, this doctoral thesis is restricted to the results reviewed above, and in this section I will discuss the potential consequences of these results.

6.1 The reluctance to use the definition

Students' reluctance to use the definition was a feature which was revealed in several connections in this study. In Articles B and D it was described how Mark and Theresa used methods which were not based on the definition and how the use of these methods led them to build structures which were in contradiction with the formal theory. Also in the study presented in Article C, the reluctance to use the definition of derivative could be seen in several cases as a significant reason for the problems encountered: Most of the students chose an informal approach in a task in which, in fact, the formal solution based on the use of the definition would have been easier. As well, in Article A, Anna seemed to have a tendency to rely on visual argumentation instead of formal reasoning, and in Task 3, the inexact visual reasoning led her to an erroneous conclusion.

It came out in several connections that even though the students were reluctant to use the definition of derivative, they were able to do it, if it was compulsory. In the written test in Task 3a, almost 80% of the subjects managed to use the definition of derivative at least for the most part correctly, and, as shown in Article E, the average success in the tasks requiring formal reasoning based on the definition cannot be considered to be at any rate lower than the average success in the tasks where informal reasoning was required. Besides, most of the students who chose the informal approach in Article C had succeeded well in the formal tasks of the written test where the explicit use of the definition was required. In Article B, Theresa managed to perform the formal calculation based on the definition when the interviewer asked her to do that. Also Mark mastered the main principle of the use of the definition. Therefore, poor skills to use the definition are not a sufficient explanation for the observed reluctance.

Previously, Vinner (1991) and Pinto (1998), among others, have reported about students' tendencies to avoid the use of definitions. Juter (2005) and Tsamir et al. (2006) have also reported on cases where university students in problem solving used arguments and methods which were not connected to theoretical knowledge about mathematical concepts. Such reasoning led to erroneous conclusions. However, in other connections it had turned out that the students possessed all the necessary theoretical knowledge. The studies by Pinto and Juter concern the concept of limit and those by Vinner and Tsamir et al. as well as my study concern the concept of derivative. Therefore, it seems that especially with these concepts the reluctance to use the definition in problem solving is a widely observed problem.

Neither this study nor the studies reviewed above directly reveal reasons for the observed reluctance to use definitions. However, something can be speculated. At first, affective factors probably had a significant effect on choosing the methods. The definitions could be thought to be troublesome or uncomfortable to use. Some references to this were found in the data: For example, in Article B, Theresa did not first rely on her abilities to handle the definition of derivative when the interviewer asked her to apply it. Also with the limit concept, the definition in the "ϵ-δ form" is often seen difficult. In addition, it has to be noted that in mathematics at upper secondary school, the definitions of such concepts as limit, continuity, derivative, and so on, have a minor role: For example, tasks at upper secondary school concerning these concepts do not require the explicit use of the definitions. This may encourage students to build their concept image of these concepts on the basis of other than formal representations, and this, furthermore, may influence their choices of methods still at the final phase of their university studies.

6.2 Neither informal nor formal reasoning must be ignored

As mentioned in Section 2.5, many researchers have been concerned about students' tendency to avoid informal reasoning, especially visualization. As well, Weber and Alcock (2004) describe cases in which students' reasoning was too much restricted around the definition. These reasonings concerned the concepts of isomorphism in group theory and convergence of a sequence. Also Vinner (1989) found that college students avoided visualization in a task concerning the mean value theorem, which was yet easier to solve visually than formally. These observations seem to be opposite to the results of this study and the results reviewed in the previous section, according to which students' reasoning is not sufficiently well connected to the formal theory. However, both kind of observations prove that students' reasoning is often too confined. Weber and Alcock consider syntactic reasoning based on the symbolic manipulation of formal statements more confined than semantic reasoning, in which the problem is seen in a more holistic way by using different kind of informal interpretations. However, reasoning may also be biassed so that the formal theory is ignored. In this study, several examples of that came out. Several reasonings based on the visual method in Article C, Theresa's reasoning in Article B and Anna's reasoning in Task 3 in Article A are examples of cases where reasoning is too much restricted to visual representations. Mark's reasoning in Articles B and D, for one, can be considered as an algorithmic reasoning which was not sufficiently well connected to the formal theory.

Therefore, it seems that, in general, students' reasoning is usually confined to one representational system, for example, to the formal, visual or algorithmic system. Even though the students know several systems, they find it difficult to change from one system to another in a problem solving situation. This difficulty is parallel to the difficulties in revising existing knowledge structures, found in studies about the *conceptual change*. According to theories about the conceptual change, the cognitive knowledge structure can be constructed either by *enriching* or by *revising* the existing structure (Merenluoto, 2001; Merenluoto and Lehtinen, 2004; Vosniadou, 1994; 2006). Enrichment is the easier way: It means that new information is added to what is already known. This requires that the new information is consistent with the existing structures. Instead, if an individual is not able to assimilate the new information to the existing knowledge, the preservation of the coherence of the knowledge structure requires that the existing structures have to be revised or the way of thinking radically changed. This is often very difficult. It may be that the reasoning inside one representational system does not require revising the knowledge structure, but when the representational system is changed, the whole problem field has to be constructed in a new way. However, good connections between different representational systems (see Section 3.2) make this reconstruction and flexible changes between representational systems easier. This makes it possible to use several representational systems simultaneously, so that reasoning based on one representational system supports and controls reasoning based on another system.

It is also important to remember that the same way of thinking is necessarily not optimal for all students. Pinto and Tall (1999; 2001; 2002) and Pinto (1998) observed two different thinking styles which were used by mathematics students in the first year analysis course at university. *Formal thinkers* attempted to base their reasoning on the formal theory and to extract meaning for the concepts on the basis of their formal definitions. Instead, *natural thinkers* attempted to give meaning to the formal theory by using their existing imagery. According to Pinto and Tall, both thinking styles may be successful or unsuccessful. Also in this study, in several connections in Articles A, B and C, it came out that some students preferred to use informal methods, whereas others rather chose a formal approach. However, as shown in Chapter 2, both informal and formal elements have a very important role in mathematical reasoning. Thus, everyone trying to succeed in mathematics studies cannot totally bypass either of these elements, but their relative importance may differ between individuals. In teaching mathematics it is important to take the different thinking styles into account. At least in the tertiary-level courses, the central role of the formal theory is usually unavoidable, but the teacher should also be aware that for many students an introduction to informal interpretations may have a significant effect on the success in studies.

6.3 The development of the coherence of the concept image

The case studies presented in Articles B and D offer examples of cases where incoherent structures were constructed. In these cases, the students initially had some erroneous conceptions, and the structures constructed were based on them. Because it is hard to reconstruct large structures, the development of the incoherent structures should be prevented in time. Therefore, teachers should be aware of students' conceptions, and it would also be important that in their studies students would meet with conflicts which would lead them to reflect their conceptions critically.

In addition to the absence of contradictions, the coherence of a concept image indicates that connections between different parts of the concept image are flexible. In fact, flexible connections can be considered as a key factor in preventing the emergence of contradictions, in revealing the existing unconscious conflicts and in solving them. Moreover, the flexible connections help students choose the most suitable methods in a problem solving situation. But how could the concept image be developed so that its elements would become connected to each other in an optimal way ? One option for this could be definition-centred thinking. It means that the definition or, in fact, its personal interpretation becomes the most essential element in the concept image, and all other elements, such as informal interpretations or calculation methods concerning the concept, are understood as consequences of the definition. In that case, the definition is the cohesive element of the concept image, and connections between the other elements of the concept image are formed through it. However, the support of definition-centred thinking in teaching is often problematic, because, at least below the tertiary level, it is usually pedagogically inappropriate to begin the learning of a new concept from the definition. For example, according to Tall's theory of the three worlds of mathematics (see Section 2.3), the learning of mathematical knowledge should begin from perceptions and actions. In the learning process, the qualities of objects are explored through them, and the knowledge is compressed through generalization and abstraction. The attainment of the formal-axiomatic world is regarded as a punchline of this process. If the teaching of mathematics is designed on the basis of this theory, formal definitions are introduced before the students see a real need for them. Correspondingly, in the research of mathematics, the definitions are not an end as such, but they are only tools for compressing the knowledge. However, when a definition is in use and fixed, it absolutely defines what is true and what is not true about the concept. Both these roles of definitions should be brought out also in teaching.

According to the *realistic mathematics education theory* (Freudenthal, 1991; Gravemeijer, 1999), the students should be given an opportunity to reinvent mathematics. The central idea is "...to allow learners to come to regard the knowledge they acquire as their own private knowledge, knowledge for which they themselves are responsible" (Gravemeijer, 1999; p. 158). Naturally, the reinvention occurs under instruction. In this theory, mathematics is not primarily regarded as a structured knowledge but as a human activity, and thus definitions and theorems are not considered as important as the ideas which have contributed to their construction. If the role of definitions is aimed to be emphasized in the way presented above, the principles of this theory could offer one baseline for the teaching of mathematics. Before introducing the definition of a concept, students should be introduced to a situation in which they see a real need to compress the knowledge by bringing a new definition into use. However, all individuals do not see the need to compress the knowledge in the same way, but the definitions which should be learnt are usually universal. This point, among others, makes the design of teaching based on the realistic mathematics education theory challenging.

6.4 Participants of the study as prospective teachers

Above, some views have been presented on how the findings of this study together with those of some previous studies concerning the same issues could increase our knowledge about the learning of mathematics and about mathematical thinking. In addition, some views based on these findings as to how the teaching of mathematics could be improved have been proposed. In this discussion, the subjects of the study have been considered as mathematics students whose skills have been tested and whose thinking has been analysed. But what does it mean that most of the students, at the moment when they participated in the study, were prospective teachers at the middle or at the final phase of their studies?

All the tasks used in the written test and at the interviews concerned real valued functions of a single variable. No peculiar tricks were needed in solving the tasks, but most of the tasks could be solved by basic skills applying the formal definition of derivative and interpreting derivative informally. In fact, a thorough understanding of the knowledge included in the curriculum of mathematics of the upper secondary school should be enough to solve all the tasks of the study, possibly excluding Task 6 in the test where knowledge of the density of rational and irrational numbers was needed. Despite this, the study revealed that many students had serious problems in solving these tasks. Even though it has to be taken into account that the students did not prepare in any way for the test or for the interviews and that the intensity of answering in the test is not known, the findings of the study raise concern about the subject matter knowledge of prospective mathematics teachers. In this respect, the observed reluctance to use the definition is especially worrying, because, as shown above, it is totally against the fundamental nature of mathematics to ignore the definitions of the concepts which are used. Furthermore, this raises a question: Have the (prospective) mathematics teachers really internalized the real nature of mathematics? For example, do they really understand the fundamental role of definitions in mathematics? The findings of this study suggest that foundations of mathematical knowledge are an issue which should be emphasized more in teaching of mathematics at university.

6.5 Practical and theoretical relevance of the study

All research of mathematics education, as well all educational studies, should have *practical* or *theoretical relevance*. That is, it should have "some positive impact on the practice of teaching" or it should "broaden or deepen our understanding of the teaching and learning phenomena" (Sierpinska, 1993; p. 38). As regards this classification, Kilpatrick (1993) divides research studies into three categories. 1) Some studies attempt to have a direct influence on teaching practices "by providing ideas and material for teachers to use and by suggesting activities they might conduct" (p. 20). For example, studies based on the principles of *design-based research* (Kelly, 2003) can be considered as such studies. 2) On the other hand, some studies suggest new ways to understand students' thinking and events in the classroom. They may have an indirect influence on teaching practices. 3) In addition, there are studies which attempt to develop the terms and the framework in which mathematics education is portrayed in publications. These studies usually have direct importance only for researchers in the discipline. In practice, most of the studies in the area of mathematics education usually have both practical and theoretical relevance at least to some extent.

It has to be noted that in this thesis the learning process in a broad sense was not studied, but almost all the data were based on observations of students' performances in the written test and at the interviews. This kind of research may reveal what skills students have and what kind of manners they tend to use in reasoning, but it does not reveal how these skills or manners have been developed. Each of the studies described in the included articles brings out features of students' reasoning which should be taken into account in the teaching of mathematics at upper secondary school and at university. These features are related to erroneous conclusions, students' skills and tendencies to apply informal and formal methods, and so on. In the studies, these features were revealed and their importance was shown by using, in addition to the collected data, the literature and the constructed theoretical framework presented in Sections 2 and 3. However, on the basis of the usable data, it is impossible to show what teaching practices could be effective with respect to them. In this thesis, some speculative ideas concerning this question and reasons for the observed features were presented, but further studies, in which the learning process is taken into account, would be needed in obtaining reasonable research-based views.

An important question concerning the practical relevance of this study is the question of the generalizability of the observed features in students' reasoning. Are these features so general that taking them into account in designing teaching practices, curricula and teaching materials is really worthwhile? Are they typical of Finnish students only, or are they shared by all students around the world? As stated in Section 4.2.1, the group of the participants in the written test cannot be considered as a representative sample of all Finnish subject teacher students in mathematics. However, at the moment when the test was arranged, the participants of the test made up a large portion of all Finnish subject teacher students in mathematics and, therefore, also a large portion of all Finnish university students in mathematics. Thus, it is very reasonable to take the findings concerning the whole sample into account when designing teaching of mathematics at university, at least in Finland.

The practical value of the case studies presented in Articles A, B, C and D, is that they make the designers of teaching aware of the existence and the importance of the observed features in students' reasoning. However, like case studies in general, they show nothing about the generalizability of these features, Therefore, the designers of teaching are responsible to decide as to to what extent and how the findings of these studies are worth being taken into account in each situation.

In the present thesis, I have generated the conceptual framework relating to informal and formal argumentation and the coherence of the concept image. These have been described in Sections 2 and 3. I consider this as the most important theoretical achievement of this thesis. When I worked with this thesis, I often met a need to contemplate over and over again questions like: What are the informal and formal sides of mathematics, what kind of arguments are informal or formal, what kind of a structure does the concept image include, what is the relationship between the concept image and the formal theory, what kind of a concept image is coherent, and so on. From the literature I found plenty of interesting material about these questions, but I did not find any complete framework which would have been suitable for my study. Thereby, on the basis of the literature about these issues in mathematics education, as well the history and philosophy of mathematics, I established the definitions of informal and formal arguments (Section 2.7) and the framework relating to the structure and the coherence of the concept image (Section 3). The development of both of these has been a long process, which can be observed also from the included articles: In Articles A and C, I have presented some views on the informal and formal sides of mathematics and on the informal and formal arguments, but the presentation is not as systematic as in Article E or in Section 2 of this thesis, which have been written later. In Article B, I have also dealt with issues concerning the coherence of the concept image, but in Article D and in Section 3 in this thesis, this framework is much more systematic. I believe that both the established definitions for informal and formal arguments and the framework relating to the concept image are usable also for further studies of mathematical reasoning, especially for studies about distinctions of the different sides of mathematics. It would be interesting to analyse the learning of mathematics on their basis. For example, it would be interesting to study how the understanding of informal and formal sides of mathematics is developed and how the connections between different sides of mathematics could be learnt. These questions could be examined either generally or by concentrating on some given concept. In addition to the tertiary level, similar questions could be studied at all levels of mathematics education. However, this might require further modification of the framework.

References

Arcavi, A.: 2003, 'The role of visual representations in the learning of mathematics', *Educational Studies in Mathematics*, 52, 215-241.

Asiala, M., Brown, A., DeVries, D. J., Dubinsky, E., Mathews, D. and Thomas, K.: 1996, 'A framework for research and curriculum development in undergraduate mathematics education', in Kaput, J., Shoenfeld, A. H. and Dubinsky, E. (Eds.), *Research in Collegiate Mathematics Education II, Issues in Mathematics Education (CBMS)*, American Mathematical Society, pp. 1-32.

Aspinwall, L., Shaw, K. L. and Presmeg, N. C.: 1997, 'Uncontrollable mental imagery: graphical connections between a function and its derivative', *Educational Studies in Mathematics*, 33, 301-317.

Bingolbali, E.: 2005, *Engineering and mathematics students' conceptual development of the derivative: an institutional perspective*, Unpublished PhD thesis, University of Leeds, UK.

Bingolbali, E., Monaghan, J. and Roper, T.: 2006, 'Service teaching: mathematical education of students of client departments', in Novotná, J., Moraová, H., Krátká, M. and Stehliková, N. (Eds.), *Proceedings of the 30th Conference of the international Group for the Psychology of Mathematics Education*, Prague, Czech Republic, Vol. 2, pp. 169-176.

Clements, K.: 1982, 'Visual imagery and school mathematics (concluded)', *For the Learning of Mathematics*, 2(3), 33-39.

Creswell, J. W. and Plano Clark, V. L.: 2007, *Designing and conducting mixed method research*, Sage publications, USA.

Davis, P. J. and Hersh, R.: 1981, *The Mathematical Experience*, Viking Penguin Inc., New York.

Dreyfus, T.: 1994, 'Imagery and reasoning in mathematics and mathematics education', in Robitaille, D. F., Wheeler, D. H. and Kieran, C. (Eds.), *Selected lectures from the 7th International Congress on Mathematical Education*, Les Presses d l'Université Laval, Quebec, Canada, pp. 107-122.

Dubinsky, E. and McDonald, M. A.: 2001, 'APOS: a constructivist theory of learning in undergraduate mathematics education research', in Holton, D. (Ed.), *The teaching and learning of mathematics at university level: an ICMI study*, Kluwer Academic Publishers, pp. 273-280.

Eisenberg, T. and Dreyfus, T.: 1991, 'On the reluctance to visualize in mathematics', in Zimmermann, W. and Cunningham, S. (Eds.), *Visualization in teaching and learning mathematics*, MAA notes number 19, Mathematical Association of America, Washington, DC, pp. 25-37.

Fischbein, E.: 1994, 'The interaction between the formal, the algorithmic and the intuitive components in a mathematical activity', in Biehler, R., Scholz, R. W., Strässer, R. and Winkelmann, B. (Eds.), *Didactics of mathematics as a scientific discipline*, Kluwer, Dordrecht, The Netherlands, pp. 231-245.

Fischbein, E.: 2001, 'Tacit models and infinity', *Educational Studies in Mathematics*, 48, 309-329.

Freudenthal, H.: 1991, *Revisiting mathematics education*, Kluwer, Dordrecht, The Netherlands.

Goldin, G. A.: 1998, 'Representational systems, learning and problem solving in mathematics', *Journal of Mathematical Behavior*, 17(2), 137-165.

Goldin, G. A. and Kaput, J. J.: 1996, 'A joint perspective on the idea of representation in learning and doing mathematics', in Steffe, L. and Nesher, P. (Eds.), *Theories of Mathematical Learning*, Mahwah, NJ: Lawrence Erlbaum, pp. 397-430.

Gravemeijer, K.: 1999, 'How emergent models may foster the constitution of formal mathematics', *Mathematical Thinking and Learning*, 1(2), 155-177.

Gray, E. and Tall, D.: 1994, 'Duality, ambiguity and flexibility: a proceptual view of simple arithmetic', *The Journal for Research in Mathematics Education*, 26(2), 115-141.

Gray, E. and Tall, D.: 2001, 'Relationships between embodied objects and symbolic procepts: an explanatory theory of success and failure in mathematics', in van den Heuvel-Panhuizen, M. (Ed.), *Proceedings of The 25th conference of the international group for the psychology of mathematics education*, Utrecht, The Netherlands, Vol. 3, pp. 65-72.

Gutiérrez, A.: 1996, 'Visualization in 3-dimensional geometry: in search of a framework', in Puig, L. and Gutiérrez, A. (Eds.), *Proceedings of The 20th conference of the international group for the psychology of mathematics education*, University of Valencia, Spain, Vol. 1, pp. 3-19.

Haapasalo, L. and Kadijevich, D.: 2000, 'Two types of mathematical knowledge and their relation', *Journal für Mathematik-Didaktik*, 21(2), 139-157.

Hanna, G.: 1990, 'Some pedagogical aspects of proof', *Interchange*, 21(1), 6-13.

Hiebert, J. and Levevre, P.: 1986, 'Conceptual and procedural knowledge in mathematics: an introductory analysis', in Hiebert, J. (Ed.), *Conceptual and procedural knowledge: the case of mathematics*, Hillsdale, NJ: Lawrence Erlbaum, pp. 1-27.

Hähkiöniemi, M.: 2006a, 'Perceiving the derivative: the case of Susanna', *Nordic Studies in Mathematics Education*, 11(1), 51-73.

Hähkiöniemi, M.: 2006b, *The role of representations in learning the derivative*, Report 104, Department of mathematics and statistics, University of Jyväskylä, Finland.

Iannone, P. and Nardi, E.: 2007, 'The interplay between syntactic and semantic knowledge in proof production: mathematicians' perspectives', in Pitta-Pantazi, D. and Philippou, G. (Eds.): *Proceedings of The Fifth Congress of the European Society for Research in Mathematics Education*, Department of Education, University of Cyprus, pp. 2300-2309.

Inglish, M., Maija-Ramos, J. P. and Simpson, A.: 2007, 'Modelling mathematical argumentation: the importance of qualification', *Educational Studies in Mathematics*, 66(1), 3-21.

Juter, K.: 2005, 'Limits of functions: traces of students' concept images', *Nordic Studies in Mathematics Education*, 3-4, 65-82.

Kelly, A. (Ed.): 2003, 'Theme issue: the role of design in educational research', *Educational Researcher*, 32.

Kilpatrick, J.: 1993, 'Beyond face value: assessing research in mathematics education', in Nissen, G. and Blomhøj, M. (Eds.), *Criteria for scientific quality and relevance in the didactics of mathematics*, Danish Research Council for the Humanities, The initiative: "Mathematics Teaching and Democracy", Roskilde university, pp. 15-34.

Kline, M.: 1972, *Mathematical thought from ancient to modern times*, Oxford University Press, New York, USA.

Krutetskii, V. A.: 1976, *The psychology of mathematical abilities in schoolchildren*, University of Chicago Press, Chicago.

Kvale, S.: 1996, *Interviews: an introduction to qualitative research interviewing*, Sage publications, USA.

Merenluoto, K.: 2001, *Lukiolaisen reaaliluku. Lukualueen laajentaminen käsitteellisenä muutoksena matematiikassa* ('Students' real number. Enlargement of the number concept as a conceptual change in mathematics'), Doctoral dissertation, Series C, Vol. 176, University of Turku.

Merenluoto, K. and Lehtinen, E.: 2004, 'Number concept and conceptual change: towards a systemic model of the processes of change', *Learning and Instruction*, 14, 519-534.

Pinto, M. M. F.: 1998, *Students' understanding of real analysis*, Unpublished doctoral thesis, University of Warwick, UK.

Pinto, M. M. F. and Tall, D.: 1999, 'Student constructios of formal theory: giving and extracting meaning', in Zaslavsky, O. (Ed.), *Proceedings of The 23th conference of the international group for the psychology of mathematics education*, Haifa, Israel, Vol. 4, pp. 65-73.

Pinto, M. M. F. and Tall, D.: 2001, 'Following students' development in a traditional university analysis course', in van den Heuvel-Panhuizen, M. (Ed.), *Proceedings of The 25th conference of the international group for the psychology of mathematics education*, Utrecht, The Netherlands, Vol. 4, pp. 57-64.

Pinto, M. M. F. and Tall, D.: 2002, 'Building formal mathematics on visual imagery: a case study and a theory', *For the Learning of Mathematics*, 22(1), 2-10.

Presmeg, N. C.: 1985, *The role of visually mediated processes in high school mathematics: a classroom investigation*, Unpublished Doctoral thesis, University of Cambridge.

Presmeg, N. C.: 1986a, 'Visualisation and mathematical giftedness', *Educational Studies in Mathematics*, 17, 297-311.

Presmeg, N. C.: 1986b, 'Visualization in high school mathematics', *For the Learning of Mathematics*, 6(3), 42-46.

Presmeg, N. C.: 1992, 'Prototypes, metaphors, metonymes and imaginative rationality in high school mathematics', *Educational Studies in Mathematics*, 23, 595-610.

Presmeg, N. C.: 1997, 'Generalization using imagery in mathematics', in English, L. D. (Ed.), *Mathematical reasoning: analogies, metaphors and images*, Mahwah, New Jersey, pp. 299-312.

Presmeg, N. C.: 2006a, 'A semiotic view of the role of imagery and inscriptions in mathematics teaching and learning', in Novotná, J., Moraová, H., Krátká, M. and Stehliková, N. (Eds.), *Proceedings of the 30th Conference of the international Group for the Psychology of Mathematics Education*, Prague, Czech Republic, Vol. 1, pp. 19-34.

Presmeg, N. C.: 2006b, 'Research on visualization in learning and teaching mathematics: emergence from psychology', in Gutiérrez, A., and Boero, P. (Eds.), *Handbook of research on the psychology of mathematics education*, Rotterdam, Sense Publishers, pp. 205-236.

Raman, M.: 2002, *Proof and justification in collegiate calculus*, Doctoral thesis, University of California, Berkeley.

Raman, M.: 2003, 'Key ideas: what are they and how can they help us understand how people view proof?' *Educational Studies in Mathematics*, 52, 319-325.

Rodd, M. M.: 2000, 'On mathematical warrants: proof does not always warrant, and a warrant may be other than a proof', *Mathematical thinking and learning*, 2(3), 221-244.

Rösken, B. and Rolka, K.: 2007, 'Integrating intuition: the role of concept image and concept definition for students' learning of integral calculus', *The Montana Mathematics education*, Monograph 3, 181-204.

Sfard, A.: 1991, 'On the dual nature of mathematical conceptions: reflections on processes and objects as different sides of the same coin', *Educational Studies in Mathematics*, 22(1), 1-36.

Sibley, T. Q.: 1998, *The geometric viewpoint -A survey of geometries*, Addison Wesley Longman, USA.

Sierpinska, A.: 1993, 'Criteria for scientific quality and relevance in the didactics of mathematics', in Nissen, G. and Blomhøj, M. (Eds.), *Criteria for scientific quality and relevance in the didactics of mathematics*, Danish Research Council for the Humanities, The initiative: "Mathematics Teaching and Democracy", Roskilde university, pp. 35-74.

Skemp, R. R.: 2006, 'Relational understanding and instrumental understanding', *Mathematics Teaching in the Middle School*, 12(2), 88-95, reprinted version, original version published in 1976, *Mathematics Teaching*, 77, 20-26.

Stylianou, D. A.: 2002, 'On the interaction of visualization and analysis: the negotiation of a visual representation in expert problem solving', *Journal of Mathematical Behavior*, 21, 303-317.

Stylianou, D. A. and Dubinsky, E.: 1999, 'Determining linearity: the use of visual imagery in problem solving', in Hitt, F. and Santos, M. (Eds.), *Proceedings of The Twenty-first Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, Columbus, pp. 245-252.

Stylianou, D. A. and Silver, E. A.: 2004, 'The role of visual representations in advanced mathematical problem solving: an examination of expert-novice similarities and differences', *Mathematical Thinking and Learning*, 6(4), 353-387.

Tall, D.: 1995, 'Mathematical growth in elementary and advanced mathematical thinking', in Meira, L., and Carraher, D. (Eds.), *Proceedings of The 19th conference of the international group for the psychology of mathematics education*, Refice, Brazil, Vol. 1, pp. 61-75.

Tall, D.: 2003a, 'Concept image and concept definition', http://www.warwick.ac.uk/staff/David.Tall/themes/concept-image.html, retrieved 5.12.2007.

Tall, D.: 2003b, 'Using technology to support an embodied approach to learning concepts in mathematics', in Carvalho, L. and Guimarães, L., *História e tecnologia no ensino da matemática, Rio de Janeiro*, Vol. 1, pp. 1-28.

Tall, D.: 2004, 'Thinking through three worlds of mathematics', in Gates, P. (Ed.), *Proceedings of the Annual Meeting of the International Group for the Psychology of Mathematics Education (PME) (28th, Bergen, Norway, July 14-18, 2004), Vol. 4, pp. 281-288.*

Tall, D.: 2005, 'A theory of mathematical growth through embodiment, symbolism and proof', *Plenary Lecture for the International Colloquium on Mathematical Learning from Early Childhood to Adulthood*, Belgium, 5-7 July 2005.

Tall, D. and Vinner, S.: 1981, 'Concept image and concept definition in mathematics with particular reference to limits and continuity', *Educational Studies in Mathematics*, 12, 151-169.

Toulmin, S. E.: 2003, *The uses of argument*, Cambridge University Press New York, Cambridge, U. K.

Tsamir, P., Rasslan, S. and Dreyfus, T.: 2006, 'Prospective teachers' reactions to Right-or-Wrong tasks: the case of derivatives of absolute value functions', *Journal of mathematical Behavior*, 25, 240-251.

Viholainen, A.: 2007, 'Suomalaisten matematiikan aineenopettajaopiskelijoiden derivaattakäsitteen informaali ja formaali ymmärtäminen', *Solmu - Fysiikan ja matematiikan aikakauslehti* 4/2007, 16-19.

Vinner, S.: 1982, 'Conflicts between definitions and intuitions: the case of the tangent', in Vermandel, A. (Ed.), *Proceedings of the 6th Annual Conference of the International Group for the Psychology of Mathematics Education*, Antwerpen, Belgium, pp. 24-28.

Vinner, S.: 1989, 'The avoidance of visual considerations in calculus students', *Focus on Learning Problems in Mathematics*, 11(2), 149-156.

Vinner, S.: 1991, 'The role of definitions in teaching and learning', in Tall, D. (Ed.), *Advanced mathematical thinking*, Kluwer, Dordrecht, The Netherlands, pp. 65-81.

Vinner, S. and Dreyfus, T.: 1989, 'Images and definitions for the concept of function', *Journal for Research in Mathematics Education*, 20(4), 356-366.

Vosniadou, S.: 1994, 'Capturing and modelling the process of conceptual change', *Learning and Instruction*, 4, 45-69.

Vosniadou, S.: 2006, 'The conceptual change approach in the learning and teaching of mathematics: an introduction', in Novotná, J., Moraová, H., Krátká, M. and Stehliková, N. (Eds.), *Proceedings of the 30th Conference of the international Group for the Psychology of Mathematics Education*, Prague, Czech Republic, Vol. 1, pp. 155-159.

Weber, K.: 2001, 'Student difficulty in constructing proofs: the need for strategic knowledge', *Educational Studies in Mathematics*, 48, 101-119.

Weber, K. and Alcock, L.: 2004, 'Semantic and syntactic proof productions', *Educational Studies in Mathematics*, 56, 209-234.

Zazkis, R., Dubinsky, E. and Dautermann, J.: 1996, 'Coordinating visual and analytic strategies: a study of students' understanding of the group D_4', *Journal of Research in Mathematics Education*, 24(4), 435-457.

Zimmermann, W. and Cunningham, S.: 1991, 'Editors introduction: what is mathematical visualization', in Zimmermann, W. and Cunningham, S. (Eds.), *Visualization in teaching and learning mathematics*, MAA notes number 19, Mathematical Association of America, Washington, DC, pp. 1-8.

APPENDIX 1: English translation of the form used in the written test

A written test concerning derivative

Dear student,

This test form is connected to a study concerning argumentation skills of subject teacher students in mathematics. The study is carried out at the Department of Mathematics and Statistics of the University of Jyväskylä. The data of the study are collected from several universities in Finland. Your answers will be considered confidentially, and single answers will not be given to any use which is not connected to the study.

Antti Viholainen
Department of Mathematics and Statistics
University of Jyväskylä

Background information

Name:

Main subject:

Secondary subjects:

Number of passed credits in mathematics:

Beginning year of the studies at university:

Describe shortly your success in studies in mathematics at university:

What kind of work (teaching at lower secondary school, teaching at upper secondary school, research, etc.) would you like to do in the future?

Would you like to receive feedback by email about your performance in this test? If you do, please give your email-address.

Definitions

Continuity

A function $f : \mathbb{R} \to \mathbb{R}$ is *continuous at a point* $x_0 \in \mathbb{R}$ if and only if the limit $\lim_{x\to x_0} f(x)$ exists and

$$\lim_{x\to x_0} f(x) = f(x_0).$$

A function is *continuous* if and only if it is continuous at all points of the domain of the function.

Derivative and differentiability

A function $f : \mathbb{R} \to \mathbb{R}$ has a *derivative* or it is *differentiable at a point* $x_0 \in \mathbb{R}$ if and only if the limit

$$\lim_{h\to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. Then the derivative of function f at the point x_0 is equal to the value of this limit.

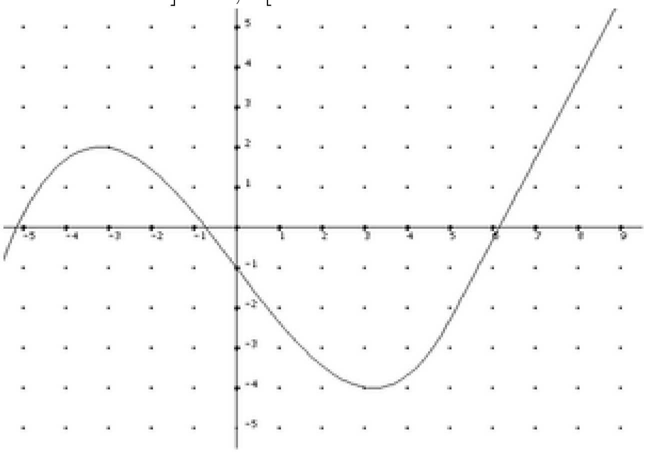
A function is *differentiable* if and only if it is differentiable at all points of the domain of the function.

Tasks

Take into account that the answering time is limited. Try to answer in some way to each question. In Tasks 1 and 2, you can write your answer on the test form; for the other tasks, please use an extra sheet.

If you feel that you are not able to solve the task properly, please describe your problems briefly: Would you need more time to solve these tasks? Do you have problems with perceiving the situation in the task? Do you have problems with computations? Or, do you have some other problems with inventing a solution? What kind of additional knowledge would you need?

1. The diagram below presents a graph of a differentiable function f defined on the interval $]-6, 9[$.

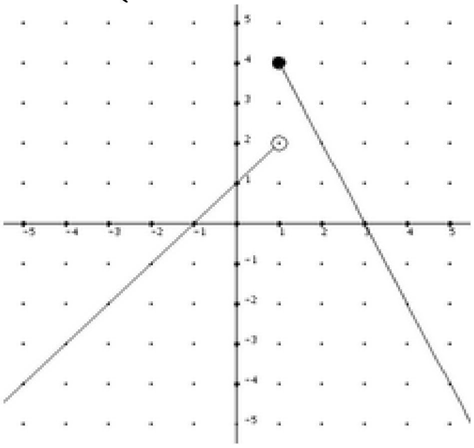


   a) On the basis of the graph, determine the value of the derivative of f when $x = -2$.

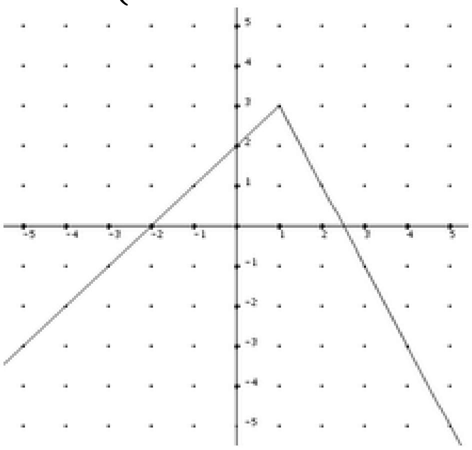
   b) Sketch the graph of the derivative function of f.

2. Which of the following functions are continuous and which are differentiable? "Yes/No" -answers are sufficient.

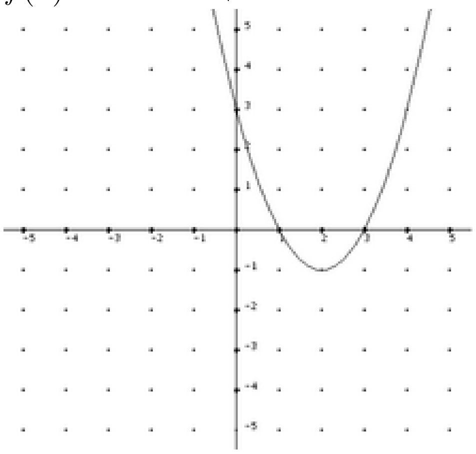
a) $f(x) = \begin{cases} x+1, & x < 1, \\ -2x+6, & x \geqslant 1. \end{cases}$



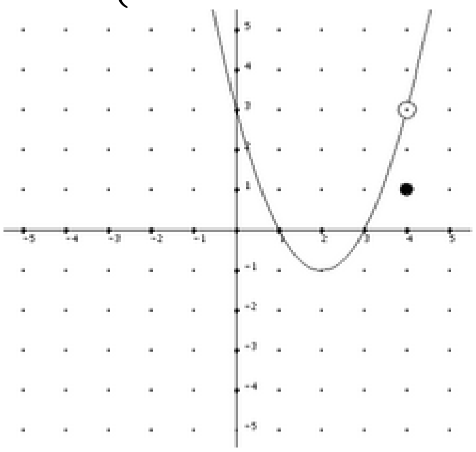
b) $f(x) = \begin{cases} x+2, & x < 1, \\ -2x+5, & x \geqslant 1. \end{cases}$



c) $f(x) = x^2 - 4x + 3$



d) $f(x) = \begin{cases} x^2 - 4x + 3, & x \neq 4, \\ 1, & x = 4. \end{cases}$



3. a) How would you explain, by using graphical interpretations, why the derivative of a constant function is equal to zero everywhere?

   b) Prove the same by using the formal definition of derivative.

4. Claim: Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function and $x_0 \in \mathbb{R}$. Then

$$\lim_{h \to 0} \frac{f(x_0) - f(x_0 - 3h)}{3h} = f'(x_0).$$

   a) How would you visually argue this claim by using a diagram?

   b) Prove the claim formally by using the definition.

5. a) How would you explain, on the basis of the definition of derivative, why the derivative of a differentiable function f at a given point x_0 is equal to the slope of the tangent line drawn to the graph of the function at the point $(x_0, f(x_0))$?

   b) How would you explain, on the basis of the definition of derivative, why the graph of a differentiable function cannot have a corner?

6. Let $f : \mathbb{R} \to \mathbb{R}$:

$$f(x) = \begin{cases} x^2 + 1, & x \in \mathbb{Q}, \\ 1, & x \notin \mathbb{Q}. \end{cases}$$

   Sketch the graph of function f and determine the points where f is differentiable. Give exact arguments.

   Remark: $\mathbb{Q}$ is the set of rational numbers.

Thank you!

APPENDIX 2: Tasks used at the interviews

1. What does the limiting process in the definition of derivative mean graphically?

2. Why cannot a function having jumps be differentiable?

3. Let $f : \mathbb{R} \to \mathbb{R}$:
$$f(x) = \begin{cases} x^4 \cos(1/x^3), & x \neq 0, \\ 0, & x = 0. \end{cases}$$
   Is function f differentiable?

4. Even function: The graph is symmetric with respect to the y-axis.
   Odd function: The graph is symmetric with respect to the origin.

   Let $g : \mathbb{R} \to \mathbb{R}$ be a differentiable even function. It is known that $g'(1) = 2$. Then, what can you say about $g'(-1)$?

APPENDIX 3: Evaluation criteria for answers of the written test

Classification criteria for the study success

The following criteria describe how the answers to the question about the study success in mathematics during university studies (see the test form in Appendix 1) were classified.

The priority order for the classification:

1. A numerical estimate for the average grade or the grade of a study module.

2. A verbal estimate or description concerning the average grade.

3. Other kind of description concerning the study success.

In the following, the criteria for each class are listed. In the classification, an answer was placed into a class if one criterion of the class in question was fulfilled.

Class 0: Unclassified cases

- The answer was missing.
- The answer was too vague.
- The answer concerned only a part of the studies.

Class 1: Poor success

- The estimated average grade was on the interval [1.00, 1.66].
- If several grades were mentioned, it was stated that mostly the grades had been on the above-mentioned interval.
- The study success was verbally described by using expressions like "bad", "poor", "quite bad", "quite poor", "the courses have been scarcely passed", "under the average level", and so on. [1]

[1]The original Finnish expressions were "huono", "melko huono", "heikko", "heikohko", "välttävä", "kurssit läpi rimaa hipoen" and "keskitason alapuolella".

Class 2: Satisfactory success

- The estimated average grade was on the interval [1.67, 2.33].
- If several grades were mentioned, it was stated that the grades had mostly been on the above-mentioned interval.
- The study success was verbally described by using expressions like "moderate", "passable", "average", "satisfactory", and so on. [2]
- The study success was described fluctuating.

Class 3: Excellent success

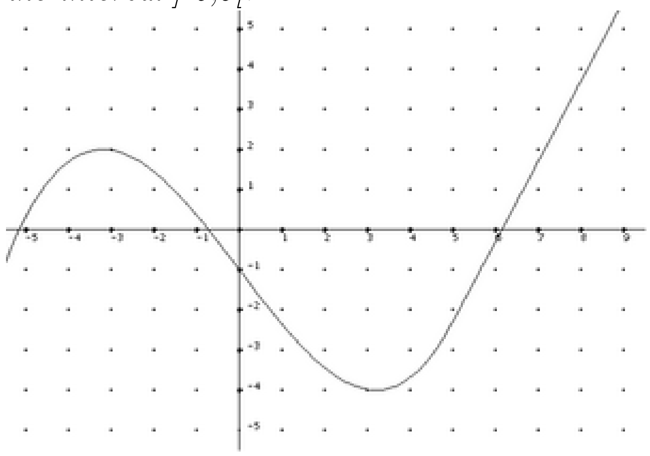
- The estimated average grade was on the interval [2.34, 3.00].
- If several grades were mentioned, it was stated that the grades had mostly been on the above-mentioned interval.
- The study success was verbally described by using expressions like "good", "quite good", "ok", "all right", and so on. [3]

[2]The original Finnish expressions were "kohtalainen", "kohtuullinen", "keskitasoinen" and "tyydyttävä".

[3]The original Finnish expressions were "hyvä", "melko hyvä", and "ihan ok".

Evaluation criteria for task answers

1. *The diagram below presents a graph of a differentiable function f defined on the interval]-6,9[.*



   a) *On the basis of the graph, determine the value of the derivative of f when x=-2.*

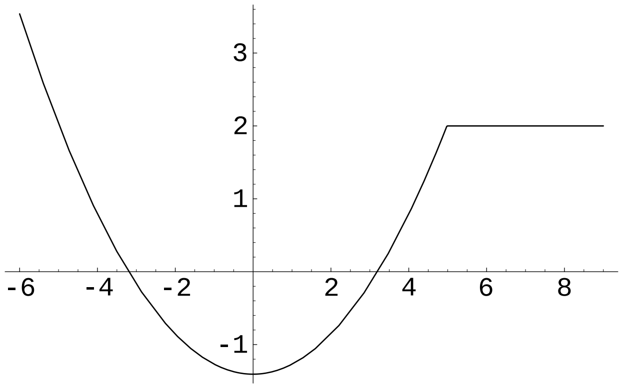
   The correct answer is about -0.86. Answers on the interval $[-1.30, -0.50]$ were considered acceptable.

   Other comments:

   - If the answer was outside the interval mentioned above but it came clearly out that the principle of the method used was correct, the answer was accepted.
   - If the answer was inside the interval mentioned above but it came out that the principle of the method used was erroneous, the answer was not accepted.
   - If the value of the derivative had been determined correctly but it had been done at a point different from $x = -2$, the answer was accepted. However, the definite requirement for acceptance was that the point at which the determination was done was clearly identified in the answer.
   - If the minus-sign in front of the answer was missing but otherwise the solution fulfilled the criteria mentioned above, the answer was accepted. This practice was chosen, because the most important goal of this task was to measure how large a proportion of the respondents understood the basic idea of a visual interpretation of derivative (for example, an interpretation of derivative as a slope of a tangent line of the graph). Indeed, the sign belongs essentially to this interpretation, but the sign is one out of the three issues which are examined in Task 1b. Thus it was decided that, in this task, a wrong sign would not lead to the rejection of the whole answer.

   b) *Sketch the graph of the derivative function of f.*

   The graph of the derivative function looks like the following:



In evaluation it was examined whether the zeros, the sign and the shape of the graph were correct.

The zeros were considered acceptable if they were (or seemed to be) on the intervals [-3.50, -2.75] and [2.75, 3.50]. (Correct values are approximately -3.20 and 3.20.) In some drawings the scale at the x-axis was inexact or it was missing. However, in most cases it was possible to reason whether the respondent had understood where the derivative was zero. If this was not possible, the zeros were not accepted.

If the graph was first from the left above, then below and then again above the x-axis, **the sign** was considered acceptable. The zeros were not required to be correct.

The shape of the graph was considered acceptable if the following criteria were fulfilled:

- The value of the derivative function had to be approximately correct everywhere. If the scales were not expressed in the diagram, attention was paid to the proportions between the values in different parts of the graph.
- Zeros had to be acceptable according to the criteria mentioned above.
- The direction of the graph (increasing or decreasing) had to be mainly correct in each part of the graph.
- The graph had to be connected.
- Appproximately from the point $x = 5.0$ to the right, the graph had to be a horizontal line.

Other comments on the evaluation of the shape:

- The acceptance of the shape was not dependent on whether the graph had any corners.
- If the sign was systematically wrong, but the shape was otherwise acceptable (the diagram presented a graph of $-f$), the shape was accepted. This was the only situation in which the shape could be acceptable even if the sign was rejected.

2. *Which of the following functions are continuous and which are differentiable? "Yes/No" -answers are sufficient.*

   a) $f(x) = \begin{cases} x+1, & x < 1, \\ -2x+6, & x \geqslant 1. \end{cases}$

      This function is neither continuous nor differentiable.

   b) $f(x) = \begin{cases} x+2, & x < 1, \\ -2x+5, & x \geqslant 1. \end{cases}$

      This function is continuous but not differentiable.

   c) $f(x) = x^2 - 4x + 3$

      This function is both continuous and differentiable.

   d) $f(x) = \begin{cases} x^2 - 4x + 3, & x \neq 4, \\ 1, & x = 4. \end{cases}$

      This function is neither continuous nor differentiable.

No special criteria were needed in the evaluation of this task.

In Tasks 3a, 3b, 4a, 4b and 5a the answers were graded by using the points 0, 1 and 2. Two points were given, if the solution fulfilled all the criteria listed below. If the main principle in the solution was correct, but the answer did not fulfill every criteria, one point was given. In the following, the criteria for two points are listed for each task. In the cases in which a formal proof was required (Tasks 3b and 4b), more detailed criteria for one point are given. In these cases, example solutions are also presented.

3. a) *How would you explain, by using graphical interpretations, why the derivative of a constant function is equal to zero everywhere?*

   Criteria for two points:

   1. An appropriate informal interpretation for a constant function had to be presented. Acceptable interpretations were, for example, that the constant function is a function whose graph is a horizontal line or a function whose values do not change.
   2. An appropriate informal interpretation for a derivative had to be presented. Acceptable interpretations were, for example, that derivative means steepness of the graph (or steepness of a tangent line) or that derivative measures the rate of change.
   3. A reasonable conclusion based on criteria 1 and 2 justifying the claim had to be presented.

   Other comments:

   - The interpretations could be presented either visually or verbally.

   b) *Prove the same by using the formal definition of derivative.*

   An example solution:

   Let $f : \mathbb{R} \to \mathbb{R}, f(x) = c, c \in \mathbb{R}$.

   For all $x_0 \in \mathbb{R}$:

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{h \to 0} \frac{c - c}{h} = \lim_{h \to 0} \frac{0}{h} = 0.$$

   Criteria for two points:

   1. A correct formal definition for the constant function had to be presented.
   2. The proof had to be based on the definition of derivative.
   3. The proof had to be general enough: It had to prove that the claim is true for all constant functions at all points of the domain.

   Detailed criteria for one point:

   1. The key argument had to be based on the definition of derivative.
   2. Deficiencies could appear in the calculation of the limit or with respect to the generality (the criterion 3 in the above list).

4. *Claim: Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function and $x_0 \in \mathbb{R}$. Then*

$$\lim_{h \to 0} \frac{f(x_0) - f(x_0 - 3h)}{3h} = f'(x_0).$$

a) *How would you visually argue this claim by using a diagram?*

Criteria for two points:

1. The given difference quotient had to be interpreted either as steepness of a secant line or as an average rate of change. These interpretations had to be justified.
2. By using the chosen interpretation, the limiting process had to be explained.
3. By using the chosen interpretation, the state after the limiting process had to be explained.
4. The derivative had to be interpreted either as the steepness of a tangent line or as the instantaneous rate of change.
5. The structure of the argument had to be coherent.

Other comments:

- When evaluating answers with respect to the criteria 2 and 3 and with respect to the interpretation of the difference quotient in the criterion 1, the respondent's answer to Task 5a could be used if this answer seemed to reveal better how the author understood the issues in question.

b) *Prove the claim formally by using the definition.*

An example solution:

Let us denote $\tilde{h} := -3h$. By using this change of a variable, we receive:

$$\begin{aligned}
\lim_{h \to 0} \frac{f(x_0) - f(x_0 - 3h)}{3h} &= \lim_{\tilde{h} \to 0} \frac{f(x_0) - f(x_0 + \tilde{h})}{-\tilde{h}} \\
&= \lim_{\tilde{h} \to 0} \frac{f(x_0 + \tilde{h}) - f(x_0)}{\tilde{h}}
\end{aligned}$$

The last expression is $f'(x_0)$, according to the definition of derivative.

Criteria for two points:

1. The definition of derivative had to have a key role in the answer.
2. The expression in the task had to be modified with appropriately reasoned steps to the form which appears in the definition of derivative. This process had to include an idea of the change of a variable.

Detailed criteria for one point:

1. It had to come out from the answer that the author has understood that the given expression has to be modified to the form which is expressed in the definition of derivative.

2. An idea about the change of a variable had to be included in the answer, either explicitly or implicitly.
3. The calculation of the limit could be unfinished or it could include erroneous phases or phases whose justifications were not detailed enough.

5. a) *How would you explain, on the basis of the definition of derivative, why the derivative of a differentiable function f at a given point x_0 is equal to the slope of the tangent line drawn to the graph of the function at the point $(x_0, f(x_0))$?*

   Criteria for two points:

   1. It had to be explained why the difference quotient can be interpreted as a slope of a secant line (the state before the limiting process).
   2. The visual interpretation of the limiting process had to be described.
   3. The state after the limiting process (the secant line has become a tangent line) had to be described.
   4. The structure of the argument had to be coherent.

   Other comments:

   - When evaluating answers with respect to the criteria 1-3, the respondent's answer to Task 4a could be used if this answer seemed to reveal better how the author understood the issues in question.

   b) *How would you explain, on the basis of the definition of derivative, why the graph of a differentiable function cannot have a corner?*

Answers to this task were not graded, but they were classified on the basis of the key argument which was presented in the answer (see Appendix 4).

6. *Let* $f : \mathbb{R} \to \mathbb{R}$:

$$f(x) = \begin{cases} x^2 + 1, & x \in \mathbb{Q}, \\ 1, & x \notin \mathbb{Q}. \end{cases}$$

Sketch the graph of function f and determine the points where f is differentiable. Give exact arguments.

An example solution:

If $x_0 \neq 0$, f is not continuous at the point x_0, because in every neighborhood of the point x_0, there exists a point x_1 such that $|f(x_1) - f(x_0)| \geq {x_0}^2 > 0$. Therefore, the only point at which function f can be continuous is zero. Because continuity is a necessary condition for differentiability, function f cannot be differentiable at any point excluding zero. At zero, the difference quotient can be approximated in the following way:

$$\left| \frac{f(0+h) - f(0)}{h} \right| \leq \left| \frac{h^2}{h} \right| = |h| \xrightarrow[h \to 0]{} 0.$$

Thus, f is differentiable at zero, and $f'(0) = 0$.

When evaluating the answers to this task, the following matters were checked:

- Is the answer correct? (Function f is differentiable at the point $x = 0$.)
- Has differentiability at the point $x = 0$ been proved correctly?
- Has it been argued why the function is not differentiable at other points? What kind of arguments have been used for this?

Incorrect solutions and their arguments were also classified (see Appendix 4).

Evaluation of the tasks used in the interviews

In this study, no general evaluation criteria were established for the tasks used in the interviews. In studies presented in Articles A-D, students' working with these tasks has been analysed mainly qualitatively, except that in Article C, students' approaches and success in Task 3 were classified quantitatively. The applied criteria are described case-specifically in each article.

APPENDIX 4: Results of the written test

In the following, the results of the written test are collectively presented. The evaluation and the scoring of the answers have been performed according to the criteria presented in Appendix 3.

Results of Task 1

a) According to the criteria presented, 126 out of 146 participants answered to Task 1a in an acceptable way.

b) In Task 1b, 99 answers fulfilled the required criteria with respect to the zeros, 99 answers with respect to the sign and 60 answers with respect to the shape.

Results of Task 2

The distribution of answers in Task 2 are presented in the following table. The numbers of the right answers are printed in boldface.

	Task 2a	Task 2b	Task 2c	Task 2d
Continuous and differentiable	0	5	**146**	1
Continuous, not differentiable	5	**130**	0	6
Discontinuous and differentiable	17	1	0	38
Discontinuous, not differentiable	**122**	9	0	**99**
Other answer	2	1	0	2

Results of Tasks 3a, 3b, 4a, 4b and 5a

The following table reveals how many respondents obtained 0, 1 and 2 points from Tasks 3a, 3b, 4a, 4b and 5a. As well, means and standard deviations of the obtained points are announced.

		Points				
		0	**1**	**2**	**Mean**	**St. dev.**
Task	3a	17	68	61	1.30	.67
	3b	31	19	96	1.45	.82
	4a	97	28	21	.48	.74
	4b	89	34	23	.55	.75
	5a	85	33	28	.61	.79

Results of Task 5b

In Task 5b, the respondents were asked to explain how it follows from the definition of derivative that the graph of a differentiable function cannot have corners. The answers were distributed in the following way:

- In 34 answers, it was explained that, at the corner point, the left-hand and right-hand limits of the difference quotient are not equal or that the limit of the difference quotient is not unambiguous. Only in three answers the inequality of the left-hand and right-hand limits was argued.
- In 26 answers, the crucial argument was that at the corner point it is impossible to draw an unambiguous tangent line to the graph. In 15 of these answers, no argument was presented for this. In the remaining 11 answers, it was argued why derivative on the basis of the definition can be interpreted as the slope of a tangent line.
- In seven cases, the answer was a mixture of the arguments presented above.
- In two answers, derivative was interpreted as an instantaneous rate of change of the values of the function. In these answers, it was explained that because the rates of change on the left and right sides of the corner are unequal, it is impossible to determine an instantaneous rate of change at the corner point.
- Ten answers made sense at least to some extent, but they could not be placed into any of the categories mentioned above.
- In 67 cases, the answer was missing or it was irrelevant.

Results of Task 6

The distribution of the answers was the following:

- The correct solution (f differentiable at the point $x = 0$) was given in 37 answers. In 16 of these answers, differentiability at zero was argued on the basis of the definition so that at least the main idea of the argument was correct. In 31 of these 37 answers, it was tried to argue why the function was not differentiable outside zero. These arguments were distributed in the following way: In 12 answers, it was explained why the difference quotient did not have a limit outside zero, in 11 answers the discontinuity of the function was the crucial argument, and in one answer both of these arguments were used. In three answers, it was stated that the function did not have a limit outside zero. In four answers the argumentation for nondifferentiability was essentially erroneous or inadequate.

- 17 respondents answered that the function was differentiable everywhere. Nine respondents argued this by stating that functions $x^2 + 1$ and 1 are both differentiable.

- 13 respondents answered that the function was not differentiable anywhere. Arguments in these answers were mainly similar to the arguments for nondifferentiability presented for the correct solutions except that the speciality of the point $x = 0$ was not taken into account.

- Ten respondents thought that the function was differentiable at the rational points. Arguments for this were quite vague or they were totally missing.

- Three respondents thought the function to be differentiable at the irrational points and at zero. Two of them thought that there are intervals consisting of irrational numbers on the real line.

- One respondent proposed that the function was differentiable at the irrational points.

- Seven respondents proposed another solution (for example, $x \neq 0$, $x \neq 1$, etc.). Arguments for these solutions were vague or they were missing.

- 58 respondents did not answer this question at all.

APPENDIX 5: Statistics about the participants

In the following, some statistics about the participants of the study is presented. The data are based on the participants' own answers to the background questions of the test form (see Section 4.2.2 and Appendix 1). The distributions in the following tables are first presented for all participants and then separately for those who were interviewed.

Table 1: Gender distribution of the participants.

Gender	**All participants**	**Interviewees**
Male	64	15
Female	53	6
Unknown	29	0
Together	146	21

The gender distribution is presented in Table 1. Because gender was not inquired in the test -which, in fact, was a deficiency of the test form- it was inferred on the basis of the name of the participant. As told in Section 4.2.1, some participants used a pseudonym instead of a real name, and this made it difficult to infer the gender in some cases. However, in some cases the pseudonym revealed the participant's gender quite reliably: For example, if the pseudonym was a male name, it was concluded that the participant probably was a man.

There were far more men among the interviewees than in the whole sample. This was not intentional, but it resulted from various issues concerning the selection of the interviewees (see Section 4.2.1).

Most of the participants were majoring in mathematics, but the sample also included a significant proportion of students majoring in physics (see Table 2). Among the interviewees, the mathematics majors were overrepresented, because almost all of the most successful students in the sample were majoring in mathematics, and it was important to choose some of them to the interviews.

Table 2: Distribution of the participants according to their main subject.

Main subject	**All participants**	**Interviewees**
Mathematics	89	17
Physics	29	4
Chemistry	9	0
Education	8	0
Other	9	0
No answer	2	0
Together	146	21

Table 3 reveals that most of the participants had passed 35-64 Finnish credits in mathematics. Subject teacher students who were majoring in math-

Table 3: Number of passed credits in mathematics (in Finnish credits).

Number of passed credits	**All participants**	**Interviewees**
20-34	23	0
35-49	52	9
50-64	52	8
65-79	7	0
80-99	3	1
100-130	7	3
No answer	2	0
Together	146	21
Mean	49.6	58.7
Median	45.5	52.0
Standard deviation	20.8	27.0

ematics were required to pass about 65-70 credits in mathematics, and students who had mathematics as a secondary subject had to pass 35 credits. Thus, it is probable that most of the participants had mainly passed their compulsory mathematics studies. Students who had passed more than one hundred credits were going to complete a post-graduate degree (licentiate or Ph.D.) in mathematics. They were usually the most successful students, and, for this reason, three of them were interviewed. All interviewees had passed at least 35 credits in mathematics.

The announced numbers of passed credits are based on students' own estimates, and thus they are not fully reliable. However, it can be assumed that most students remembered it at least approximately, and thus the distribution probably does not include significant errors. Yet, it has to be noticed that the number of passed credits cannot be considered as an absolute measure of the amount of studies, because the contents of the studies varied.

The test form included a question about the beginning year of university studies. From the answer to this question, it was deduced how many years a student had studied at university. However, it is not known if the student had missed some years after he/she began the studies. Thus the distribution presented in Table 4 is not very reliable. Alternatively, the number of study years could have been asked, but this question could have as well been problematic: For instance, in some cases students had possibly studied only part-time during some years. The distribution in Table 4 can be considered as a distribution with respect to the total length of the study history at university.

The information about the success in mathematics studies is based on the participants' verbal descriptions in the test form (see Appendix 1). Thus it can be considered only suggestive, and due to that, only a three-class scale was used in coding the variable dealing with this information. If the answer contained an estimate on the average grade, this was the primary criterion in the classification. If no estimate on the average success was presented,

Table 4: Duration of studies in years before the academic year 2004-2005.

Duration of studies (in years)	**All participants**	**Interviewees**
1-2	24	0
3-4	71	11
5-6	31	8
7 or more	19	2
No answer	1	0
Together	146	21
Mean	4.8	4.9
Median	4.0	4.0
Standard deviation	4.2	1.6

the classification was made on the basis of verbal descriptions. A more detailed description of the classification criteria is presented in Appendix 3. The distribution resulting from this analysis is presented in Table 5. Among the interviewees, there were students from all study success classes. In fact, the class of poor success was a little overrepresented, whereas the class of excellent success was underrepresented respectively.

Table 5: Distribution of the study success in mathematics.

Study success class	**All participants**	**Interviewees**
Poor	27	5
Satisfactory	68	9
Excellent	43	5
Unclear or missing answer	8	2
Together	146	21

Table 6: Distribution of the participants with respect to their aims for the future.

Aim	All participants	Interviewees
Teaching at comprehensive school	13	1
Teaching at upper secondary school	31	3
Teaching at comprehensive school or at upper secondary school	53	10
Research at university	10	1
Teaching, the school level missing, or several levels mentioned	16	0
Research or teaching	7	1
Another work	5	2
Unclear aim or answer missing	11	3
Together	146	21

The participants' plans and aims with respect to their future career are presented in Table 6. Naturally, most of them were planning on teaching at school, either at comprehensive school or at upper secondary school. Thus, it is very probable that, to a great extent, the sample really consisted of prospective mathematics teachers, which, in fact, was the original goal.

The data concerning the number of passed credits and the study success have been used in Article E. The other parts of the data concerning the backgrounds have not been analysed in this thesis further.

APPENDIX 6: Extra statistical analysis relating to the study presented in Article E

In Section 5.2 in the present thesis and in Article E, two subgroups of the sample are mentioned: the group of studens whose good success in the informal tasks can be explained neither by good study success nor by the large number of passed credits, and the group of students whose good success in the formal tasks cannot be explained by good study success but who had significantly more passed credits than the whole sample on average. However, neither in Section 5.2 nor in Article E has it been described how these subgroups were found. In the following, a statistical analysis about the relationships between the obtained points from the informal and formal tasks and the estimates of the number of passed credits and the study success is described. In this analysis, we study how the circled classes in Tables 1 and 2 differ from the whole sample with respect to the number of passed credits. The classes in the lower left corners in each tables are the subgroups mentioned above.

Table 1: Crosstabulation: Scores from the informal tasks vs. estimated study success

		Estimated success			
		Poor	**Satisfactory**	**Excellent**	**Total**
Informal	**0**	3	9	2	14
	1	14	22	16	52
	2	7	23	7	37
	3	2	10	11	23
	4	1	4	7	12
Total		27	68	43	138

Table 2: Crosstabulation: Scores from the formal tasks vs. estimated study success

		Estimated success			
		Poor	**Satisfactory**	**Excellent**	**Total**
Formal	**0**	10	14	3	27
	1	4	8	3	15
	2	9	22	17	48
	3	2	15	12	29
	4	2	9	8	19
Total		27	68	43	138

As mentioned in Section 3.3 in Article E and in Appendix 5, the estimates about the number of passed credits and the study success are not very reliable. Therefore, the analysis which is based on these variables can convey only to suggestive conclusions.

A crosstabulation between the obtained points from informal tasks and the estimated study success is presented in Table 1 and a crosstabulation between the obtained formal points and the estimated study success is presented in Table 2. Tables 1 and 2 are the same as Tables 3 and 4 in Article E. These crosstabulations show that, both in the informal tasks and in the formal ones, some students obtained high points even if their study success had been at most at the satisfactory level, and, on the other hand, especially in the informal tasks, several students with excellent study success obtained low points. In these cases, the study success and the test results were opposite (one was poor and the other was good), and thus it is improbable that these factors had causality between them. The goal of the following analysis is to explore whether the number of passed credits might be a factor explaining the test results in these cases. In addition, the cases in which success in the studies and the results of the test are parallel (both poor or both good) are explored in order to find out whether the number of passed credits could in these cases have any effect on the success in the test. In order to study these questions, the classes defined in Table 3 were formed. Each class consists of students in one corner in Table 1 or in Table 2 (see the circlating).

In this connection, it is important to notice that statistical tests can reveal only dependences between variables, but they do not show causalities between them. Therefore, it is only justifiable to speak about possible effects and possible explanatory factors when interpreting the dependences which were found.

Table 3: The definitions of the classes that were formed on the basis of the estimated success in mathematics studies and on the basis of the results of the test.

Group	Study success	Test result
Inf-lower-left	poor or satisfactory	informal ≥ 3
Inf-upper-right	excellent	informal ≤ 1
Inf-upper-left	poor	informal ≤ 1
Inf-lower-right	excellent	informal ≥ 3
Form-lower-left	poor or satisfactory	formal ≥ 3
Form-upper-right	excellent	formal ≤ 1
Form-upper-left	poor	formal ≤ 1
Form-lower-right	excellent	formal ≥ 3

The students with satisfactory success were included in the classes Inf-lower-left and Form-lower-left, because, otherwise, these classes would have been too small. Due to this expanding, most students in the classes Inf-lower-left and Form-lower-left had satisfactory study success, but, in any case, the students who had excellent study success were omitted. Class Form-upper-right consists only of six students, and, therefore, reliable conclusions on the basis of this class cannot be drawn. Despite that, this class was not expanded, because the aim in the case of the class Form-upper-right - like in the case of the class Inf-upper-right - was to examine the students who clearly had done well in their studies but whose success in the tasks of the test was poor. Satisfactory study success is to some extent insufficient, and thus it cannot be considered as a very valid factor for explaining the good success in the test, but, on the other hand, it can have a significant influence on the failures. This assumption is an additional reason for the decision to take students with satisfactory study success into the classes Inf-lower-left and Form-lower-left but not into the classes Inf-upper-right and Form-upper-right.

Means, medians, ranges and standard deviations of passed credits were calculated for each of the formed classes (Table 4). The distributions with respect to the quartiles of the whole sample, presented in Table 5, reveal how many students in each class had passed relatively small or relatively large numbers of credits. Moreover, it was studied by the one-sample mean test and by the binomial median test how significantly the means and the medians in each of the formed classes differed from the mean and the median of the whole sample (Table 6). On the basis of this analysis, it was possible to find out whether the small or large number of passed credits was a representative feature of the class at issue. Primarily, these conclusions were made on the basis of the results of the mean and the median tests, but because the frequencies of the classes were small, the quartile distributions in Table 5 and the statistics in Table 4 were also taken into account.

Table 4: Frequencies, means, medians, ranges and standard deviations of the number of passed credits in the formed classes and in the study-success classes.

	Passed credits				
Group	**Freq.**	**Mean**	**Median**	**Range**	**St. dev.**
Inf-lower-left	17	51.3	50.0	30-90	16.7
Inf-upper-right	18	44.2	35.5	29-85	16.9
Inf-upper-left	17	38.7	38.0	23-60	11.6
Inf-lower-right	18	60.5	55.0	35-130	28.3
Form-lower-left	28	58.9	57.8	30-129	20.4
Form-upper-right	6	40.0	35.0	35-60	10.0
Form-upper-left	14	35.7	36.0	23-55	9.4
Form-lower-right	20	64.0	57.0	29-130	28.9
Poor success	27	41.0	40.0	23-70	12.5
Satisf. success	67	49.1	50.0	20-129	18.7
Excellent success	43	53.5	47.0	29-130	23.7
Whole sample	144	49.6	45.5	20-130	20.8

Table 5: Distributions of students in each class with respect to the quartiles of the distribution of the number of passed credits in the whole sample. It has to be noted that in the whole sample, as many as 20 students had, according to their own estimates, passed exactly 35 credits, which is just the upper limit of the first quartile. Owing to that, the first quartile is bigger than the second one. In fact, the first quartile includes 29.9% and the second quartile 20.1% of the students of the whole sample.

	Number of students having passed credits			
Group	**20-35**	**36-45**	**46-59**	**60-130**
Inf-lower-left	5	2	6	4
Inf-upper-right	9	3	2	4
Inf-upper-left	7	6	2	2
Inf-lower-right	2	6	2	8
Form-lower-left	2	5	8	13
Form-upper-right	4	1	0	1
Form-upper-left	7	5	2	0
Form-lower-right	2	4	5	9
Poor success	9	9	6	3
Satisf. success	21	10	19	17
Excellent success	12	9	8	14
Whole sample	43	29	36	36

Table 6: Significances of the differences between the classes and the whole sample.

Group	**Mean** One-sample t-test Test value: 49.59	**Median** Binomial test Cut value: 45.50
Inf-lower-left	.685	.629
Inf-upper-right	.191	.238
Inf-upper-left	.001	.049
Inf-lower-right	.119	.815
Form-lower-left	.023	.013
Form-upper-right	.066	.219
Form-upper-left	$< .001$	.013
Form-lower-right	.039	.115
Poor success	.001	.122
Satisf. success	.832	.625
Excellent success	.291	1.000

In the classes Inf-lower-left, Inf-upper-right, Form-lower-left and Form-upper-right, the students' study success and their test success were contrary, and, thus, in these cases, the level of study success could not be considered as an explanatory factor for the test results. On the basis of the statistical analysis presented in Tables 4-6, the following conclusions from the effects of the amount of passed credits could be drawn:

- The amount of passed credits does not explain why the students in the class Inf-lower-left did well in the informal test tasks.

- In the case of the class Form-lower-left, the large number of passed credits is a possible explanatory factor for the good success in the formal test tasks.

According to the mean and the median tests, the class Inf-upper-right did not differ statistically from the whole sample. However, half of the students in this class were in the first quartile in Table 5. Thus, one could not draw reliable conclusions about the effect of the amount of passed credits. No valid conclusions about the class Form-upper-right could be drawn, either, due to the small frequency of this class.

In the classes Inf-upper-left, Inf-lower-right, Form-upper-left and Form-lower-right, the students' study success was parallel with their test success, and, thus, in these cases, the level of study success could be considered as a possible explanatory factor for the test results. The following conclusions from the effects of the amount of passed credits could be drawn:

- With the classes Inf-upper-left and Form-upper-left, the small number of passed credits is a possible factor, along with poor study success, which might explain the poor test success.

- The amount of passed credits does not explain the good test success of the classes Inf-lower-right and Form-lower-right.

According to the one-sample mean test, the difference between the class Form-lower-right and the whole sample was statistically almost significant. However, it has to be taken into account that some students in this class had a very large number of passed credits, which were strongly weighted in this test. Instead, according to the binomial median test, the difference was clearly not significant. Thus, it is not justifiable to consider the large number of passed credits in this class as a possible factor explaining the good success in the formal tasks. On the other hand, it has to be noted that almost half of the students both in the class Inf-lower-right and in the class Form-lower-right were in the fourth quartile in Table 5.

How should the dependence between the number of passed credits and the study success be taken into account when assessing the validity of the results of the analysis presented above? These variables had a positive correlation (Spearman's rho .215) with p-value .011. To what extent were the observed differences between the classes and the whole sample caused by this correlation? In order to explore this, the same statistical analysis which was made for the "corner-classes" above was also made for the study success classes. The results are included in Tables 4, 5 and 6, and, on the basis of them, it is very clear that there was no significant difference between the students with excellent study success and the whole sample with respect to the number of passed credits. Thus the dependence mentioned cannot have any significant effect on the results of the classes Inf-upper-right, Inf-lower-right, Form-upper-right and Form-lower-right. Instead, the mean of the number of passed credits of the students with poor study success differed statistically significantly from the corresponding mean of the whole sample. In addition, the quartiles in Table 5 revealed that the distribution was skew to the right. However, the difference between the medians was not statistically significant. In any case, the difference between the means could have some influence on the results of the classes Inf-lower-left, Inf-upper-left, Form-lower-left and Form-upper-left. In the cases of the classes Inf-upper-left and Form-upper-left, this dependence may partially explain why these classes differed from the whole sample with respect to the number of passed credits. This dependence may also partially explain why the numbers of passed credits in the class Inf-lower-left were not significantly larger than in the whole sample. In the case of the class Form-lower-left, this dependence strengthens the result which was found: Despite the positive correlation between the number of passed credits and the study success, the students with poor or satisfactory study success who succeeded well in the formal tasks had passed significantly more credits than the whole sample on average. The number of passed credits between the class Form-lower-left and all students with poor or satisfactory study success was also compared: The p-value of the difference between the means was .004 and the p-value of the difference between the medians was .013. This means that, among the students having poor or satisfactory study success, the students who succeeded well in the formal tasks had passed significantly more credits in mathematics than the other students.

It is important to take into account that the above analysis was made statistically at the class level, not at the individual level. Thus, it is very likely that every class contained students for whom the class-level conclusions are not valid. For example, in the class-level analysis, it turned out that the amount of passed studies did not explain why the students in the class Inf-lower-left succeeded well in the informal tasks. In spite of that, it is possible that for some individuals in this class the large number of passed credit was an important factor which made succeeding in these tasks possible.