# Finnish mathematics teacher students' informal and formal arguing skills in the case of derivative

Antti Viholainen

In this study, formal and informal reasoning skills of 146 Finnish subject-teacher students in mathematics are investigated. The students participated in a test in which they were asked to argue two claims concerning derivative both informally and formally. The results show that the success in the formal tasks and the success in the informal tasks were dependent. However, there were several students who did well in the formal tasks despite succeeding poorly in the informal tasks. The success both in the formal tasks and in the informal tasks was dependent also on the amount of passed studies in mathematics and on the success in these studies. Moreover, these factors could have a stronger effect on the formal than on the informal reasoning skills.

In principle, mathematical concepts are both formal and abstract. However, they can often be illustrated by concrete interpretations which are more or less informal. The concrete interpretations have also an essential role in thinking and problem solving processes of mathematicians (Raman, 2002; Stylianou, 2002). Mathematical claims can also be argued either formally or informally: They can be proved by using definitions, axioms, previously proven theorems and formal language, or they can be justified by using visual, physical or other more concrete interpretations. The formal proof is usually demanded for the final form of the argument, but explanations based on concrete interpretations can help an individual to understand, even at a more conceptual level, why a claim is really true.

This study concerns Finnish mathematics teacher students' abilities to produce informal and formal arguments in the connection of the concept

**Antti Viholainen**
*University of Jyväskylä*

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

of derivative. The study is based on a quantitative analysis of students' answers to a written test in which students were asked to argue two mathematical claims concerning the concept of derivative both informally and formally. The aim of the study was to find answers to the following questions:

– Is it easier for the students to produce informal or formal arguments for the given claims?

– How are the abilities to produce informal and formal arguments dependent?

– How are these abilities dependent on the amount of passed studies in mathematics and on the success in these studies?

The concept of derivative is essential in the basic analysis, but often its understanding has proved challenging for students. The limiting process has an essential role in construction of the concept of derivative, and thus the difficulties met in understanding the concept of derivative are often connected to epistemological obstacles of the concept of limit. These obstacles, furthermore, may be due to problems in understanding properties of real numbers, the concept of infinity, and so on. Several studies about these issues are mentioned, for instance, in Parameswaran (2007). However, in the present paper the main attention is not directed to the epistemology of derivative or to the learning problems as such. Instead, the main goal is to examine the relationship between the informal and formal reasoning. Presumably, the results of this analysis are dependent on the mathematical context and on its epistemology.

## Informal and formal arguments in mathematical thinking

### *Definitions of 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 necessitates 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. In the classification of arguments used in this study, the arguments are classified on the basis of their warrants. The definitions of informal and formal arguments are stated in the following way:

An argument is *informal,* if its warrants are based on a use of *informal interpretations* of concepts or situations which the argument concerns. The informal interpretations can be *visual*, *physical* or other interpretations about the meaning of the concepts or situations. An argument is, for one, *formal,* if its warrants are based on the elements of the formal axiomatic system of mathematics, that is to say, it is based only on the formal definitions, axioms and theorems. A formal argument explicitly shows how the conclusion logically follows from these elements and from the given data. In addition, a formal argument has to be systematic and rigorous.

The concept of derivative can be interpreted formally, visually and physically, for instance, in the following ways:

*Formally,* a value of the derivative of a function $f : \mathbb{R} \to \mathbb{R}$ at a point $x_0 \in \mathbb{R}$ equals with the limit

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

assuming that this limit exists. Alternatively, the limit of the difference quotient can be presented in the following form:

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

*Visually,* derivative describes the *steepness* of a 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 is in the graph. To be more exact, derivative is the slope of the tangent line drawn to the graph. *Physically,* derivative can be interpreted as an instantaneous rate of change of the dependent variable. For example, an instantaneous speed is the derivative of the passed distance, an instantaneous acceleration is the derivative of the speed, the electric current is the derivative of the flowing electric charge through a surface etc. Usually, derivative has been seen as a mathematical tool, which is needed when physical magnitudes are explored.

However, visual and physical interpretations can also be used as thinking tools in consideration of pure mathematical problems concerning derivative.

On the basis of the above presented definitions, arguments cannot be absolutely divided into informal and formal ones. In the cases of some arguments, it may depend on personal interpretations whether a warrant of an argument is seen to be bound to definitions, axioms and theorems or whether it is seen 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*) that the conclusion is compatible with the formal axiomatic system and thus a part of it. On the other hand, 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 otherwise the structure of the proof is formal.

### *The purpose of informal and formal arguments*

The variety of arguments which may be used in mathematical reasoning is very wide (cf. Harel & Sowder, 1998). In practice, mathematical arguments may be based even on such non-analytical elements as intuition, immediate perceptions or relying on some authority. However, the truth value of a mathematical proposition should – at least in principle – be fully independent of individual preferences or interpretations, social conventions, negotiations and subjective conceptions (Goldin, 2003), and formal definitions, proofs and axioms should have a conclusive role in determining what is finally true and what is false in mathematics.[1] In this way, the exact and unambiguous nature of mathematical knowledge can be preserved. On the other hand, a prerequisite for creative mathematical thinking is that an individual is able to interpret abstract mathematical concepts and claims in a meaningful way.

Final products of mathematical calculations or proofs often appear to be based on symbolic manipulation. Visual or physical interpretations or holistic views about the situation do not usually come up from them. However, these aspects are important in the producing process. According to Weber and Alcock (2004), *syntactic proof production* means a pure symbolic manipulation, which is made in a logically permissible way, whereas in *semantic proof production* the prover uses instantiations of mathematical objects in order to suggest and guide the formal inferences that he/she draws. If the problem is simple or the prover masters well the needed procedural skills, it may be possible to produce a proof by using only syntactic proof production, but, otherwise, proceeding in the process may be difficult without any semantic view. It is also notable that semantic proof production offers explanations, often at a more global and more intuitive level than syntactic proof production for why a result is true. Thus, a proof produced by using aspects of the semantic proof production may be both convincing and explanatory to the prover.

Raman (2002, 2003) distinguishes *private arguments* from *public arguments*. She defines that a private argument is an argument engendering understanding and that a public argument is an argument having sufficient rigor for a particular mathematical community. Correspondingly, Rodd (2000) defines that *a justification* is an argument which at an intuitive level gives a reason to believe that a claim ought to be true and that *a warrant* is an argument which exhibits a logical inference chain which shows that the claim is undoubtedly true [2]. Both these classifications, like Weber's and Alcock's classification concerning the proof production, are based on the observation that mathematical arguments presented in a formal form are rarely explanatory in the intuitive level, and thus another kinds of arguments are needed.

Due to the nature of mathematics, it is reasonable to suppose that subject matter teachers in mathematics know the criteria of formal arguments and are able to produce correct formal arguments. In addition to that, it is important that they are aware of the role of visual and physical interpretations of mathematical concepts and that they are aware of their usage in mathematical reasoning. All these are essential factors of a teacher's subject matter knowledge, but the knowledge about different interpretations and their potential in mathematical thinking can also be seen as a factor of *pedagogical content knowledge* (Shulman, 1986, 1987).

## Method

### *Participants*

The data of this study was collected by a written test during the period October 2004 – March 2005. The test was arranged in six out of seven Finnish universities which give mathematics teacher education. In each university, the test was arranged in a lecture which was a part of the mathematics teacher education study program. All participants of each lecture participated in the test, and the total number of participants was 160. However, it was intended that the participants of the study had experience about mathematics studies in university and that they had encountered the concept of derivative in these studies. For that reason, answers of 14 students who had passed less than 20 Finnish credits (about 35 ECTS credits) were removed from the study, and thus the final number of subjects was 146. This is yet quite an extensive sample, because in Finland about 150–250 subject teacher students who have mathematics either as a major or minor subject graduate yearly.[3] However, it cannot be considered in any sense as a representative sample, because nothing is known about the students who did not participate to the lectures in which the test was arranged.

64 participants were male, 53 were female and in 29 cases the gender was, due to the use of pseudonyms, unknown. 89 participants were majoring in mathematics, 29 in physics and the rest in some other subject. Majority of the participants aimed for a teaching career. Statistics about the number of passed credits in mathematics and about the success in these studies are presented in the section "Relationship between students' study history and test success".

### *Measures*

The test was arranged at overseen occasions, and pen and paper were the only equipment that was allowed in the answering. The goal was not to test whether the students could remember the definitions of derivative and differentiability. For this reason, these definitions were given on the questionnaire.

The test included several tasks, but in this paper we analyze only those that measured informal and formal reasoning skills. These tasks were:

1a 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.

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

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.

For practical reasons, it was not possible to get data from the student registers of the participating universities. Thus the information concerning students' background, such as the total number of passed credits in mathematics and the study success in mathematics had to be asked in the questionnaire. The following questions were included into the questionnaire:

> The number of your passed credits in mathematics:
>
> Please describe shortly your success in your studies in university mathematics.

### *Procedure*

The answers to each of tasks 1a, 1b, 2a and 2b were graded by giving each of them 0, 1 or 2 points. The main grading principle was the following: Two points were given if the answer was fully acceptable, and one point was given if the main idea was correct but the answer did not fulfil all criteria which were required for complete acceptance. The detailed criteria for full acceptance are presented in appendix 1. All criteria were not required to come out explicitly from the answer, but the examiner tried to conclude if the student had understood the content of the criteria. The scores from tasks 1a and 2a were added together resulting in a variable which indicates the ability to argue claims informally. Respectively, the scores from tasks 1b and 2b produce a variable which indicates the ability to prove claims formally. The values of these variables vary then between 0–4.

The answers to the question concerning the study success were classified into three categories: poor, satisfactory or excellent success. If the answer contained an estimate of the average grade, this was the primary criterion in the classification. If no estimate about the average success was presented, the classification was made on the basis of verbal descriptions. The detailed criteria for this classification are presented in appendix 2.

Because the variables concerning the total number of passed credits and the study success in mathematics are based on students' own estimates, their reliability is not very good. In any case, these can be considered suggestive. As well, the number of passed credits cannot be considered as an absolute measure of how much the students have studied mathematics, because the raw number of passed credits does not reveal the composition of the studies: The courses and their contents vary between universities, and, in addition, study programs in all universities include some optional courses.

## Results

### *Students' success in the test*

**Observation 1**: *Task 2 was clearly more difficult than task 1.*
The distributions presented in table 1 verify this presupposed observation. According to Wilcoxon's signed ranks test (two-tailed), the differences between tasks 1a and 2a, between tasks 1b and 2b and, as well, between the total scores (1a + 1b) and (2a + 2b) were all very significant ($p < .001$ in all cases). In addition, it turned out that 109 out of 146 students (74.6 %) received at least one point both from task 1a and from task 1b. That is to say, in task 1, most of the students were able to present both an informal and a formal argument for the claim at least in a partially acceptable way. Instead, in the case of the second claim, 71 students (48.6 %) presented neither an informal nor a formal argument in a way which was to any extent acceptable. They received zero points both from task 2a and from task 2b.

Table 1. *Distributions (%), means and standard deviations of the received scores*

| | Points | | | | |
|---|---|---|---|---|---|
| | **0** | **1** | **2** | **Mean** | **St. dev.** |
| **Task 1a** | 11.6 | 46.6 | 41.8 | 1.30 | .67 |
| **1b** | 21.2 | 13.0 | 65.8 | 1.45 | .82 |
| **2a** | 66.4 | 19.2 | 14.4 | .48 | .74 |
| **2b** | 61.0 | 23.3 | 15.8 | .55 | .75 |
| **1a + 2a** | | | | 1.78 | 1.14 |
| **1b + 2b** | | | | 1.99 | 1.29 |

**Observation 2**: *The difference in success between the informal tasks and the formal tasks was not significant.*
As presented in table 1, for both claims the mean of the received scores from the task requiring informal argumentation (task 1a/2a) was somewhat smaller than the corresponding mean of the task requiring formal proving (task 1b/2b). According to Wilcoxon's signed ranks test, neither the difference between tasks 1a and 1b, between tasks 2a and 2b nor between the total scores (1a + 2a) and (1b + 2b) can be considered statistically significant.

**Observation 3**: *The success in the informal tasks and the success in the formal tasks were dependent.*
In the cases of both tasks, there were several students who succeeded even only in the informal part or only in the formal part of the task. However, in the cases of both tasks, the correlation (Spearman's rho) between informal and formal scores was very significant ($p < .001$). The correlation between tasks 1a and 1b was .302 and between tasks 2a and 2b .337. However, the difference between these correlations is not statistically significant. As well, the correlation between the combined scores (1a + 2a and 1b + 2b) was very significant ($p < .001$). The value of this Spearman's rho was .419. However, due to different scales of the variables, this value is not comparable with the two above mentioned values of Spearman's rho.

The dependence between the combined scores can be observed also from the crosstabulation in table 2.

Table 2. *Crosstabulation: Scores from informal tasks vs. scores from formal tasks*

| | Formal | | | | | |
|---|---|---|---|---|---|---|
| | **0** | **1** | **2** | **3** | **4** | **Total** |
| **Informal 0** | 11 | 2 | 1 | 1 | 1 | 16 |
| **1** | 9 | 9 | 22 | 7 | 5 | 52 |
| **2** | 7 | 3 | 14 | 9 | 6 | 39 |
| **3** | 1 | 2 | 11 | 7 | 5 | 26 |
| **4** | 0 | 0 | 4 | 5 | 4 | 13 |
| **Total** | 28 | 16 | 52 | 29 | 21 | 146 |

**Observation 4**: *A poor success in the formal tasks seemed to imply a poor success also in the informal tasks, but a good success in the formal tasks did not indicate a good success in the informal tasks.*
The crosstabulation between the combined scores (table 2) reveals that none of the students who received 0 or 1 points from the formal tasks received 4 points from the informal tasks, and only three of them received 3 points from the informal tasks. This suggests that students who had serious problems in the formal tasks probably had problems also in the informal tasks. On the other hand, there were together 14 students whose success in the formal tasks was good (3 or 4 points), but the success in the informal tasks was poor (0 or 1 points). In addition, the crosstabulation reveals that among the students who received 3 or 4 points from

the formal tasks, the scores received from the informal tasks were quite equally distributed between the points 1–4. Thus, a good success in the formal tasks did not seem to imply a good success in the informal tasks.

### *Relationship between students' study history and test success*

Because the validity and the reliability of the data concerning the number of passed credits and the study success are not very good (see above), the findings presented in this section can be considered only suggestive. However, they offer some baselines for the discussion about the effect of education on the informal and formal reasoning skills. In order to get more exact results, further studies based on broader and more exact data are needed.

**Observation 5**: *Both the informal scores and the formal scores were dependent on the amount of passed credits in mathematics. The dependence with respect to the formal scores was possibly stronger.*
The number of passed credits among the participants ranged between 20 and 130, the median was 45.5, the mean 49.6 and the standard deviation 20.8. All these values are in Finnish credits. Scatter diagrams in figure 1 show that the number of passed credits had a clear influence both on the score received from the informal tasks and on the score received from the formal tasks. As well, correlations (Spearman's rho) were very significant ($p < .001$) in both the cases. The values of the Spearman's rhos were .348 (informal) and .496 (formal). The scatter diagram in the case of the informal score looks more fragmented. As well, on the basis of the values of the Spearman's rho, it seems that in the case of the formal tasks the dependence on the number of passed credits is stronger. The difference

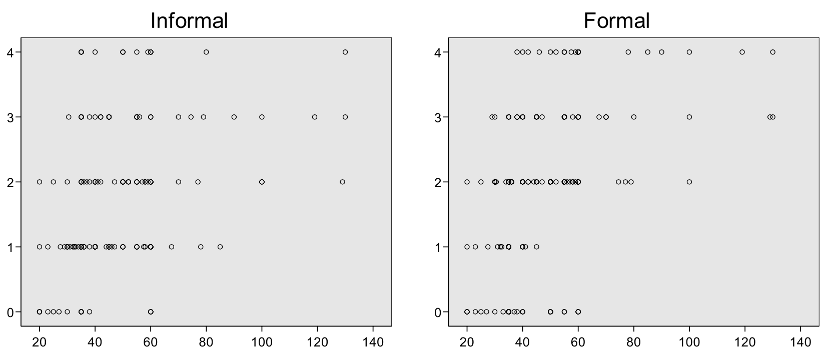


Figure 1. *Scatter diagrams describing the dependence of the received scores on the number of passed credits*

between the Spearman's rhos cannot yet entirely be considered statistically significant (p = .065). However, the size of the sample is small for the significance test.

**Observation 6**: *Both in the informal tasks and in the formal tasks, at least satisfactory study success appeared to be needed for a good success. A good study success possibly had more influence on the success in the formal tasks than on the success in the informal tasks.*
According to the criteria mentioned in appendix 2, the students in the sample were divided into classes on the basis of their study success in mathematics. 27 students (18.5 %) were placed in the class "poor success", 68 students (46.6 %) in the class "satisfactory success", and 43 students (29.5 %) in the class "excellent success". Eight students (5.5 %) could not be placed in any of these classes.

Table 3. *Crosstabulation: Scores from the informal tasks vs. estimated study success*

| | Study 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 |

The crosstabulations in tables 3 and 4 reveal one feature which is common for both cases: Only a few students with poor study success managed to receive high scores in the test. From the students with poor study success, three students (11.1 %) received 3 or 4 points from the informal tasks, and, respectively, four students (14.8 %) received 3 or 4 points from the formal tasks. Therefore, it seems that at least satisfactory study success was a prerequisite for a good success both in the informal and formal tasks. Among the students with excellent study success, only a few (six students, 14.0 %) received a poor score (0 or 1 points) in the formal tasks, but in total 18 (41.9 %) of them received 0 or 1 points in the informal tasks. These observations suggest that good study success did not guarantee a good success in the informal tasks, but it could have a stronger effect on the achievements in the formal tasks.

Table 4. *Crosstabulation: Scores from the formal tasks vs. estimated study success*

| | Study 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 |

Correlation (Spearman's rho) between the informal scores and the study success was 0.214 ($p = .012$) and between the formal scores and the study success .295 ($p < .001$). The difference between the correlations is not significant.

The crosstabulations show that both in the informal tasks and in the formal tasks some students received a high score even if their study success had been at most on the satisfactory level, and on the other hand, especially in the informal tasks, some students with excellent study success received low scores. In these cases the study success and the test results are opposite (one is poor and the other is good), and thus it is improbable that these factors have causality between them. Could the number of passed credits might be a factor explaining the test results in these cases? In order to study this question, the groups consisting of students in the lower-left and upper-right corners in tables 3 and 4 were formed, and the distributions of the number of passed credits in these groups were compared to the corresponding distribution in the whole sample. The analysis was based both on a raw (ocular) compasion of distributions and on statistical parameters concerning differences between the means and between the medians. The detailed description of this a quite complicated analysis is omitted from this paper, but two interesting findings concerning the groups in the lower-left corners are presented in the following.

**Observation 7**: *The amount of passed credits and the study success did not in all cases explain the success in the informal tasks.*
It turned out that in a group consisting of those 17 subjects who had received three or four points from the informal tasks but whose study success was poor or satisfactory (the lower-left corner in table 3), the distribution of the number of passed credits did not differ significantly from the corresponding distribution in the whole sample. Therefore, the

number of passed credits either cannot be considered as an explanatory factor for the good success in the informal tasks. This suggests that the success in the informal tasks cannot be explained necessarily at all only on the basis of the number of passed credits and the level of the study success.

**Observation 8**: *The amount of passed credits could have had an important effect on the good success in the formal tasks even though the study success had been at most satisfactory.*
It turned out that in a group consisting of those 28 subjects who had received three or four points from the formal tasks but whose study success was poor or satisfactory (the lower-left corner in table 4), the high number of passed credits is a possible explanatory factor for the good success in the formal tasks. In this group, the mean of the numbers of passed credits was almost significantly higher than the mean in the whole sample on average ($p = .023$). The difference of the medians between this group and the whole sample can be considered significant ($p = .013$). These results suggest that the amount of passed studies may have an important role in the development of formal proving skills.

## Discussion

At least for two reasons, hardly any valid conclusions about students' absolute reasoning skills or about their absolute mastery of the concept of derivative can be drawn on the basis of the above presented analysis. First, the data is contentually narrow, because students' success has been analysed only in two tasks. Secondly, it is not known how intensively the students answered to the test. The test came unexpectedly to the students, and the students were aware that the test results do not have any influence on their grades or on their life at all. However, it is more reliable to assume that the answering intensity was substantially in the same level in all tasks of the test. Thus, it was appropriate to examine the relative success between the tasks.

Epistemological differences between the tasks may at least partially explain why task 2 was more difficult for the students than task 1 (cf. observation 1). Task 1 considers a function which can be explicitly expressed by using such concrete representations as a formula or a graph. Instead, in task 2, the function has to be considered as an abstract concept. In addition, solving task 2 requires a deep understanding of the concept of the limit, whereas in task 1 no consideration of the limiting process is in fact needed at all.

The comparison between the success in the informal tasks and the success in the formal tasks revealed that a good success in the formal tasks does not necessarily imply succeeding in the informal tasks (cf. observation 4). One possible explanation for this may be the emphasis of formal elements in mathematics education both in school and in university. Informal reasoning tasks, like tasks 1a and 2a in the test of this study, are very seldom considered in textbooks or in lectures, and they are rarely included in exams. Thus it is possible that the students who succeeded well in the formal tasks but poorly in the informal tasks may have found the form of the question in the informal tasks odd, whereas the formal proving tasks have probably been more familiar to them. On the other hand, it appeared that difficulties in the formal tasks imply difficulties also in the informal tasks, and, in other words, success in the informal tasks implies success also in the formal tasks. One might interpret these findings to mean that only those students whose skills are throughout at a high level are capable to present acceptable informal arguments. However, Hähkiöniemi (2006a, 2006b) has shown that the learning of derivative can very well begin from informal elements by perceiving the increase, steepness and local straightness of a graph of a function. Therefore, it is not justified to consider informal reasoning skills as the most advanced achievement of the learning process. More like, a deep understanding of the connections between informal and formal representations could be the punchline of the learning process. Especially, in task 2, the claim was presented in a formal form, and thus in the informal argumentation the claim had to be translated to the informal form. This required quite a deep understanding of the connections between informal and formal representations. Furthermore, it is obvious that the deep understanding of the connections between informal and formal representations requires a mastery of these both representations. This may explain why a good success in the informal tasks seemed to require a good success also in the formal tasks.

In university-level studies in mathematics in Finland, the derivative of the real-valued function of a single variable is taught mainly at the beginning of the studies. This happens typically before passing 20 credits in mathematics. Therefore, on the basis of the findings about the effect of the number of passed credits, it seems that the general training of mathematics, which concerns also other issues than derivative, has a notable effect on the ability to reason claims concerning derivative. One could assume that students having, for example, 20–30 credits in mathematics would remember better the things which they have learnt about derivative than the students who are further in their studies, because a shorter time has probably passed since their studies of derivative. The results of this study are contrary to this speculation: Succeeding in the tasks clearly increased when the number of passed credits increased. The growth of general mathematical skills might be an important factor explaining this. Giving the definition of derivative on the questionnaire decreased the effect of memory. In addition, it is probable that the students had later in their studies met with situations where they had needed derivative: For example, many of the students who had a lot of credits in mathematics had probably passed a vector-calculus course, and maybe some students had taught the concept of derivative during teacher training.

It is not surprising that both succeeding in the informal and succeeding in the formal tasks showed to depend both on the number of passed credits in mathematics and on the study success in mathematics (cf. observations 5 and 6). However, in several connections during the analysis, there appeared the impression that these factors have a stronger effect on the formal than on the informal reasoning skills. Certainly, on the basis of this study, these conclusions can be considered only suggestive. However, these conclusions are understandable due to the different nature of informal and formal reasoning. Formal reasoning is technically often more demanding than informal reasoning: It requires skills to carry out calculations and other technical procedures. This, for one, demands the correct and fluent use of mathematical symbols and mathematical language, and, in addition, the structure of mathematical knowledge has to be internalized. These are skills which an individual can hardly learn without training and on whose learning education has therefore a crucial effect. Instead, informal reasoning does not demand as much procedural skills but it is based on holistic understanding, and several factors may have influence on its development. Especially, it is essential how informal elements like visualization are emphasized in teaching. Undoubtedly, formal proving has a central role almost in all teaching at the university level, but it crucially depends on the lecturer how the informal reasoning is considered in teaching. The data used in this study reveal nothing about that. In addition, factors which are independent of the education, such as students' personal interests and talents, may have influence on the development of the informal and the formal reasoning skills. Explorative qualitative studies would be needed to reveal these factors and their effects.

Because the informal reasoning has an important role in creative mathematical thinking, it is also important that its role is emphasized in teaching of mathematics. Several researchers, for instance, Arcavi (2003), Dreyfus (1994) and Rodd (2000), have proposed that especially development of the visual reasoning should be seen as an intended goal of mathematics education. However, it is important as well that the informal reasoning is not taught in isolation from the formal reasoning. On the contrary, the connections between them should be especially emphasized. In teacher education students should first reflect their own mathematical reasoning and recognise explicitely the importance of the informal elements in it. After that they may develop methods to give support to students' reasoning. Nowadays the information technology offers various resources to teaching of informal reasoning. It is yet important that these resources are utilized in a pedagogically appropriate way. Therefore, pedagicical frameworks and research concerning these issues are needed.

## References


Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, 52 , 215–241.

Dreyfus, T. (1994). Imagery and reasoning in mathematics and mathematics education. In D.F. Robitaille, D.H. Wheeler & C. Kieran (Eds.), S*elected Lectures from the 7th International Congress on Mathematical Education* (pp. 107–122). Quebec: Les Presses d l'Universit´e Laval.

Dreyfus, T. (2000). Some views on proofs by teachers and mathematicians. In A. Gagatsis (Ed.), *Proceedings of the Second Mediterranean Conference on Mathematics Education, Vol. I, Nikosia, Cyprus* (pp. 11–25). The University of Cyprus.

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

Hanna, G. (1991). Mathematical proof. In Tall, D. (Ed.), *Advanced mathematical thinking* (pp. 54–61). Dordrecht: Kluwer.

Harel, G. & Sowder, L. (1998). Students' proof schemes: results from exploratory studies. In A. H. Schoenfeld, J. Kaput & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234–283). Providence, RI: American Mathematical Society.

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.

Parameswaran, R. (2007). On understanding the notion of limits and in-finitesimal quantities. *International journal of Science and mathematics Education,* 5 , 193–216.

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

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.

Shulman, L. S. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15 (2), 4–14.

Shulman, L. S. (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review*, 57 (1), 1–22.

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.

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

Weber, K. & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, 56 , 209–234.

## *Notes*

1 In practice, most of the proofs made by professional mathematicians include several informal elements, and the social convincingness among other mathematicians is the crucial factor in the acceptance of a proof (Hanna, 1991; Dreyfus, 2000).

2 The term warrant does not have the same meaning here and in Toulmin's model: In Toulmin's model, a justification defined according to Rodd could also be treated as a warrant.

3 In Finnish universities, the subject-teacher students have one subject as a major subject and the other subjects are minor subjects. Usually, the major subject is one of the subjects taught in school, but in some universities, the science of education may also be the major subject. Students majoring in mathematics are required to pass 140 ECTS credits in mathematics, and students who have mathematics as a minor subject have to pass 60 ETCS credits in order to receive the qualifications of a subject teacher in mathematics.

## Appendix 1

Solutions for the tasks in the test and evaluation criteria for the students' answers

1a *How would you explain, by using graphical interpretations, why the derivative of a constant function is equal to zero everywhere?*

Criteria for two points:

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.

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.

A reasonable conclusion based on criteria 1 and 2 which justifies the claim had to be presented.

Other observations:

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:

A correct formal definition for the constant function had to be presented.

The proof had to be based on the definition of derivative.

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:

The key argument had to be based on the definition of derivative.

The deficiencies could appear in the calculation of the limit or with respect to generality (criteria 3 in the above list).

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

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.

By using the chosen interpretation, the limiting process had to be explained.

By using the chosen interpretation, the state after the limiting process had to be explained.

The derivative had to be interpreted either as the steepness of a tangent line or as the (instantaneous) rate of change.

The structure of the claim had to be coherent.

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

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

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

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

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

Criteria for two points:

The definition of derivative had to have a key role in the answer.

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:

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

## Appendix 2

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 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.

At that time in Finland, accepted studies at university were graded by using the scale from 1 to 3. One step in this scale was one fourth of a point.

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.
  (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.
  (The original Finnish expressions were "kohtalainen", "kohtuullinen", "keskitasoinen" and "tyydyttävä")
- 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.
  (The original Finnish expressions were "hyvä", "melko hyvä", and "ihan ok")

Antti Viholainen

Antti Viholainen is a PhD candidate at University of Jyväskylä, Finland. Since the year 2003, he has worked as a researcher in the area of mathematics education. His main interests are the nature of mathematical knowledge, informal and formal reasoning and the coherence of mental structures. In his studies, he has especially concentrated on the informal and formal representations regarding the concept of derivative.

Antti Viholainen
Department of mathematics and statistics
University of Jyväskylä
Finland
anvihola@maths.jyu.fi

# Sammanfattning

I denna studie undersöktes den formella och informella argumentationsfärdigheten hos 146 finska blivande ämneslärare i matematik. Studenterna deltog i ett test, där de skulle argumentera både formellt och informellt för två påståenden gällande derivatan. Resultaten visar, att framgången i formella uppgifter och i informella uppgifter beror av varandra. Studenterna kunde emellertid klara sig bra i formella uppgifter utan att ha likadan framgång i informella.

Resultaten i båda uppgiftstyperna var beroende både av mängden av och framgången i tidigare matematikstudier. Dessutom kan de här faktorerna ha haft en starkare effekt på de formella argumentationsfärdigheterna.