# Students' Feedback on Teaching Mathematics Through The Calculational Method 

João F. Ferreira and Alexandra Mendes<br>School of Computer Science, University of Nottingham, joao@joaoff.com, afm@cs.nott.ac.uk


#### Abstract

This paper describes a study conducted at the University of Nottingham, whose goal was to assess whether the students registered on the first-year module "Mathematics for Computer Scientists" appreciate the calculational method. The study consisted of two parts: "Proof Reading" and "Problem Solving". The goal of "Proof Reading" was to determine what the students think of calculational proofs, compared with more conventional ones, and which are easier to verify; we also assessed how their opinions changed during the term. The purpose of "Problem Solving" was to determine if the methods taught have influenced the students' problem-solving skills. Frequent criticisms of our approach are that we are too formal and that the emphasis on syntactic manipulation hinders students' understanding. Nevertheless, the results show that most students prefer or understand better the calculational proofs. On the other hand, regarding the problemsolving questions, we observed that, in general, the students maintained their original solutions.


Index Terms - Calculational Method, Computer Science Education, Problem Solving, Teaching Mathematics.

## I. INTRODUCTION

Reasoning rigorously and effectively in mathematics is essential for studies in computing and engineering. However, most students are not adequately skilled in formal reasoning and proof. As a result, they have difficulties in employing mathematics to solve new problems. We believe that the state of affairs can be improved by using the calculational method [1], where most proofs are reduced to elementary syntactic manipulation.

Calculational proofs are usually written using a uniform format, where each step is accompanied by a hint justifying the validity of the step. For example, we write

$$
\begin{aligned}
& A \\
= & \{p\} \\
= & \{q\}
\end{aligned}
$$

to prove that $\mathrm{A}=\mathrm{C} . \mathrm{A}, \mathrm{B}$ and C are expressions, and p and q are hints why $A=B$ and $B=C$, respectively. Some relevant advantages of this format are that the hints reduce the search space, there are no repeated intermediate expressions, and we can immediately conclude that $A=C$, just by looking at the first and last expressions and at the relations connecting
them. Another important aspect of this format is that it forces the writer to provide explanation for each step.

Although calculational proofs are widely used by computer scientists, and some of them even use them to teach mathematics [2] [3], there is little information on what students think of them, especially when compared with traditional proofs.

In this paper, we describe a study conducted at the University of Nottingham, whose goal was to assess whether the students registered on the first-year module "Mathematics for Computer Scientists" appreciate calculational proofs.

The only related studies that we are aware of were recently done in Turku, Finland, and show that a calculational approach can be beneficially used in highschool and university [4]. Although our experiment is done on a much smaller scale, we believe that it complements their studies by asking the students to compare traditional and calculational proofs and by studying how their opinions change during the term.

The details of the study and the results are discussed in Section II. Frequent criticisms of our approach are that we are too formal and that the emphasis on syntactic manipulation hinders students' understanding. Nevertheless, the results show that most students prefer or understand better the calculational proofs. For instance, we were surprised to observe that, in the first coursework, when the students had no prior knowledge of the calculational method and notations, more than two thirds of them preferred the calculational proof to its traditional counterpart. On the other hand, regarding the problem-solving questions, we observed that, in general, the students maintained their original solutions. In the conclusion, we elaborate on these results and we discuss some future work.

## II. DESCRIPTION OF THE STUDY and Results

The study was conducted at the University of Nottingham with the students registered on the first-year module "Mathematics for Computer Scientists" (G51MCS) in the academic year 2008/2009.

The module is targeted at undergraduate students in the School of Computer Science and in the School of Mathematics, and it covers basic concepts in mathematics of relevance to the development of computer software (Boolean algebra, simple number theory, sets, functions and relations, quantifiers, and simple induction on natural numbers).

There were a total of 135 students registered on the module. They had two lectures per week, associated coursework and weekly tutorials. Their feedback was collected through supplementary questions included in seven of the nine courseworks released. The participation was on a voluntary basis, but the students would have extra marks if they expressed their opinions.

The study consisted of two parts: "Proof Reading" and "Problem Solving". In "Proof Reading", we have shown calculational and traditional proofs for the same theorem and we have asked the students which one they preferred. We repeated the same questions later in the term to measure how their opinions changed. In "Problem Solving", we have asked them to solve the same problems at the beginning and later in the term, so that we could compare the solutions and determine if our methods have influenced them.

## A. Part 1 - Proof Reading

1) Research Questions: The goal of the "Proof Reading" part was to determine what the students think of calculational proofs, compared with more traditional (and informal) ones. In particular, we have addressed the following research questions:

- How do students react when confronted for the first time with calculational proofs and with its proof format?
- Which type of proofs do the students prefer or understand better?
- How did students' opinions change during the term?
- Which type of proofs is easier to verify and in which one is it easier to detect mistakes?

2) Study Description: In total, we have shown the students three different propositions, each with two different proofs: a calculational and a more conventional one. The question was the same for the three propositions:

For the following theorem, we show you two different proofs. Read both of them carefully and say which one you prefer or understand better. Justify your answer. If there is any step that you don't understand, please mention that in your answer. You are encouraged to suggest comments and improvements to the proofs presented.

The theorem and proofs given in coursework 1 are shown in Figure 1. We have repeated the same question in coursework 6. Proof 0 was taken from a chapter on mathematical proofs of a secondary school book [5]. Proof 1 is essentially the same proof, but rewritten in the calculational format described in the introduction. Note that the calculational format forces the writer to provide explanations for all the steps; that is why Proof 1 has more hints.

In coursework 2, we have shown the proofs for the irrationality of $\sqrt{2}$ depicted in Figure 2. Proof 0 was taken from [3] and we believe it is a much more goal-oriented and motivated proof than the conventional proof by contradiction. We did not repeat the same question; instead, for coursework 4, we changed it and we gave the false proposition that $\sqrt{4}$ is not a rational number, together with the same proofs used in coursework 2, but with all the
relevant occurrences of 2 replaced by 4 (the calculational proof uses the wrong assumption that exp. $4=1$; the conventional one wrongly assumes that q is even). The goal of this "trick question" was to assess which type of proof is easier to verify.

```
Theorem: If \(\sqrt[n]{a}>1\) then \(a>1\).
- Proof 0
    If \(\sqrt[n]{a}>1\), then \(\sqrt[n]{a}=1+h\), with \(h>0\). Therefore,
    \(a=(\sqrt[n]{a})^{n}=(1+h)^{n}=1+{ }^{n} C_{1} h+{ }^{n} C_{2} h^{2}+\ldots>1+n h \geq 1\).
- Proof 1
    If \(\sqrt[n]{a}>1\), then \(\sqrt[n]{a}=1+h\), with \(h>0\). Therefore,
            \(a\)
        \(=\quad\left\{\quad a=(\sqrt[n]{a})^{n} \quad\right\}\)
        \((\sqrt[n]{a})^{n}\)
    \(=\{\quad\) Assumption: \(\sqrt[n]{a}>1\), thus \(\sqrt[n]{a}=1+h\),
                with \(h>0 \quad\}\)
        \((1+h)^{n}\)
    \(=\{\) We use the binomial theorem:
                        \((a+b)^{n}={ }^{n} C_{0} a^{n}+{ }^{n} C_{1} a^{n-1} b+\ldots+{ }^{n} C_{n} b^{n}\),
            with \(a:=1\) and \(b:=h \quad\}\)
            \(1+{ }^{n} C_{1} h+{ }^{n} C_{2} h^{2}+\ldots+{ }^{n} C_{n} h^{n}\)
    \(>\quad\left\{\quad{ }^{n} C_{m}>0\right.\) and \(\left.h>0 \quad\right\}\)
        1 .
                                    FIGURE 1
        THEOREM AND PROOFS GIVEN IN COURSEWORKS 1 AND 6
```

3) Sample: In total, 115 students answered the questions of courseworks 1 or 6: 28 students answered only coursework 1, 18 only coursework 6, and 69 both. So, there were 97 students answering the question in coursework 1 and 87 answering the question in coursework 6.

Regarding the other two courseworks, 123 students answered the question of coursework 2 and 100 answered the question of coursework 4.
4) Results:
a) How do students react when confronted for the first time with calculational proofs and with its proof format?: We have categorized the answers to coursework 1 as depicted in Figure 3. Surprisingly, most students (67, 69\%) preferred or understood better the calculational proof (e.g. "I prefer proof 1 , because it clearly highlights what has been used, which is not as obvious in proof 0. ."). Twenty-six students (27\%) preferred the conventional proof (e.g. "I prefer Proof 0, because it keeps explanation to a minimum and it is easier to understand the steps."). Four students (4\%) did not understand the proof or did not express their preference (e.g. "I don't understand binomial expansion. I do not like or understand either proof.").

Frequent reasons for preferring the calculational proof included expressions as "more detailed", "steps more explained", "more understandable format", and "easier to understand". Regarding the conventional proof, the reasons included "easier to understand", "clearer" and "short".


Thus $p^{2}$ is even. The only way this can be true is that $p$ itself is even. But then $p^{2}$ is actually divisible by 4 . Hence $q^{2}$ and therefore $q$ must be even. So $p$ and $q$ are both even which is a contradiction to our assumption that they have no common factors. The square root of 2 cannot be rational!

FIGURE 2
THEOREM AND PROOFS GIVEN IN COURSEWORK 2
Clearly, the results show that the students had no problems with the calculational format. Although they have never used or seen it before, most of them preferred it to the conventional style. The main reason for their preference was the use of hints explaining the relation between the steps.
b) Which type of proofs do the students prefer or understand better?: According to the results obtained in coursework 1, we would expect the students to prefer the calculational proof shown in coursework 2. However, 77 students (63\%) preferred the proof by contradiction (e.g. "(...) The reason I prefer Proof 1 is because although it is quite short it gets straight to the point. Proof 0 drags on and begins to talk about exponential functions, which to be honest, confuses me. (...)").


Forty students (32\%) preferred the calculational alternative (e.g. "I understand Proof 0 more than Proof 1. This is because proof 1 is too short and doesn't explain fully each step like proof 0 does. Even though Proof 0 brings in other functions, it explains why they are being used and the properties that the function has. (...) Overall even though Proof 0 is longer it is easier to understand and doesn't make any assumptions."). Six students (5\%) did not express any preference (e.g. "I totally understood Proof 0, due to the use of extended English in order to explain the steps being taken. I also understand Proof 1."). Figure 4 depicts these results.


The difference between the results of coursework 1 and coursework 2 show that the preferences of the students depend on the proof being considered. The major problem with the calculational proof was the use of the exp function, which confused the students. We also believe that the implication step (second step) may have caused difficulties.

In section II.A.4.d, we discuss some results that may indicate that a significant number of the students who preferred the proof by contradiction may not have understood it properly.
c) How did students' opinions change during the term?: To answer this question, we have analysed the answers of the students who have answered both courseworks 1 and 6 . We have classified the answers as depicted in Figure 5.


Out of the 69 students that answered both courseworks, 46 (67\%) maintained the calculational proof as their favourite, and 5 (7\%) changed their preference from the calculational to the conventional proof. Some of the answers of these students who changed opinion are confusing and

October 18-21, 2009, San Antonio, TX
39 ${ }^{\text {th }}$ ASEE/IEEE Frontiers in Education Conference
contradictory; for example, one student wrote in coursework 1 that "In proof 0, I don't understand after ( $1+h$ ). I understand proof 1. ., but in coursework 6 the same student wrote "I prefer proof 0 because it is clearer. Proof 1 is complicated, I don't understand from the second step. It is confusing.".

Five students (7\%) kept the conventional proof as their favourite, and 11 (16\%) changed their preference from the conventional to the calculational version. Most of these students who changed opinion justify the change with proof 1 having hints between the steps. For example, one student wrote in the first coursework "I prefer Proof 0 because it is brief and concise", and he used conciseness as a reason to choose proof 1 in coursework 6: "Proof 0 is more concise than Proof 1. Proof 1 contains many explanations unlike Proof 0 that contains none. I prefer Proof 1 better because in the event that an error occurred, the error could be easily seen and corrected. Also by making the calculations explicit another person can easily read and understand the solution of a problem.".

There were 2 students (3\%) who did not express any preference at all.

We see that from the 51 students that initially preferred the calculational version, 5 (10\%) changed opinion; this is a relative small number of students, when compared with the 11 (69\%) out of the 16 students that changed their preference from the conventional to the calculational proof.

However, considering the results obtained in coursework 2, we know that their preferences depend on the examples. We address again this issue in the conclusion.
d) Which type of proofs is easier to verify and in which one is it easier to detect mistakes?: Recall that, to answer this question, we have included a false proposition in coursework 4, together with two incorrect proofs. Surprisingly, 74 (74\%) students did not detect any mistake; moreover, 44 (59\%) of these 74 preferred the conventional proof, and 30 (41\%) the calculational one.

Figure 6 illustrates how the 26 (26\%) students that detected the mistake are divided.


The details are given below:

- 9 (34\%) preferred the calculational proof; 8 of them detected the mistake only in the calculational version (e.g. "I prefer the set up of Proof 0 , however I don't understand one step (that exp. $4=1$ ). Although I do understand proof 1, I would not personally work it out that way."); one student detected the mistake in both.
- 8 (31\%) preferred the traditional proof, but all of them detected the mistake only in the calculational proof (e.g.
"Proof 1 is by far the easier to understand, looking at Proof 0 I can't even see that the answer is correct as exp. $4+$ exp. $q^{2}$ does/can not equal $1+2 \times$ exp.q?").
- all the others did not express any preference:
- $2(8 \%)$ detected the error in both proofs;
- 2 (8\%) detected the error in the calculational proof;
- $2(8 \%)$ detected the error in the traditional proof;
- $\quad 3$ (12\%) realized that the proposition was false and did not comment on the proofs (e.g. "Theorem is incorrect. Because $2^{2}=4$, so $\sqrt{4}$ must be a rational number.").
Note that there were no students detecting the error in the conventional proof and choosing it as their favourite. Moreover, only 2 (8\%) students detected the error in the traditional proof, against 18 (69\%) who detected it in the calculational one. These results strongly suggest that the calculational format is better to verify proofs and to detect mistakes. The reason is clear: since all the steps are justified and are easily identified, the reader can easily check each step individually.

We now show two excerpts of the answers of two different students showing that the informal style of the proof by contradiction obfuscates the error. The first student prefers the proof by contradiction (failing to see that it is wrong), because the calculational proof has an error:
"I think there is a mistake in it [Proof 0] when: 'Let exp.k be the number of times that 2 divides k. (e.g. $\exp .8=3$ )’ it is then applied to the left hand side $\exp .4+\exp . q^{2}$ to give $1+2 \times$ exp.q but 2 goes into 4,2 times not one. It should be 'let exp.k be the number of times 4 divides k' then it would be correct. Given the above I prefer proof 1. It's short, takes logical steps that are well defined."

The second student also detects the error in the calculational proof, but he writes that the informal proof "clearly shows that $\sqrt{4}$ is not a rational number":
" (...) However I believe that $\sqrt{4}$ is a rational number because from the Proof 0 you can clearly see that: [student shows calculation with exp. 4 replaced by the correct value,
Although it is very risky to challenge the proofs, I believe that Proof 1 is actually correct. Making my life harder, I have concluded that Proof 0 actually might show that $\sqrt{4}$ is a rational number and Proof 1 clearly shows that $\sqrt{4}$ is not a rational number."

These results also suggest that a significant number of the students who preferred the proof by contradiction in coursework 2 (see section II.A.4.b) may not have understood the proof properly.

## B. Part II - Problem Solving

1) Research Questions: The goal of the "Problem Solving" part was to assess how the students' problemsolving skills changed during the term. We have addressed the following research questions:

- How was the students’ ability to solve problems at the beginning of the term?
- Did the students' ability to solve problems change during the term?
- Did the students adopt what was taught in the module to solve the problems?

2) Study Description: We have asked the students to solve three different problems. All the problems were prefaced by the following text:

Please try to solve the following problem, showing all the steps that justify your answer. The problem can be solved in different ways and it is important for us that you make a serious attempt to solve it. Remember that we are just assessing your problem-solving skills.
The problem that we gave in coursework 1 was:
Is the sum of all the numbers $n$, where $1982 \leq n \leq 2008$, an even or an odd number?
We have repeated the same problem in coursework 7. We included this problem because it can be solved by a simple counting argument that follows from a calculational rule that we teach. The rule is that the predicate even distributes over Boolean equality, i.e.:

$$
\text { even. }(\mathrm{m}+\mathrm{n}) \equiv \text { even. } \mathrm{m} \equiv \text { even. } \mathrm{n}
$$

(even.n evaluates to true if n is even and it evaluates to false if it is odd.) A consequence of this rule is that a sum of numbers is even if the number of odd summands is even.

In coursework 2, we asked the students to solve the following logic puzzle:

The island of knights and knaves is a fictional island that has two types of inhabitants: 'knights', who always tell the truth; 'knaves', who always lie. Three of the inhabitants$\mathrm{A}, \mathrm{B}$, and C -were standing together in the garden. A stranger passed by and asked A, "Are you a knight or a knave?". A answered, but rather indistinctly, so the stranger couldn't make out what was said. The stranger then asked B, "What did A say?". B replied, "A said that he is a knave". At this point the third, C, said "Don't believe B; he's lying!". The question is, what are B and C?

We repeated the same problem in coursework 9. We use logic puzzles as examples for demonstrating the effectiveness of the calculational logic: we replace the traditional case analysis by straightforward calculations. We included this problem to assess whether the students consider the calculational alternative effective.

The third problem was posed in coursework 5:
Prove that the product of four consecutive non-zero numbers cannot be the square of an integer. Hint: the property that there are no two consecutive non-zero squares can be useful.

We did not repeat this problem. We included it to determine if, at the middle of the term, the students would use the calculational format to record their proofs.
3) Sample: In total, 117 students tried to solve the problems of courseworks 1 or 7: 37 tried only the problem in coursework 1, 8 only the one in coursework 7, and 72 tried both. So, there were 109 students trying to solve the problem in coursework 1 and 80 attempted to solve the one in coursework 7.

The problem of coursework 2 was attempted by 124 students and from these only 12 attempted the problem again in coursework 9.

Finally, there were 48 students that tried the third problem in coursework 5.
4) Results:
a) How was the students' ability to solve problems at the beginning of the term?: To answer this question, we have analysed the students' solutions to courseworks 1 and 2. In coursework 1, 88 ( $81 \%$ ) students gave the right answer: 31 used a counting argument, 23 used the formula for the arithmetic series, 18 added all the numbers, and 19 gave an unsatisfactory argument (e.g. " $1982+\ldots+2008$ is an odd number"). Also, 18 students (16\%) solved it incorrectly (e.g. "The sum of numbers will be even because there are more even numbers in the given range; the probability of even numbers will be higher.") and 2 (3\%) were not able to answer (e.g. "I don’t know.").

In coursework 2, 87 (69\%) students solved the problem correctly, 37 (29\%) solved it incorrectly, and 2 (2\%) did not know how to solve it.

Of the 37 students who solved the problem incorrectly, most of them (14, 38\%) were not able to determine the value of the variables (e.g. "If B is a knave then C is a knight. If C is a knave then B is a knight."); 13 (35\%) gave the right answer, but a wrong justification (e.g. " B is a knave, C is a knight. This is because like the problem states, knights always tell the truth. C has not been asked any question, but instead C just came out with "Don't believe B, he's lying". So therefore C has to be a knight for pointing out the truth!"); 6 (16\%) gave a wrong answer; and 4 (11\%) just gave the final answer, with no explanation.

The results clearly show that, in general, the students were able to solve the problems.
b) Did the students' ability to solve problems change during the term?: To answer this question, we have checked how the answers changed from coursework 1 to coursework 7 and from coursework 2 to coursework 9.

Although most of the students maintained their initial solutions, a significant number of them improved. Out of the 72 students who answered both courseworks 1 and 7, 45 (63\%) maintained their initial solution, 21 (29\%) improved their solution, while 6 (8\%) had a worse solution in the end (Figure 7).

Of the 12 students who answered both courseworks 2 and 9, 7 (58\%) maintained their initial solution, 3 (25\%) improved, and 2 (17\%) worsened (Figure 8).

The results suggest that the students' ability to solve problems improved. In the following paragraph, we give more details on the improvement.


Evolution of the students' answers (Coursework 1 TO 7)

c) Did the students adopt what was taught in the module to solve the problems?: Out of the 21 students who improved from coursework 1 to coursework 7, 17 (81\%) changed their initial solution to a correct counting argument, which suggests that they used what was taught during the term. On the other hand, in coursework 7, only 4 (6\%) students used the calculational format for solving the problem.

Regarding the problem in coursework 5, most of the students (33, 69\%) gave a wrong answer; only 7 (14\%) answered it correctly. Eight students (17\%) did not know how to solve it. Of the 33 that failed to solve the problem, 13 (39\%) used (in one way or another) the calculational format, with hints justifying the steps; of the 7 who solved the problem, 4 (57\%) used the calculational format.

Finally, in coursework 9, only one student tried to solve the problem calculationally and he failed. All the other students provided an informal argument, which clearly suggests that they did not find our calculational solutions to logic puzzles good enough.

Overall, we can say that they have adopted what was taught in the module. However, their use of the calculational style was not effective enough.

## III. CONCLUSION

The results shown above indicate that, in general, the students preferred the calculational format. However, according to the results of coursework 2, it is not enough to provide detailed hints; we also have to explain and motivate the techniques that we use. We believe that providing more motivation for the introduction of the function exp in coursework 2 would have changed the results.

The results also indicate that the calculational format is better for detecting errors in arguments. In particular, the results of coursework 4 (and the excerpts shown) suggest that informal proofs by contradiction may obfuscate the error. In fact, we are led to believe that a significant number of students who preferred the proof by contradiction did not understand it.

The relatively large number of students changing their preference from the conventional to the calculational proofs also shows that, as the students got more familiar with the calculational format, they found it better. Indeed, we think that the students would be prepared to switch to the calculational format if given more practice.

Regarding the part on problem solving, we see that most students had no difficulties solving the first two problems. Even so, a significant number of them improved their solutions in their second try. Generally, they have adopted what was taught in the module for solving problems, but their use of the calculational style was not effective.

In the future, we plan to do further studies at the university level and in Portuguese secondary schools. We are currently creating packages of teaching scenarios, which are fully worked out solutions to problems together with detailed guidelines, to use in these future experiments. Finally, we are also developing a structure editor of handwritten mathematics [6] that supports a calculational approach to teaching mathematics.

## Acknowledgment

We would like to thank to all the students who participated in the experiment and to Professor Roland Backhouse, the module convenor, for letting us do the study and for his feedback. This research was supported by FCT (Portuguese Foundation for Science and Technology) with grants SFRH/BD/24269/2005 and SFRH/BD/29553/2006, and done in the context of the MathiS project (contract PTDC/EIA/73252/2006).

## References

[1] Backhouse, Roland, D. Gries, W. M. Turski, and F. Dehne "Special Issue on the Calculational Method", Information Processing Letters, Vol. 53, No. 3, 10 February 1995, pp. 121-172
[2] Gries, David and Fred B. Schneider, "Teaching Math More Effectively, Through Calculational Proofs", The American Mathematical Monthly, Vol. 102, No. 8, October 1995, pp. 691-697
[3] Backhouse, Roland, "Program Construction", John Wiley and Sons, Inc., 2003.
[4] Back, Ralph-Johan, Linda Mannila, Patrick Sibelius and Mia Petomälki "Structured Derivations: a Logic Based Approach to Teaching Mathematics" In FORMED 2008: Formal Methods in Computer Science Education. 29 March 2008.
[5] Lima, Yolanda and Francelino Gomes, "Xeqmat - Matemática $12^{\circ}{ }^{\circ}$, Editorial O Livro, 1999, p. 383
[6] Mendes, Alexandra, "Work in Progress: Structure Editing of Handwritten Mathematics" In Frontiers in Education Conference 2008. 22-25 October 2008.

## AUTHOR INFORMATION

João F. Ferreira, PhD Student, University of Nottingham, joao@joaoff.com

Alexandra Mendes, PhD Student, University of Nottingham, afm@cs.nott.ac.uk

