Spring 2011

Teaching and learning of proof in the college curriculum

Maja Derek
San Jose State University

TEACHING AND LEARNING OF PROOF IN THE COLLEGE CURRICULUM

A Thesis

Presented to

The Faculty of the Department of Mathematics

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Maja Derek

May 2011

The Designated Thesis Committee Approves the Thesis Titled

TEACHING AND LEARNING OF PROOF IN THE COLLEGE CURRICULUM

by

Maja Derek

APPROVED FOR THE DEPARTMENT OF MATHEMATICS

SAN JOSÉ STATE UNIVERSITY

May 2011

Dr. Joanne Rossi Becker	Department of Mathematics
Dr. Richard Pfiefer	Department of Mathematics
Dr. Cheryl Roddick	Department of Mathematics

ABSTRACT

TEACHING AND LEARNING OF PROOF IN THE COLLEGE CURRICULUM

by Maja Derek

Mathematical proof, as an essential part of mathematics, is as difficult to learn as it is to teach. In this thesis, we provide a short overview of how mathematical proof is understood by students in K-16. Furthermore, we answer questions about mistakes students usually make in the transition period from high school to college in understanding mathematics and mathematical proof. Through a case study, we learned that deduction mistakes characteristic for early mathematical education, such as arguing from an example, can be abandoned very easily as students begin to understand the inadequacy of one, or finitely many, examples when arguing about a general mathematical claim. Furthermore, students accept basic procedures and different methods of proof, but they experience difficulties when faced with new or complicated mathematical topics to prove, such as those concerning the floor function introduced during the proof teaching sessions. Also, we observe the students' progress during the teaching sessions for a specific proving method. Finally, we discuss grounds for further investigation about learning and teaching mathematical proof. For example, introduced are ideas of how to alter research instruments and/or modify the group studied to be able to answer more specific questions about mathematical proof in the college curriculum.

DEDICATION

This work is dedicated to my family. To Lovro and Franka who inspired and motivated me, giving me strength not to give up when the times were hard. To Ante for support, for believing in me, for sharing my hopes and dreams.

ACKNOWLEDGEMENTS

I would like to thank my mentor, Dr. Joanne Rossi Becker, for support and guidance through the years of study.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1

INTRODUCTION

This thesis examines the concepts of teaching and learning of mathematical proof in the college curriculum. To emphasize the importance of studying mathematical proof at the college level, as opposed to mathematical proof in K-12 education, we provide a summary of how proof is defined and understood through grades K-16. Also, the evolution of our current understanding of mathematical proof in K-16 is outlined.

In Chapter 3, we focus our attention college-level proof and look deeper into the college curriculum examples used to teach mathematical proof. The role of mathematical proof at the college level is to use it as a tool to validate mathematical conjectures but also to develop the sense of why something is true. Proof is the only tool of validation in mathematics that is accepted by mathematical society, and as such every science student should realize its importance.

Following the theoretical discussion of mathematical proof in the college curriculum is the case study described in Chapter 4. The case study addresses students' transition from high-school to a more formal, college level and rigorous way of understanding and constructing a proof. The focus of the case study is on the mistakes students make while attempting to construct a valid argument. We categorize those mistakes and procedures students employ using language and notation adopted from their textbook (Epp, 2004).

In Chapter 5, we present results from each data instrument used in the case study: pre-teaching and post-teaching questionnaires; two quizzes; and the midterm exam. The results are presented in two ways: tables showing percentage of correct answers and tables with listed mistakes accompanied by the percentage of occurrence of each mistake.

In Chapter 6, we discuss the results, concentrating on the meaning of the mistakes. Also, in this part of the thesis, we give examples of students' actual work to provide for better understanding of the classification of the mistakes as well as to visualize students' reasoning processes.

Finally, in the last Chapter 7, we report on the data collected during the case study and compare our observations to the norms and standards students should meet during early college mathematics education. As a part of that last chapter, we also state questions that, if studied further, would provide a better understanding of students' proof and the comprehension of proving methods. Finally, suggested are some directions for improvement of instruction of proof at the college level.

CHAPTER 2

LITERATURE REVIEW ON MATHEMATICAL PROOF K-12

2.1 What is mathematical proof?

"Prove it!" is a very common phrase in every day speech, and we all use it when asking for an explanation or a validation of an argument. On the other hand, almost no one thinks of it as something special or worthwhile to study.

What do we actually mean when asking that something be proved? Usually, we just ask for evidence. Very often we are convinced by one fact or one occurrence or just one person exhibiting certain behavior or experiencing a certain phenomenon.

On the other hand, mathematical proof is much more than just one or a few examples supporting a mathematical statement. In sciences such as biology, medicine and chemistry, experiments and their outcomes are the primary methods of validating an hypothesis, but still scientists look for explanations of why something happens and what elements could change the result. In such practice, numerous examples and experiments are necessary before establishing some process to be understood. Often we rely on statistics, though we could never be 100% certain of something. That is why it is common to hear about 95% confidence intervals, or if an event occurs there is a 75% chance of it occurring again, and similar statements. Such statements can be very misleading since there are many hidden variables that we do not know about, such as the size of the sample or how the sample was chosen. An example is building statistics on a sample with a built-in bias (Huff, 1954), focusing only on a part of the group and neglecting other subsets that might exhibit different behavior from that reported.

In mathematics, we do not take chances, and we do not trust a group of facts. In mathematics, we like to prove things to be true in well defined environments with well known properties. When something is considered true in mathematics there is no chance that another scientist anywhere in the world, or space, could prove us wrong. That is why proofs in mathematics are essential, and non-proved facts are either omitted in mathematics curriculum or left as open problems to study further.

Even though it sounds like a very simple process - you either prove a mathematical statement or you do not - mathematical proof might be a very complicated and time-expensive process depending on the complexity of the problem. Sometimes it takes years to prove a conjecture and sometimes even centuries. A well known problem that preoccupied many mathematicians for centuries is "Fermat's Last Theorem," or perhaps we should say Fermat's famous conjecture. In 1673, Pierre de Fermat stated the following conjecture:

> Equation
> $$x^n + y^n = z^n$$
> has no non-zero solutions for x, y, z and n integers, $n > 2$.

The conjecture was finally proved in 1995 by Andrew Wiles (NCTM, 1999). Originally, he announced his proof in 1993, but a serious flaw was discovered by one of the reviewers. It took Wiles almost two years to analyze and reexamine his work so he could finalize his proof.

Having an example like Fermat's Last Theorem in mind, it is clear that learning how to construct proofs in mathematics is not a simple task, but it is more of a lifetime adventure that starts early in the elementary school curriculum. Students learn to explain, justify, validate, and finally prove their results throughout their education, and most of them never master or even appreciate the efforts of proving.

In this thesis, our central goal is to gain an insight into early college students' understanding of mathematical proof. We do not look only for the answer as to whether they are able or unable to construct proof but also for the obstacles preventing them from succeeding.

Finally, to answer "What is mathematical proof?" we can simply say: "Mathematical proof is a valid argument." But then we need to define what a valid argument is. There are many definitions of mathematical proof, and each of them is characteristic of a certain level of mathematical maturity.

A poetic description of what mathematical proof is can be found in Schoenfeld (2009):

> If problem solving is the "heart of mathematics" then proof is its soul. (p. xii)

2.2 Valid argument across the grades in K-12

In mathematics, proof has several functions. As suggested by de Villiers in Harel and Sowder (2007), there are six roles of mathematical proof, and they are not mutually independent. Mathematical proof (Harel & Sowder, 2007, p. 819) serves us as:

- verification

- explanation
- discovery
- systematization
- intellectual challenge
- communication

"The notion of proof is not absolute..." (Hersh, 2009, p. 17). The understanding and the function of proof changes with mathematical developmental stages until it reaches the point of formal and rigorous proof as understood and accepted by researchers and mathematical society. Within the above framework, in elementary school, proving can be understood as "sense making," and it relies on informal mathematical reasoning and argumentation. As long as the argument is valid, non-contradictory, related to the subject and it yields a right conclusion teachers accept such argument as a mathematical proof at that level. Moreover, justifications by specific cases are very common and even desirable at this level. Later in this thesis, we refer to such validation as arguing from the example or proof by example. Encouraging students to explain their ideas and conclusions nurture three out of six roles described by de Villiers: verification, explanation and communication. Through exchanging their ideas and explanations, students and teachers form a purposeful mathematical communication. Even informal proof gives an explanation of the problem itself, and in a way it offers the verification of the conclusions. In the elementary grades, students see mathematics as something useful and practical and are unable to implement abstract thinking. Hence, proof by example is acceptable even though, later in our work, we characterize proof by example as the main mistake and misunderstanding that students have about mathematical proof.

Similar practice is common through the middle school, while at the same time students are introduced to symbols and formal symbolic mathematical notation. The process of proving remains of an empirical nature. An interesting study of how students in seventh grade construct proof is described by Boaler and Humpreys (2005). The teacher in the study is guided by the description of the proving process given by Mason, Burton, and Stacey (Mason, Burton, & Stacey, 1982, p. 103) who note three phases of proving:

> Convince yourself.
>
> Convince a friend.
>
> Convince an enemy.

The teacher adopts the process of proving to a reasonable skeptic, finding the argument to be valid only when it suffices to convince a skeptic. The teacher recalls a problem from the previous session, asking students to validate the conjecture they came up with. The conjecture to be proved is:

$$2(n-1) = 2n - 2.$$

Students are divided into groups, and by that organization they have been given an opportunity to follow the process above: to convince themselves, convince a friend (group members), and finally as a group to convince a skeptic (the teacher). All students started in the same way, validating the conjecture on specific examples. Only after the teacher stated the question: "How many numbers do you have to try out to be convinced?" did some students start to think in a more general way, and the reply she got was: "All numbers!". From this example we can see that students in the middle school can comprehend the necessity to validate general conjectures on all numbers, though very often they do not know how to accomplish such a complex goal. Very few students tried to use symbolic notation in order to represent any number, and with guidance and a lot of help from the teacher's side the class constructed the proof.

Looking forward into the case study described in more detail and discussed in Chapter 4, we can provide an example of K-8 reasoning when given the following problem.

Example 2.2.1. Prove if n is odd, then $n^2 + 1$ is even.

The problem itself is fairly simple, and most students in sixth grade would be able to understand and tackle it in some way. Most of them would try it out first using examples, but some might look further and try to characterize odd numbers using symbolic notation.

Using the National Council of Teachers of Mathematics *Curriculum Standards for School Mathematics* (1989), we can follow the expectations and standards in reasoning in grades K-4, 5-8, and 9-12. Thus in the lowest level, students should (NCTM 1989, p. 29):

- draw logical conclusions about mathematics
- use models, known facts, properties, and relationships to explain their thinking
- justify their answers and solution processes
- use patterns and relationships to analyze mathematical situations
- believe that mathematics makes sense.

Furthermore, in grades 5 through 8, students should (NCTM 1989, p. 81):

- recognize and apply deductive and inductive reasoning

- understand and apply reasoning processes, with special attention to spatial reasoning and reasoning with proportions and graphs
- make and evaluate mathematical conjectures and arguments
- validate their own thinking
- appreciate the pervasive use and power of reasoning as a part of mathematics.

Finally, the high-school mathematics curriculum should include various and numerous examples that will help students extend logical reasoning so that by the end of the 12th grade they should be able to (NCTM 1989, p. 143):

- make and test conjectures
- formulate counterexamples
- follow logical arguments
- judge the validity of arguments
- construct simple valid arguments.

Also, advanced, college oriented students should be given an opportunity to learn about indirect proofs and proofs by mathematical induction.

As we can see from the standards above transition to high-school understanding of mathematical proof consists mainly of using symbolic notation when validating general conjectures, and thus, building the road to formal proof. According to NCTM (NCTM, 2000a) students in secondary school should be able to

> ...justify and prove mathematically based ideas..

To summarize, in K-12 mathematics education by *Principles and standards for School Mathematics* (NCTM, 2000a) students should develop reasoning skills so that they can:

- Recognize reasoning and proof as fundamental aspects of mathematics
- Make and investigate mathematical conjectures
- Develop and evaluate mathematical arguments and proofs
- Select and use various types of reasoning and methods of proof.

Proof in K-12 can be found mostly as a part of the problem-solving section of curriculum. As such we can show various problem-solving examples that can be traced in K-12 depending on how the problem is stated at each level as well as on the questions asked.

Example 2.2.2 (Triangle and square).

grades K-2 Is the white triangle smaller than the black square in Figure 2.1?

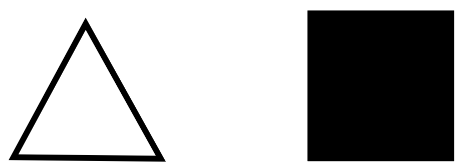


Figure 2.1: Triangle and square comparison for grades K-2

Children at the youngest age need to manage the models in order to compare two shapes. They are still learning and accepting the meaning of small-large, and they have difficulties to assign values to the shapes, as for example, length of the side of the triangle is three inches. Thus, we should not expect from them more than the simplest comparison between two objects described in the problem.

Their response and reasoning should be based on overlapping the objects as shown in the following Figure 2.2:

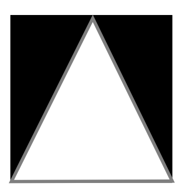


Figure 2.2: Answer to the triangle and square comparison for grades K-2

grades 3-4 Is the the triangle with equal sides of length three inches smaller or larger than the square with the sides of length of three inches?

Advancing to third grade students start managing numbers and they learn to recognize numbers representing objects and their characteristics.

Also, they are becoming familiar with the properties and names of shapes. For example, they learn that a shape with four equal sides and four equal angles is called a square. Also, they learn the meaning of area. Still, they are unable to use abstract thinking and problems they encounter should be very specific accompanied by figures or models as true as possible.

grades 5-8 Is the area of the equilateral triangle smaller or larger than the area of the square with the same sides?

At this level students should be able to recognize the features of the shapes named in the problem and construct figures representing them on their own. Also, they should have knowledge of the formulas representing areas of each geometrical shape in the problem and be able to compare those formulas algebraically in order to answer the question. The most common line of reasoning would be to start with specific numbers and conjecture the relation in general. It is possible for students to reason from general formulas but intuitive justification using visualizing methods or software should be accepted as valid arguments.

grades 9-12 Show that the area of an equilateral triangle with sides n is smaller than the area of a square with sides n.

Again, it is common to start with specific examples but the generalizing should be immediate and students should be able to justify their reasoning using area formulas only.

The problem described and analyzed in example 2.2.2 can be observed within the framework of van Hiele's model of reasoning as described by Burger and Culpepper (1993). According to van Hiele's model there are five levels of reasoning in geometry

(1) visualization

(2) analysis

(3) abstraction (informal deduction)

(4) deduction

(5) rigor

By the van Hiele theory students at the two lowest van Hiele levels of reasoning are unable to construct any formal type of mathematical proof. Even the students at the third van Hiele level are not expected to manipulate with rigorous proving processes but might be able to do only short proofs based on the empirically derived premises. Finally, students at the van Hiele levels four and five are expected to be able to provide consistently formal proofs. Such hypotheses are partially supported by the research reported by Senk (1989). As reported, students enter mathematical education at the first, ground level. The second and third levels are characteristic for the high school students, but only those advanced to the third level might be successful in writing mathematical proofs. This level is called a transitional level between formal and informal geometry. Currently most high school students are at or below the second van Hiele level, indicating that most of them are unable to understand, appreciate or construct a formal mathematical proof. On the other hand, if looking at the NTCM standards (NCTM, 1989) it is expected that high school students be at least at the third van Hiele level upon graduation.

2.3 Summary

In this chapter we have given a definition of what mathematical proof is and how it is understand at different ages. Also, we have learned about six roles of proof and have seen how some of them are conveyed to students during the exemplar teaching session described by Boaler and Humphreys (2005). Finally, we have observed one problem through different grades and how the same problem can be restated to challenge students at each stage. We have also discussed the van Hiele levels of reasoning and in short assigned K-12 grades to the first three levels. Before entering college students are expected to be at the third level but as supported by research (Senk, 1989) we have seen that most students are at the second or lower level by the end of their high school education. On the other hand, the fourth and fifth levels of reasoning are required to construct and understand mathematical proof at the college level.

CHAPTER 3

MATHEMATICAL PROOF IN THE COLLEGE CURRICULUM

The teacher's role in a student's learning the process of mathematical proof, its importance and its functions, is crucial. According to the *Principles an Standards for School Mathematics* (NCTM, 2000a):

> Effective mathematics teaching requires understanding what students know and need to learn and then challenging and supporting them to learn it well.

Also, to be able to teach effectively a teacher needs to understand what students know, as well as what students do not know, and what they do not understand. Having this in mind it is logical to search for the answers of what students know and do not know about mathematical proof among the students themselves.

In Section 3.2, we see what is expected from college students to know coming from high school as well as what they should learn about mathematical proof and what new proving skills they should develop during early college. To define what is the old knowledge, inherited from earlier education, and what are the newly acquired abilities, we look for the errors students make when constructing mathematical proof. In the next Chapter 4, we concentrate on 33 college students and analyze their work during four weeks, eight sessions (seven teaching sessions). Our main goal is to identify and classify mistakes they make using language and categorization adopted from the textbook, *Discrete Mathematics with Applications* (Epp, 2004). Such classification allows us to discover and define at which level a student gets "stuck," for example if a student persists on *proof by example* we say that s/he did not evolve from middle school comprehension of mathematical proof. Being stuck might be a frustrating situation for both students and teacher; however it should serve as a starting point for learning and teaching.

3.1 The role of proof in the college curriculum

We need to distinguish between undergraduate students that are math majors and others, usually science, majors. Other programs do not involve a lot of mathematics and do not require rigorous knowledge and understanding of proof. However, logical reasoning characteristic to mathematics, and validation in arguments, should be implemented in every day life, not only mathematics education. Because of the differences between math majors and other majors our focus is restricted to the early college curriculum, e.g., the first two years, when most students who have math requirements share similar mathematical courses.

The difference between K-12 and college mathematics is in the complexity of the problems students are confronted with. The problems they need to solve take more time to resolve, include different approaches, require different methods and very often there is more than one way to reach a solution. In most cases students are asked to prove their answer and to justify their methods, steps and algorithms. Proofs by example, informal arguments and similar methods are no longer accepted. Students are asked to use axioms and definitions to prove simpler statements, and furthermore, to dissect more complex problems into parts that could be proved using only primary sources such as axioms and definitions. In advanced mathematical courses, students are required to prove theorems, lemmas, and corollaries from statements and theorems previously proved.

In the first year, students take lower division mathematical courses; for example calculus (I, II, III, with precalculus) and discrete mathematics (requirements in BS in mathematics, BS in computer science, BS in computer engineering) [1] . In calculus students see many proofs, especially of the facts they were using through high-school or earlier and sometimes they are asked to provide proofs; but in discrete mathematics students actually learn more about what proof is, different methods of proof, and how to construct a proof.

3.2 Methods of proof in the college curriculum

In this section, we report on the methods usually taught in a discrete mathematics course. The differences of what and how the course is taught to students between different colleges, or different teachers, are inevitable and sometimes the same teacher decides to take a different approach and emphasize some methods over others. However, in general, all teachers at most colleges discuss some of the nine following techniques:

- Trivial Proof
- Vacuous Proof
- Direct Proof
- Indirect Proof

[1] All requirements listed are from degrees and programs offered at a large state university in northern California, year 2010.

    - ∗ Contraposition
    - ∗ Contradiction

- Proof by Cases
- Counterexample
- Proof by Exhaustion
- Existence Proof
    - ∗ Constructive Proof
    - ∗ Non-constructive Proof
- Mathematical Induction

In the later sections, we limit our investigation only to the seven methods considered in the textbook (Epp, 2004):

(1) Direct proof

(2) Indirect proof by contraposition

(3) Indirect proof by contradiction

(4) Proof by cases

(5) Proof by counterexample

(6) Proof by exhaustion

(7) Existence proof

Table 3.1: Truth table for the implication relation $P \to Q$

P	Q	$P \to Q$
T	T	T
F	T	T
T	F	F
F	F	T

The definitions and examples of the nine methods listed above are following:

Trivial Proof and *Vacuous Proof* are the simplest proving techniques and both are based on the implication truth table.

Suppose we want to prove a theorem of the form $P \to Q$ where P is a hypothesis and Q a conclusion. Then we have the following definitions.

Definition 3.2.1 (Trivial Proof)**.** When the conclusion Q is already known to be true it follows from the truth table 3.1 that the implication statement $P \to Q$ is always true. In this case we need to show that Q is true.

Example 3.2.2. Prove that if $x \in \mathbb{R}, x \geq 0$ then $x^2 \geq 0$.

But from calculus we already know that for all real numbers x, $x^2 \geq 0$ so the implication $x \geq 0 \Rightarrow x^2 \geq 0$ is trivially true.

Definition 3.2.3 (Vacuous Proof)**.** The implication $P \to Q$ is always true if the hypothesis P is false. Thus we need to show that P is false.

Example 3.2.4. Prove: If $x \in \mathbb{R}$ such that $x^2 + 2 = 0$ then $x > 0$.

Since the hypothesis that a real number has a negative square is false there is nothing that can be concluded from it. Hence there is nothing to be proved.

Both Trivial and Vacuous Proof are often omitted in the college curriculum and it is hard to find their definitions or examples in discrete mathematics textbooks.

Definition 3.2.5 (Direct Proof)**.** We assume that the hypothesis P is true. Using rules of inference and theorems already proved we can show that Q must be true as well.

This is the most common proving method but its name might be misleading at first. The directness of such proof comes from the final written form but its progress usually is much different and often starts in reverse order. When trying to prove Q from P mathematicians might work their way backward to see how Q follows from P, or in other cases the proof is the result of interchanging forward and backward steps.

Example 3.2.6. For $n \in \mathbb{N}$ such that n even, show that n^2 is even.

n even $\Rightarrow \exists k \in \mathbb{N}$ so that $n = 2k$

It follows that $n^2 = n \times n = (2k) \times (2k) = 2(2k^2) = 2m$ for $m = 2k^2 \in \mathbb{N}$. Thus by the definition of even numbers n^2 is even.

Definition 3.2.7 (Indirect Proof: Contraposition)**.** Because implications $P \to Q$ and $\neg Q \to \neg P$ are logically equivalent it follows that $P \to Q$ is valid when $\neg Q \to \neg P$ and vice versa.

Example 3.2.8. Prove: for $n \in \mathbb{N}$, n^2 odd $\to n$ odd.

The contrapositive of the statement n^2 odd $\to n$ odd is n not odd $\to n^2$ not odd. In other words: n even $\to n^2$ even.

We have already proved that implication in Example 3.2.6 thus by Proof by Contraposition we can conclude that $n \in \mathbb{N}$, n^2 odd $\to n$ odd is true.

Definition 3.2.9 (Indirect Proof: Contradiction)**.** We assume that the hypothesis P is true while assuming at the same time that the conclusion Q is false. The proof is completed when we arrive at contradiction such as $\neg P$, Q or $R \wedge \neg R$ starting from $P \wedge \neg Q$.

Example 3.2.10. A common problem provable by contradiction is: Prove that $\sqrt{2}$ is irrational.

Our hypothesis is that 2 is rational number, and the conclusion is $\sqrt{2}$ is irrational.

Suppose 2 is rational and $\sqrt{2}$ is rational. By the definition of rational numbers $\sqrt{2}$ is rational if $\exists m \in \mathbb{Z}$ and $n \in \mathbb{N}$, $\gcd(m,n) = 1$ such that $\sqrt{2} = \frac{m}{n}$. Now:

$$\sqrt{2} = \frac{m}{n} \Leftrightarrow 2 = \frac{m^2}{n^2} \Leftrightarrow 2n^2 = m^2$$

$$\Rightarrow m^2 \text{ is even} \Rightarrow m \text{ is even}$$

Next, m even $\Rightarrow \exists k \in \mathbb{N}$ s.t. $m = 2k \Rightarrow m^2 = 4k^2$.

Finally,

$$m^2 = 2n^2 \Leftrightarrow 4k^2 = 2n^2 \Leftrightarrow 2k^2 = n^2$$

$$\Rightarrow n \text{ is even}$$

And we have reached the contradiction with $\gcd(m,n) = 1$. In this case we have arrived at the contradiction of the form $R \wedge \neg R$ ($\gcd(m,n) = 1$ and both m and n are even i.e. $\gcd(m,n) = 2k, k \in \mathbb{N}$).

Definition 3.2.11 (Proof by Cases)**.** In this case the proof is conducted by breaking down the original implication into two or more cases and proving each case separately. $P \to Q$ becomes $P_1 \to Q \wedge P_2 \to Q, \ldots, P_k \to Q$ where $P_1 \cup P_2 \cup \ldots \cup P_k = P$.

Example 3.2.12. For every $n \in \mathbb{N}$ $n^2 + 1$ is not divisible by 4.

We can break this problem into two cases by investigating n even and n odd separately, but very soon we discovered that this can not lead to the conclusion.

Thus, the implication should be broken into four cases:

(1) $n \equiv 0 \mod 4 \Rightarrow n^2 \equiv 0 \mod 4 \Rightarrow n^2 + 1 \equiv 1 \mod 4$

(2) $n \equiv 1 \mod 4 \Rightarrow n^2 \equiv 1 \mod 4 \Rightarrow n^2 + 1 \equiv 2 \mod 4$

(3) $n \equiv 2 \mod 4 \Rightarrow n^2 \equiv 0 \mod 4 \Rightarrow n^2 + 1 \equiv 1 \mod 4$

(4) $n \equiv 3 \mod 4 \Rightarrow n^2 \equiv 1 \mod 4 \Rightarrow n^2 + 1 \equiv 2 \mod 4$

Since every natural number n falls in one of the four cases above we have shown that for every n, 4 does not divide $n^2 + 1$.

Definition 3.2.13 (Proof by Counterexample)**.** This method is used to disprove statements of the form $\forall x$, $P(x)$ is true (or to prove that $\forall x\ P(x)$ is false). The proof is completed when we can provide an element a such that $P(a)$ is false.

Example 3.2.14. Disprove: Every $p \in \mathbb{N}$, p prime $\Rightarrow p$ is odd.

2 is a natural number that is prime but even $\Rightarrow$ implication above is false.

Definition 3.2.15 (Proof by Exhaustion)**.** We use this method to show that $\forall x$ $P(x)$ is true by showing that $P(x)$ is true for each x independently. This is possible only when x takes only finitely many different values.

Example 3.2.16. For $n \in \mathbb{N}$, $1 < n \leq 3$, $2^n - 1$ is prime.

Since n can only be 2 or 3 it can be easily checked if the the proposition is valid.

$n = 2 \Rightarrow 2^2 - 1 = 4 - 1 = 3$ prime

$n = 3 \Rightarrow 2^3 - 1 = 8 - 1 = 7$ prime

The proposition is proved.

Existence proofs are methods to prove statements such as: $\exists x$ such that $P(x)$. There are two ways to prove this type of statement.

Definition 3.2.17 (Existence proof: Constructive proof)**.** In constructive proof the strategy is to find or construct an element a such that $P(a)$ is true.

Example 3.2.18. There is a natural number n such that 2^n is a prime number.

2^n is an even number for every $n \geq 1$, thus we are looking for an even, prime number. The only such number is 2 so we need to find n such that $2^n = 2 \Rightarrow n = 1$.

Definition 3.2.19 (Existence proof: Non-constructive proof)**.** As opposed to constructive proof we use non-constructive proof when we are unable to find or construct an element a such that $P(a)$ is valid. In this case we assume that there is no such element a and we arrive at a contradiction. Thus we show that there must be some a such that $P(a)$ is true.

Example 3.2.20. There are irrational numbers a, b such that a^b is rational.

This is a well-known example of non-constructive proof.

Let us consider the number $m = \sqrt{2}^{\sqrt{2}}$. Now m is either rational or irrational. We already know that $\sqrt{2}$ is irrational, thus if m is rational we have shown the existence. On the other hand if m is irrational then for $a = m$ and $b = \sqrt{2}$ we have $a^b = 2$. Hence, either way there are such a and b.

Definition 3.2.21 (Mathematical Induction)**.** We want to prove that the statement $P(n)$ holds for all natural numbers $n \geq m$ for some $m \in \mathbb{N}$. There are two steps to the proof:

- *Basis* Show that the statement holds for some $n = m$.
- *Inductive Step* Assuming that the statement holds for $n - 1$ we need to show that it is valid for n as well.

Example 3.2.22. Show that for every $n \in \mathbb{N}$, $3|(n^3 + 3n^2 + 2n)$.

- *Basis* For $n = 1$ we have $n^3 + 3n^2 + 2n = 6$ and $3|6$.

- *Inductive Step* Assume that for an arbitrary n, $3|n$ then $n = 3k$.

  Now for $n + 1$ we have:

$$\begin{aligned}(n+1)^3 + 3(n+1)^2 + 2(n+1) &= n^3 + 6n^2 + 11n + 6 \\ &= (n^3 + 3n^2 + 2n) + 3(n^2 + 3n + 2) \\ \text{by assumption} &= 3k + 3(n^2 + 3n + 2) \\ &= 3(k + n^2 + 3n + 2) \\ &\Rightarrow 3|[(n+1)^3 + 3(n+1)^2 + 2(n+1)] \\ &\Rightarrow 3|(n^3 + 3n^2 + 2n),\ \forall n \in \mathbb{N}\end{aligned}$$

We have defined and provided an example for each of the methods but in future sections, we limit our discussion and report only to the seven proving methods that are covered and investigated in Epp's textbook (Epp, 2004). Thus, methods that are discussed in the following chapters are:

- Direct proof
- Indirect proof: contradiction and contraposition
- Proof by exhaustion
- Existence proof: constructive proof
- Proof by cases
- Proof by counterexample

3.3 How to write a proof?

Methods of proof as well as the importance of mathematical proof are introduced to students through numerous examples. So they are expected to mimic simple proofs that are done in the classroom at first and then start writing proofs by themselves using methods and results presented in the classroom. Since the role of mathematical proof is to convince oneself and others of the truthfulness of one's proposition, the proof should be readable and understandable to the wider audience, not to the author only. For that purpose it should be clear where the proof starts, where it ends, what we know to be true and what we need to show. Also, it is very important to provide a reasonable and clear justification of each step in the proof. In most of the textbooks on discrete mathematics it is common to find the "recipe" with steps of how to write a proof in order to produce a structurally readable mathematical proof. For example, in (Epp, 2004) the following steps are discussed:

(1) Copy the statement of the theorem to be proved on your paper

It should be clear to the reader what the assignment is.

(2) Clearly mark the beginning of your proof with the word Proof

Just to have a neat start, this is important in long and complex proofs.

(3) Make your proof self-contained

Clearly state the definitions and axioms used in the proving process, as well as the supporting claims that might be proved elsewhere; clearly state where and when, or prove them here.

(4) Write your proof in complete sentences

For readability purposes it is desirable to have explanations and transition between ideas stated in sentences rather than symbolic notation only.

(5) Give a reason for each assertion you make in your proof

Explain and justify each declaration used in the proof. Furthermore, provide a reason why such declaration should be valuable for the proof.

(6) Include the "little words" that make the logic of your arguments clear, e.g. if, then, now, such as, follows, therefore, for, let us assume, this means, by assumption, by definition. (Epp, 2004, p. 134)

We also believe that the following should be added

(7) Mark the end of the proof using one of the common end notations, such as: Q.E.D. or □

Having all these steps does not guarantee a complete and valid proof but following the prescribed structure can help one to start and focus on what needs to be proved. Also, following the proof scheme allow readers, teachers and peers to comprehend one's reasoning and to identify flaws if any.

3.4 Common mistakes

In her textbook, Epp also lists the most common mistakes students make (Epp, 2004, p. 135). The following Table 3.2 is a summary of common mistakes listed in the textbook. The table consists of two categories: Mistakes and Grade level. The grade level category represents the educational stage at which students learn how to overcome the mistake. In an earlier stage such a mistake might be tolerable, such as arguing from examples being acceptable in grades K-5.

Table 3.2: Common mistakes in constructing mathematical proof

Mistake	Grade level
Arguing from example	K-6
Same identifier	6-8
Jumping to a conclusion	8-12
Begging the question	8-12
Misuse of the word “if”	11-12

3.5 Summary

In this chapter we have laid the base for our case study by providing a full list of proving methods taught at colleges accompanied with the definition of each method and an illustration of their employment on the examples. Furthermore, we have listed common mistakes students make according to the textbook used during the case study (Epp, 2003). Also, based on the literature research we have assigned an educational level to each mistake category, as in Table 3.2.

In the following chapters, we analyze case study results by discussing mistakes categorized by Table 5.1. Also, we compare the occurrence of each mistake between different proving methods. Furthermore, brought up are conclusions about students' comprehension of proof based on mistakes they make and their occurrence.

CHAPTER 4

CASE STUDY ON THE TEACHING AND LEARNING OF PROOF

The case study, as the central part of this thesis, was conducted in one of the two sections in the Discrete Mathematics course, at a large state university in northern California, during Spring term of 2010. Subjects of the study were students enrolled in the course during the term. In total, 33 students participated in the study, and they were all presented with the same research instruments described in Section 4.2 below.

The study was completed in four weeks, and consisted of the following phases:

(1) Survey and consent forms

(2) Pre-teaching questionnaire

(3) Observation and two quizzes

(4) Post-teaching questionnaire

(5) Midterm exam

By its nature, this study was a systematic research design as characterized in Wiliam (1998, p. 7). The researcher enters the classroom and investigates the subjects' performance in their authentic environment using described instruments. The main goal of the research is to learn and understand the way in which mathematical proof is taught, as well as to learn about the difficulties students have while learning to construct a valid argument.

4.1 Aim

The aim of this case study is for the researcher to gain an insight into the process of learning/teaching mathematical proof at the early college level. From personal experience the researcher is aware of long, exhausting and very often unsuccessful attempts to learn how to construct a proof. At the same time, the researcher lacks the knowledge of how other students overcome common mistakes in proving and how they develop the sense for a valid argument.

The ultimate goal is to categorize common mistakes students make, describe difficulties students encounter and to identify gaps in mathematical knowledge inherited from earlier education. Finally, the following questions are pursued:

- *What type of proof do students accept as the most practical?*
- *What proof method do students find most complicated to use?*
- *What common mistakes do students make*?
- *Which mistakes exhibit a tendency to increase/decrease during and after the teaching sessions*?
- *Which difficulties do students encounter when attempting to construct a valid mathematical proof in the early college curriculum*?

4.2 Methods

The techniques of data collection employed in the present study are the most common techniques in mathematics education research as described by Zevenbergen (1998). These are: participant observation and text documents. It is important to say that the researcher enters the classroom in the nonparticipatory role in order to observe the daily school life. The role of this technique is to learn from the participants without imposing the researcher's opinion. As the part of the context the researcher is expected to "understand the research setting, its participants, and their behavior" (Glesne & Peshkin, 1992, p. 42) . The teaching sessions were slow paced and such a setting allowed the researcher to take notes by hand. Also, only the teacher wrote on the blackboard while students' participation consisted only of oral suggestions and comments. Furthermore, examples and problems used during the teaching sessions were taken from the textbook by S. Epp (2003) which made taking notes easier: by having the problems already written down the researcher focused easily on the students' participation.

The second data collection technique used in the reported case study are the following text documents:

- Survey
- Pre-teaching questionnaire
- Class quizzes and Midterm test
- Post-teaching questionnaire

Survey The questions students answer in the survey are not only for identifying and matching purposes but also to get an insight into the trends between different groups characterized by gender, college major, English proficiency and math courses taken with or before Math 042. The survey in its full extent is enclosed in the Appendix A.

Pre- and Post- teaching questionnaires are similar in their context and serve us to see if students' understanding of mathematical proof advances during instruction or the instruction itself has no significant impact on their deductive reasoning. The survey and questionnaires were constructed by the researcher.

Class quizzes and tests are part of the observation phase and their main purpose is to track the teaching/learning process over the six instruction sessions. Also, since quiz questions are of the same nature as the questionnaire questions it is possible and reasonable to compare the results on the quizzes to those on the questionnaires. Both questionnaires can be found in later sections 4.4.1 and 4.4.2. The data collected from the questionnaires are presented in two ways, qualitative and quantitative. The first quiz was constructed by the class instructor, while the second one was designed by the researcher.

Quantitatively we report the number of students who have the ability to solve the problem and correctly explain their answer by providing a mathematical or English proof. By English proof we mean a logical explanation in English that does not necessarily use mathematical symbols. The reason to accept the English proof is not to discredit students who understand what needs to be done in order to validate their answer even if they lack the ability and/or knowledge to express themselves using formal mathematical language.

Additional quantitative data derive from the questionnaires, providing categories of the mistakes students make when explaining their answers. The mistake in the explanation does not mean that the student got the wrong answer to the question, but the mistake shows the student's inability to prove s/he is correct. For example, if a student's answer to the question is correct but the provided proof is a proof by example, this indicates the level of understanding has not evolved from the lover level of mathematical maturity linked to the middle school level. For example, a pre-teaching questionnaire problem is:

For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$,
$(2n+1)(2^n+1)$ is ODD.

The statement is:

(1) True

(2) False

Answer:

Please explain your answer.

Qualitative aspects of the collected data are in the description of mistake categories, information gleaned about classroom atmosphere during observations and description of students' work on different tests.

Finally, to draw conclusions of how and when students start developing their understanding of mathematical proof at the higher level we implement both techniques and combine the results into a single report.

4.3 Lessons

The instruction observation is a substantial part of the case study. Therefore, we provide a summary of the teaching from the researcher's point of view, emphasizing students' responses to the new methods and topics, instead of the teacher's performance.

In total there were six sessions dedicated to teaching and learning new proving methods. Two out of six sessions were testing sessions; in other words, students had only four teaching sessions.

In the first session the teacher, hereafter refereed to as Dr. G, introduces proving techniques using homework examples, making a smooth transition from the previous topic.

Example 4.3.1.

$$D = \{-48, -14, -8, 0, 1, 3, 16, 23, 26, 32, 36\}$$

$\forall x \in D$ if x is odd then $x > 0$.

True or false?

When the teacher discusses the homework problem he mentions that this is true and it could be proved using proof by exhaustion.

Based on the silence and students' indifference to Dr. G's monologue the researcher concluded that students were not interested in the word "proof" or why the teacher mentions that it could be proved using a method with a certain name. Even after explaining what the method means and how it could be employed students had no comments and let the teacher proceed.

Following the example, Dr. G introduced his way of teaching in a philosophic manner saying:

> "Teaching a proof if well prepared is not teaching but reading what someone else has done before."

and he continued with examples from the textbook always letting students provide an answer before showing the proof himself. Another technique the teacher employed was student oriented in that he let students guide him in the proving process even when knowing that it is wrong. He wanted students to realize what and where the proof went in the wrong direction. The following example 4.3.2 illustrates such a process. Note that as the observations were not taped, the dialogue is reproduced from careful notes taken during class and represents the spirit of what occurred in class.

Example 4.3.2. *Teacher*:

Let's prove the following theorem.

Theorem: *The product of two odd numbers is odd.*

Teacher: The easiest way to begin proving to yourself that something is true is by looking at an example. So for example we have $3 \cdot 5 = 15$, where both 3 and 5 are odd, and we can see that their product is odd. Now, we'll try to prove this claim for any two odd integers.

Proof: Start with the definition of what you have:

Definition: n being an odd integer means that there is an integer k such that $n = 2k+1$ or in symbolic notations we can write this definition as: n is odd $\Leftrightarrow \exists k \in \mathbb{Z}$. s.t. $n = 2k + 1$.

Teacher: Next you need to translate what odd is into math language.

We need to prove that for two odd integers, n_1 and n_2 odd, their product is odd, i.e. $n_1 \cdot n_2$ is odd.

Do you have any ideas how to continue the proof?

Student 1: We can write $n_1 = 2k + 1$ and $n_2 = 2k + 1$ and then we have $n_1 \cdot n_2 = (2k + 1) \cdot (2k + 1)$

Teacher: Ok, so if we continue we get:

$$n_1 \cdot n_2 = (2k+1) \cdot (2k+1) = 4k^2 + 2k + 1 = 2(2k^2 + k) + 1$$

Is this an odd integer?

Group of students: Yes.

Teacher: Do you believe that we have proved our statement?

Group of students: Yes

(though they hesitated when answering.)

Teacher: How many of you believe that we have provided a correct proof?

(Ten out of 33 students raised their hands.)

Teacher: And how many believe that we are wrong?

(16 out of 33 students raised their hands. Seven students could not decide. None of the students who were against the proof could explain why it was wrong. So the teacher asked students to substitute k for some integers and to analyze the numerical examples. Only after taking $k = 1, 3, 5$ some of the students realized that what they proved is if n is an odd integer then n^2 is odd.)

Teacher: Well, not bad but you should be aware that k could not be the same for both n_1 and n_2, so we should put:

$n_1 = \text{odd} \rightarrow n_1 = 2k_1 + 1$

$n_2 = \text{odd} \rightarrow n_2 = 2k_2 + 1.$ for k_1, k_2 integers.

Any ideas how to proceed?

(This particular problem, later characterized as the *same identifier* mistake, was not discussed any further. It is important to mention that students were reintroduced to symbolic notation and its usage in previous sessions at the beginning of the semester. So the teacher felt that they should have been familiar with identifiers. Furthermore, as we have said before, Dr. G prefers to teach students using examples. Therefore, a few more examples with similar problems were introduced on the blackboard and each time Dr. G only mentioned the necessity to use appropriate symbolic notation.)

Student 2: Multiply n_1 and n_2,

$$n_1 \cdot n_2 = (2k_1 + 1) \cdot (2k_2 + 1) = ...$$

Teacher: Is there a problem?

Student 3: ...

$$n_1 \cdot n_2 = (2k_1 + 1) \cdot (2k_2 + 1) = 4k_1k_2 + 2k_1 + 2k_2 + 1$$

but how do we know that this is an odd integer?

Teacher: Can you relate this expression to the definition of an odd integer?

Student 3: Oh, $4k_1k_2 + 2k_1 + 2k_2 + 1 = 2(2k_1k_2 + k_1 + k_2) + 1$, it is similar but...

Teacher: Can anyone translate our expression into the rigid definition from the beginning?

Silence... so the teacher continues:

We can rewrite our expression in a way: $2(2k_1k_2 + k_1 + k_2) + 1 = 2k + 1$, for $k = 2k_1k_2 + k_1 + k_2$ integer.

To finish the proof we simply write the final statement that shows what needed to be proved, i.e. $n_1 \cdot n_2 = (2k_1 + 1) \cdot (2k_2 + 1) = 4k_1k_2 + 2k_1 + 2k_2 + 1 = 2k + 1$ for some $k = 2k_1k_2 + k_1 + k_2$ integer thus by the definition, product of two odd integers is an odd integer. □

Just for your information we use symbols as: □, "q.e.d." or simply "end" to note the end of the proof.

The type of proof we just used to prove the theorem is called a direct proof.

Example 4.3.2 is a good example of how students misuse symbolic notation, i.e. it is an example of the *same identifier* mistake as described in Table 5.1.

During all four teaching sessions Dr. G never introduced a proving method at first, but after proving an example problem he mentioned what the method is called and then, if possible, provided the general steps in the process.

Even though the first few examples may seem very simple and almost trivial, students had a hard time tackling the proving process for each of the problems. Example 4.3.3 shows four problems similar in difficulty to the first one, and reports on students' reactions and/or related questions.

Example 4.3.3.

- There is no smallest integer.

  When asked if they believe that statement is true all students answered "YES" in one voice, but when asked why they believe so no one had an answer.

- For all integers n, $n \leq n^2$. As in the previous problem students were very sure of the correctness of the statement and when asked to justify three out of 33 students provided numerical examples but when asked to offer a proof or at least to start the proof there were no responses. Thus, Dr. G provided proving steps using the "Proof by contradiction" method.

- The sum of an odd and an even number is odd.

  As before, after stating the problem the teacher let students to lead the proof. In this problem one student offered the beginning such that:

  Student 1: $n_1 = 2k + 1$ and $n_2 = 2k$, thus $n_1 + n_2 = 2k + 1 + 2k = 4k + 1$, so it is an odd number.

Teacher: Do you all agree?

Student 2: No, we can not use same k for both n_1 and n_2.

Teacher: Correct.

As we can see in the example 4.3.3 some students continue to make the *same identifier* mistake, but at the same time another student realizes the mistake and makes the correction without Dr. G's intervention. In their attempt to offer the proof to the stated problems students made several more mistakes besides *same identifier*. Often students do not connect lines of the proof with the equality sign to indicate equivalence or during the discussion they forget what needs to be proved.

Examples of similar difficulty were introduced to students in the following sessions, and often we observed similar scenarios. At first students let the teacher show them the first example and afterwards they try to mimic his methods and steps in order to provide proofs. An unfortunate observation during the teaching sessions is that only five students were active and joined the teacher in providing proofs. They either suggested proving steps or corrected their peers when they believed they were wrong. But the other 28 students quietly copied examples from the blackboard or answered in choir when asked fairly simple questions. Another observation that caught our attention is that only male students were active in discussions but the classroom structure was in a ratio that can not be statistically significant, in other words there were only five female students in the group of 33 students.

Students also showed insecurity in mathematical topics and definitions and even when they knew how to start the proof they were unable to define certain expressions. Thus through the four teaching sessions students asked the following questions and exhibited intimidation with the new topics:

(1) Prove: If n is any odd integer, then $-1^n = -1$.

After showing examples and breaking out the expression using the definition of odd numbers and the exponential laws we heard the question:

Student: What are exponential laws?

(2) Prove that the statement is false.

There exist $k \in \mathbb{Z}$ s.t. $k \geq 4$ and $2k^2 - 5k + 2$ is prime.

After factoring, which was done by the teacher since the students forgot how to factor binomials [no one even remembered to search for roots of the quadratic equation] one student asked:

Student: I have a question? When defining, a prime number is not 1 and not composite? So 1 is not prime.

Teacher: Yes.

Student: What is a composite?

(Surprisingly, none of his classmates provided the definition of a composite number. Some students gave numerical examples but no one was confident enough to formally explain what a composite is.)

(3) Prove 4 does not divide $n^2 + 1$.

Teacher: We can do this in two ways: observing "even and odd cases" or in a way we just learned using mod notation, mod 4. Which do you prefer?

Students: Even and odd cases.

(The answer students gave shows that students rather use methods they are more familiar with and procedures they are more confident with than learning and employing new methods.)

(4) Prove, $3 \not| n^2 + 1\ \forall n$.

This problem is within the same topic as the previous problem and when asked how to start the proof, students suggested to observe for n being even and odd. Very soon they discovered that "even or odd" cases are not sensitive enough to cover all the possibilities so they have to employ newly acquired knowledge about divisibility. As soon as the teacher suggested using $\mod$ notation, the students withdrew and let the teacher proceed on his own.

We observed similar behavior during the next session as well. Students were introduced to a new topic, floor and ceiling functions. After going over definitions and basic properties the teacher started with the simplest proving tasks.

(5) *Teacher*: Example: If n is even, then

$$\lfloor \frac{n}{2} \rfloor = \frac{n}{2}$$

are you ready for this?

Let us look at the example: $n = 4$, we have

$$\lfloor \frac{n}{2} \rfloor = \lfloor \frac{4}{2} \rfloor = 2 = \frac{4}{2}.$$

This is so trivial. I'll show you two proofs.

What is the obvious way to start this?

Students: $n = 2k$ .....

(Now they knew how to start if they have even or odd numbers in the conditions, but the problem was how to continue. The newly introduced floor function definition became an obstacle. None of the 33 students felt confident enough to suggest the next step, so the teacher proceeded on his own.)

4.4 Research instruments

In this section, described are all research instruments used in the case study. Moreover, we provide the required or expected answers to sample questions and present a selected student's work to illustrate how their work has been evaluated and mistakes categorized. Very often students provide right answers but still make significant mistakes in justification.

4.4.1 Pre-teaching questionnaire

The pre-teaching questionnaire or Questionnaire 1 consists of three questions. All three are related to mathematical proof and understanding of the same. Students are presented with fairly simple number theory problems and asked to answer whether the statement is true or false or to choose the correct answer in a multiple choice question. Either way, we are using two different approaches to the questions, one is to evaluate if students answer correctly, and the other is to see their reasoning behind the answer. Also, when no explanation is given we are assuming that the student is guessing without understanding the background of the problem. In the pre-teaching questionnaire we are trying to determine whether students consider it important to explain their answers and how close those explanations are to the actual proofs of the problems.

Through their explanations we can see how students are working on solving the problem. Using the list of actions needed to solve the problem as explained in Polya (1971, p. xvi), we can decode students' process of solving the given problem.

For example:

(1) For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$,

$(2n + 1)(2^n + 1)$ is ODD.

The statement is:

(a) True

(b) False

Answer: ________________

Please explain your answer.

A student familiar with the process of proof and its purpose usually follows the steps described by Poyla in *How to Solve It* (Polya, 1971):

- Understanding the problem

  What is unknown in the problem? *The parity of the expression.* What are the conditions? *n is a natural number, $n > 0$.* Does it seem true? *Try out for a few natural numbers, $n = 1$ we get 9, for $n = 4$ we get $9 \cdot 17 = 153$.*

- Devising a plan

  Find the connection between the unknown and the conditions by answering the following questions: Have you seen the same or similar problem before? *In the class we talked about the definition of odd and* *even numbers, also we have talked about the properties of products and sums of odds or evens.* Can you use a previous problem to solve this one? Can you use results of the previous problem solved? Can you use the methods of solving the previous problem? *Here, I can use both.*

- Carrying out the plan

  How they carry out the plan of the solution shows in the explanation students provide for their answer. Most students choose to give an English explanation as below:

  $2n + 1$ is always ODD

  $2^n + 1$ is always ODD

  product of two ODD numbers is always ODD

  In this example it is clear that the student is going through the two previous phases but what is missing here are validations of the claims stated while carrying out the plan.

- Looking back

  The final and crucial phase in validating the answer is to go back to your answer, and to check if every step is validated and that anybody who is reading the solution should be convinced that this is the right answer.

4.4.2 Post-teaching questionnaire

In the post-teaching questionnaire the problems were more direct and formal proofs were required. Three out of four questions were tightly related to the topics covered during seven teaching sessions. The fourth problem required students to read and understand the well known Pythagorean Theorem and to state its converse. Furthermore, they were asked to use the converse to justify their answer about the given triplet that happened to be a Pythagorean triplet. The objective of that question is to determine whether students are able to transfer their new knowledge about mathematical proof to a topic that was not covered during lecture.

An exemplary problem in the post-teaching questionnaire was:

> Prove that for all integers a and b if $a|b$ then $a^n|b^n$ for all $n \in \mathbb{N}$.
> *Proof*:

The desired proof type was a direct proof. If the students choose to follow the same Polya recipe as described above, the complete justification would be as follows:

- Understanding the problem

  What is unknown in the problem? *The divisibility of one integer by another having certain properties.* What are the conditions? *n, a and b are integers such that $a|b$.* Does the claim seem true? *Try out for a few numbers: $n = 2$, $a = 3$, $b = 12$ we get $a^2 = 9$, $b^2 = 144$ and* $144 : 9 = 16 \Rightarrow 9|144$; or *$n = 3$, $a = 2$, $b = 6$ we get $a^3 = 8$, $b^3 = 216$ and* $216 : 8 = 27 \Rightarrow 8|216$ *It might be true.*

- Devising a plan

  Find the connection between the unknown and the conditions by answering the following questions: Have you seen the same or similar problem before? *In the class we talked about the definition and conditions for divisibility.* Can you use a previous problem to solve this one? Can you use results of the previous problem solved? Can you use the methods of solving the previous problem? *Absolutely. I should start with the definition and carry on from there.*

- Carrying out the plan

  How they carry out the plan of the solution shows in the explanation students provide for their answer. Most students tried to provide a direct proof from the definitions:

$$a|b \Rightarrow \exists k \in \mathbb{Z} \text{ so that } b = a \times k$$
$$\text{Thus } b^n = (a \times k)^n = a^n \times k^n \Rightarrow a^n|b^n \text{ by the definition}$$
$$\text{for there is an integer } z = k^n \text{ such that } b^n = a^n \times z.$$

4.4.3 Quizzes during the teaching sessions

During the seven lecture sessions students had an opportunity to see their progress by solving two short quizzes. Each quiz had only one problem.

Quiz 1 Prove that if n is an odd integer then n^2 is odd.

Since the quiz problem was similar to the teaching examples that were proved using direct proof, the students were expected to provide a direct proof as well, using the definition of an odd integer. The proof itself follows easily from the definition. Using the previously described Polya recipe we can present the proof in short form:

- Understanding the problem

  Unknown: *The parity of* n^2. Conditions: n *is odd.* Does the claim seem true? *Examples:* $n = 3$ *and* $n^2 = 9$ *or* $n = 13$ *and* $n^2 = 169$, *hmm both are valid. Might be true.*

- Devising a plan

The connection between unknown and conditions: Similar problem? *During the previous session students were introduced to the following problem: The product of two odd integers is always odd.* Can you use the previous problem to solve this one? Can you use results of the previous problem solved? Can you use the methods of solving the previous problem? *Yes to all. Students can mimic the proof by adjusting both odd integers to be the same integer, or they can argue from the proved conjecture about the two odd integers.*

- Carrying out the plan

  Most students tried to provide a direct proof from the definitions by mimicking proof of the conjecture about the product of two odd integers.

The second quiz was presented to students after five sessions. In addition to the different methods of proof, they were also introduced to a new topics in number theory such as floor and ceiling functions. Accordingly, their second quiz was on that topic and they were asked to provide a direct proof.

Quiz 2 For n an odd integer prove that

$$\lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}.$$

The best solution would follow the sequence of steps as described by Polya. Thus we would expect to see the following work:

- Understanding the problem

  Unknown: *The value of floor function for certain integers.* Conditions: *n is odd.* Does the claim seem true? *Examples: $n = 3$ and $\lfloor \frac{3}{2} \rfloor = 1$ while $\frac{3-1}{2} = 1$. Or $n = 23$ and $\lfloor \frac{23}{2} \rfloor = 11$ while $\frac{23-1}{2} = \frac{22}{2} = 11$. It works on examples.*

- Devising a plan

  The connection between unknown and conditions: Similar problem? *During the previous session students were introduced to the the definition of floor function as well as to various similar problems. They have also seen a few solutions and proofs manipulating the floor definition.* Can you use a previous problem to solve this one? Can you use results of the previous problem solved? Can you use the methods of solving the previous problem? *Yes to all. Students should start from the definition and in a step or two they should employ their knowledge about the values of floor function for integers.*

- Carrying out the plan

  Most students tried to provide a direct proof from the definition.

4.4.4 Midterm exam

Chronologically the last data source in the case study is the midterm exam. After the last teaching session students had one review session where they had an additional opportunity to ask questions on the proof topics and to review problems they found to be challenging for them to solve in class or homework. Also, the teacher solved at least one problem for each type of question that would appear on the exam. The exam itself consisted of nine questions, and some questions had multiple subproblems. Five out of nine questions were proof related and we focus only on these problems. Proof related exam questions were a collection of problems similar to those in two quizzes and two questionnaires introduced to students during the teaching sessions. Thus, the midterm exam summarizes students' evolution in understanding and employing methods of mathematical proof.

The topics that appeared in the midterm exam are as follows:

(1) even/odd

(2) divisibility

(3) divisibility properties

(4) floor function

(5) proof of theorem

The full midterm exams, two versions, can be found in Appendix A.3.

4.5 Summary

The main part of this thesis is the case study described in Chapter 4. Here we have described continuity of the teaching lessons, problems and teaching examples introduced to students, and provided a brief description of the teacher's teaching philosophy. Furthermore, parts of teacher-students' dialogues are included in Section 4.3 to gain a better perspective of the teaching lessons.

In great detail we have described each of five data instruments: two quizzes, two questionnaires and a midterm test. A few individual problems are discussed together with desirable, expected solutions.

All data instruments in their original form, as given to students, can be found in Appendix A.

CHAPTER 5

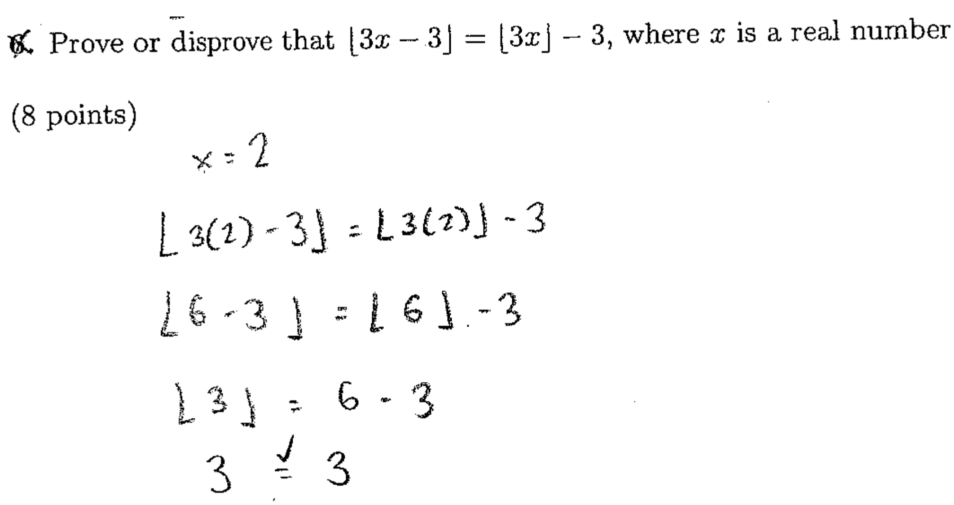
RESULTS

In this chapter, we report on the results collected from the research documents described in Section 4.4. First, we present the extended table of common mistakes students do. The full list of mistakes is the result of analyzing students' work during the case study. The full description of mistake categories can be found in Table 5.1. Mistakes and their meaning are used to analyze and discuss results and students' work and progress during the case study. Furthermore, the results are presented with two tables for each data instrument, one with the percentages of correct answers and the other with the list of all mistakes for each question. We also trace the occurrence of each mistake, or better to say mistake category.

5.1 Common mistakes

In this section, we provide an extended list of common mistakes students make. The author (Epp, 2004, p. 135) lists five mistakes, see page 27. After conducting the study we have categorized four more. Furthermore, each mistake is illustrated with an example extracted from the students' work during the teaching sessions. To identify and match different data sources students have chosen a pseudonyms, therefore we use the pseudonyms to match each example to the student.

(1) arguing from example(s), as illustrated in Figure 5.1. This is a problem from the midterm test, and the student who argued from an example used a pseudonym of his/her choice; "006620939".


6. Prove or disprove that $\lfloor 3x - 3 \rfloor = \lfloor 3x \rfloor - 3$, where x is a real number

(8 points)

$x = 2$

$\lfloor 3(2) - 3 \rfloor = \lfloor 3(2) \rfloor - 3$

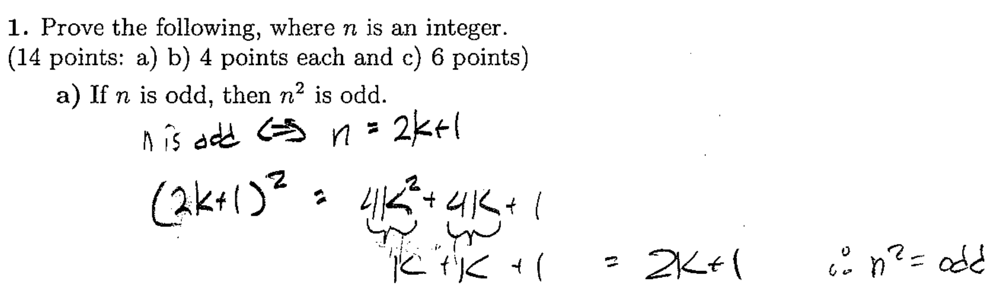
$\lfloor 6 - 3 \rfloor = \lfloor 6 \rfloor - 3$

$\lfloor 3 \rfloor = 6 - 3$

$3 \stackrel{\checkmark}{=} 3$

Figure 5.1: Example of *argue from example* mistake

(2) using the same identifier to mean two different things. Again, we have used a midterm problem from the student known to researcher under the code "3.1415".


1. Prove the following, where n is an integer.
(14 points: a) b) 4 points each and c) 6 points)

a) If n is odd, then n^2 is odd.

n is odd $\Leftrightarrow$ $n = 2k+1$

$(2k+1)^2 = 4k^2 + 4k + 1$

$k + k + 1 = 2k + 1$ $\therefore n^2 =$ odd

Figure 5.2: Example of *same identifier* mistake

(3) jumping to a conclusion; as shown in Figure 5.3. To illustrate this mistake category we have used a first problem in the pre-teaching questionnaire from a student with the code "penpen."

1. For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$, $(2n+1)(2^n+1)$ is ODD.
   The statement is:
   (a) True
   (b) False

Answer: true

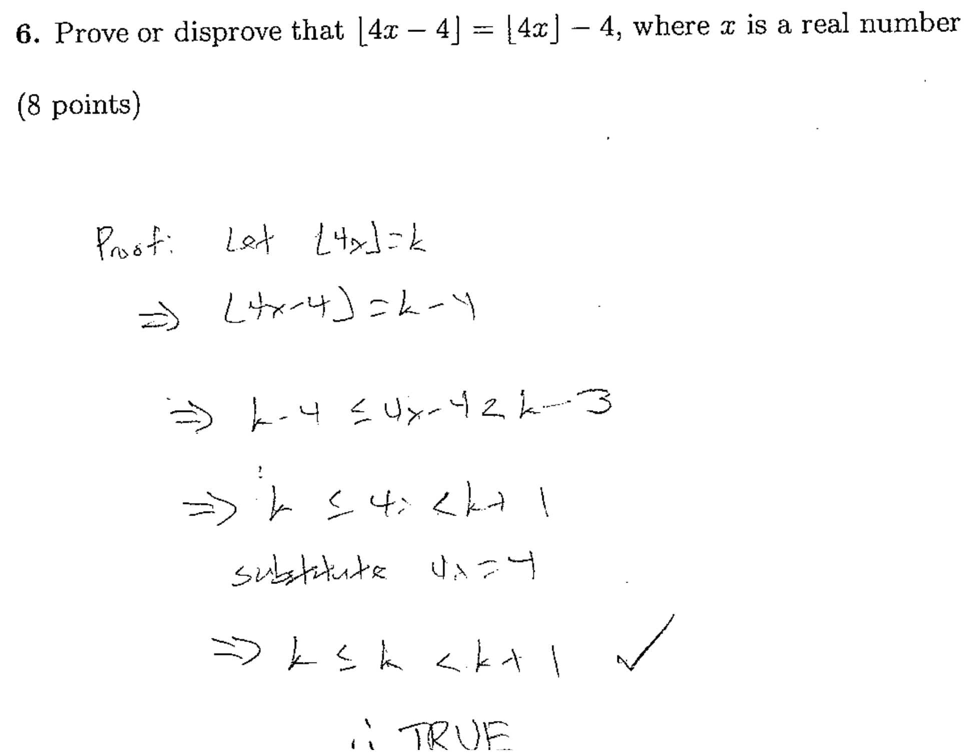
Please explain you answer.

because 2n+1 is an odd number
and because 2^n, no matter what number n would be 2^n would
be even, +1 will be odd.
odd x odd = odd.
therefore this statement is true

Figure 5.3: Example of *jumping to conclusion* mistake

(4) begging the question; assume what is to be proved, as in Figure 5.4. At the first glance *begging the question* and *jumping to conclusion* might look similar, but through the examples we can see their main difference. *Jumping to conclusion* is when a student starts an argument correctly from the definition or any other primary source, but then after a few steps comes to a conclusion without justifying all his steps. On the contrary, when making *begging the question* mistake, the student assumes what needs to be proved in the first step and then proceeds his argument from there. Such work is well illustrated by the student "jwild37" in his/her work on the sixth problem in the midterm test shown in

Figure 5.4.


6. Prove or disprove that $\lfloor 4x - 4 \rfloor = \lfloor 4x \rfloor - 4$, where x is a real number

(8 points)

Proof: Let $\lfloor 4x \rfloor = k$

$\Rightarrow \lfloor 4x - 4 \rfloor = k - 4$

$\Rightarrow k - 4 \leq 4x - 4 < k - 3$

$\Rightarrow k \leq 4x < k + 1$

substitute $4x = 4$

$\Rightarrow k \leq k < k + 1$ ✓

$\therefore$ TRUE

Figure 5.4: Example of *begging the question* mistake

(5) misuse of the word “if”

In the case study described in Chapter 4 we can see a few more errors arising frequently. Thus additional categories are:

(6) trying to solve for the unknown instead of proving the claim for the unknown

The student treats the statement as the problem to be solved for the unknown instead of trying to prove the general claim. An example can be seen in Figure 5.5, work extracted from the midterm test by a student known under the pseudonym “lenlen.”

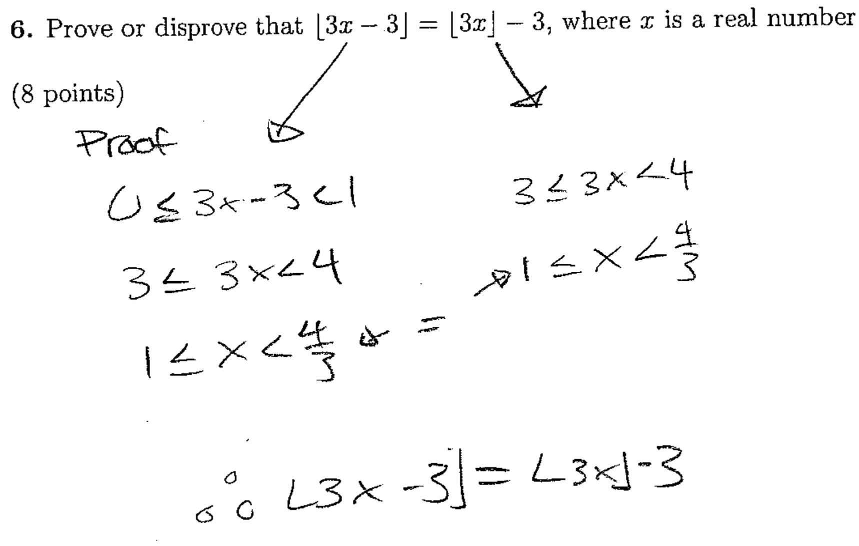


Figure 5.5: Example of *solving for the unknown* mistake

(7) intuitive or English proof

The student understands what needs to be proved and is able to build up a logical reasoning process but lacks the ability to express the process in mathematical/symbolic language, for example see Figure 5.6 presenting the third midterm problem solved by a student named “Jillian.”

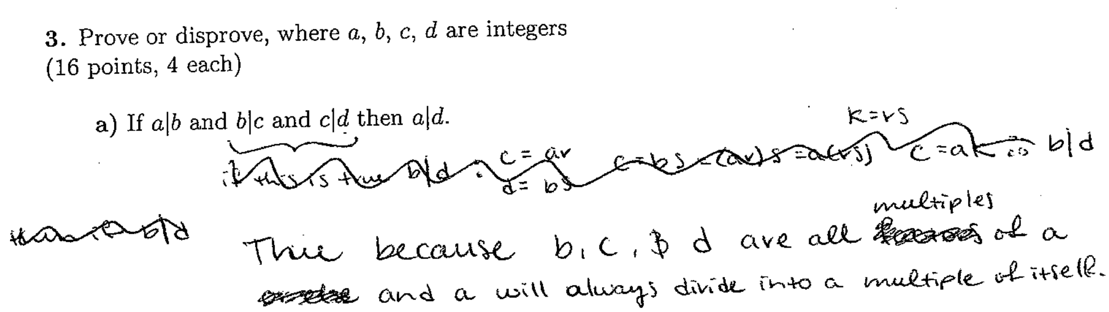


Figure 5.6: Example of *intuitive proof* mistake

(8) wrong conclusion or no conclusion

In this category we look for a complete lack of argument, an invalid argument and/or a wrong conclusion following a correct argument. No conclusion can be due to the following reasons: misunderstanding of what needs to be proved; lack of prerequisite knowledge of definitions and axioms needed to prove the claim; or, inability to use definitions, axioms, previously proved theorems in the proving process. An invalid argument usually follows incorrect definitions or illogical reasoning. A wrong conclusion might be the consequence of a computational mistake, misinterpretation of the previously proved results and/or undeveloped logical reasoning. One of the examples can be seen in Figure 5.7. Here we are presenting the problem from the second quiz by a student behind the code "m8G04."

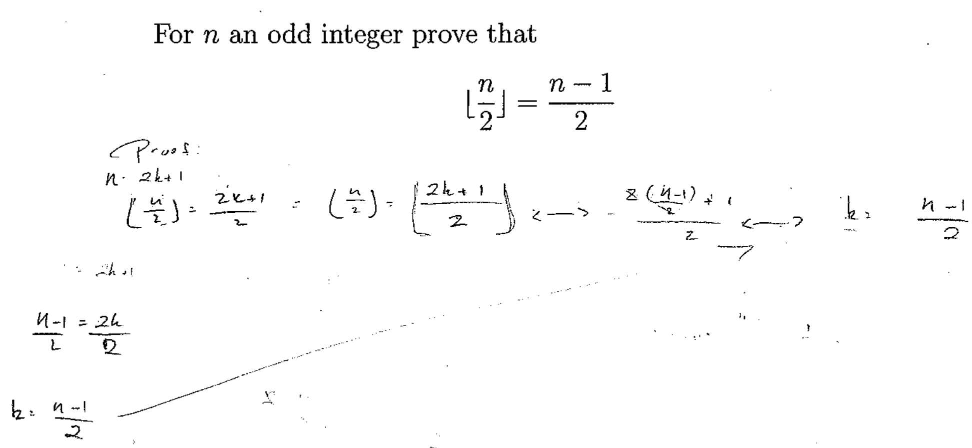


Figure 5.7: Example of *wrong (no) conclusion* mistake

(9) computational mistakes

Most computational mistakes are caused by carelessness and have no significant meaning in our study.

The following Table 5.1 is a summary of common mistakes listed above. The table consists of three categories: Mistakes, Grade level and Difficulties. As before, the grade level category represents the educational stage at which students learn how to overcome the mistake. The third category, difficulties, lists rationales behind the mistake. The table, especially difficulties represented by each mistake category, are the result of the class observation and case study results. After observing students' work during the teaching sessions and analyzing the collected data the researcher categorized all the mistakes and constructed the table. Furthermore, it is possible for a student to make more than one mistake simultaneously. In this case, we record only the one with the higher priority. For example, if a student argues from an example and makes a computational error in the same problem we disregard the computational mistake and account only for the other. In fact, we use the hierarchy from Figure 5.8 when prioritizing mistakes. Also, it should be noticed that there is no misuse of the word "if" in the figure. We have omitted that mistake since all the students in the study avoided using little words, thus every student made that particular error in each problem. The hierarchy depicted in Figure 5.8 is based on Table 5.1, more specifically on the grade level assigned to the mistake category. The highest priority has a mistake that is common among lower grades, and that should be abandoned as student advances to the higher level of education. The only exception to this reasoning is the intuitive proof that is ranked lower on the scale.

Table 5.1: Extended list of common mistakes in constructing mathematical proof

Mistake	Level	Difficulties
Arguing from example	K-6	Inability to generalize Uncomfortable with symbolic notations Difficulty to employ abstract reasoning
Same identifier	6-8	Uncomfortable with symbolic notations Uncertain of general definitions Doesn't understand the relation between symbols and numbers they replace
Jumping to conclusion	8-12	Inability to think abstractly Doesn't understand process of justification Unsure of what needs to be proved Disregarding some cases
Begging the question	8-12	Unsure of what needs to be proved Inability to see the difference of what is given and what asked
Misuse of the word "if"	11-12	Doesn't understand process of justification
Intuitive proof	K-5	Insufficient knowledge about the topic Inability to think abstractly Doesn't understand process of justification Inability to manipulate with the symbolic notation
Solving for unknown	5-10	Doesn't understand process of justification Inability to generalize Unsure of what needs to be proved
Wrong conclusion or no conclusion (no justification)	5-12	Unsure of what needs to be proved Doesn't understand the claim Doesn't understand the conditions Inability to make a connection between the conditions and claim Proving irrelevant claim
Computational mistakes	5-12	Mindless mistakes Disregarding conditions

5.2 Pre-teaching questionnaire

A total of 33 pre-teaching questionnaires were analyzed. It is possible to record a correct answer to the true/false or multiple choice question, while at the

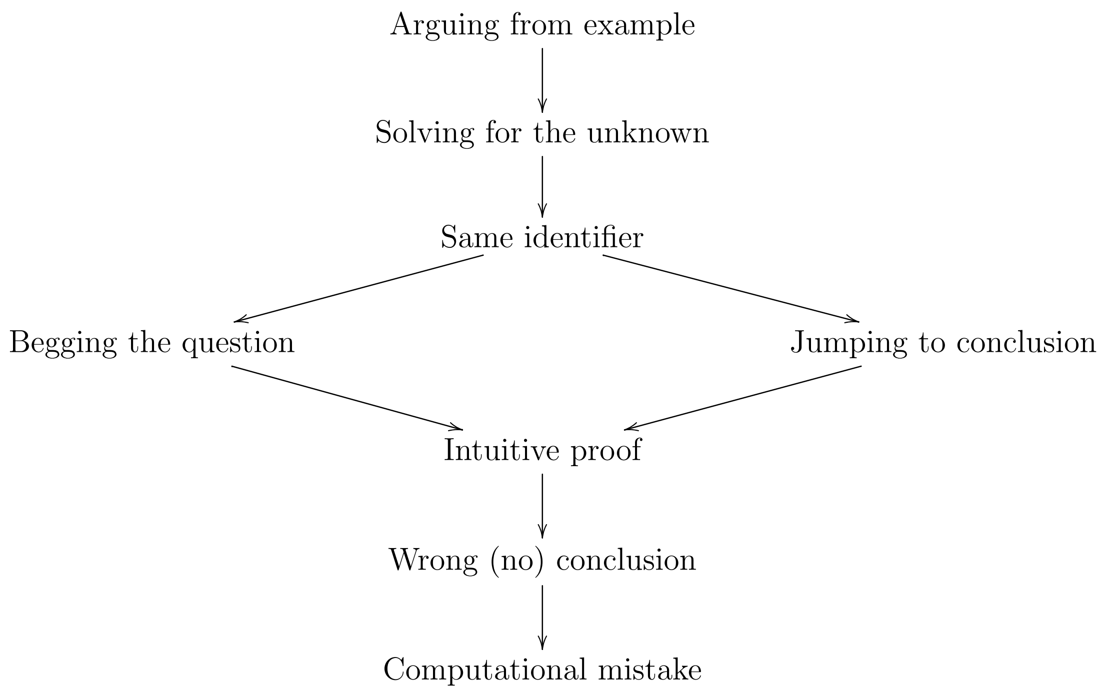


Figure 5.8: Mistakes priority

same time the justification to the answer might be inaccurate. Thus, we can notice different numbers in the tables, i.e. as we can see in the pre-teaching questionnaire tables, Table 5.2 and Table 5.3, there were 26 correct answers out of 33, while we have counted 28 proving mistakes. The number of correct answers per question can be found in Table 5.2.

Furthermore, questions on the first questionnaire were:

(1) For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$,
$(2n+1)(2^n+1)$ is ODD. (True or False?)

(2) For $n = 0$ we have $(2 \cdot 0 + 1)(2^0 + 1) = 1 \cdot 2 = 2$ and 2 is even.
How does this fact relate to the previous problem?

(a) It is irrelevant to the previous problem.

(b) It is a counterexample we can use to prove that the statement in 1. is not true.

(c) It is a special case of the previous problem.

(3) For all positive real numbers a and b, the following is true:

$$\sqrt{a+b} < \sqrt{a} + \sqrt{b}.$$

Which of the following statements can be deduced:

(a) There exist $a, b > 0$ such that $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

(b) There are no $a, b > 0$ such that $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

(c) For all $a, b \in \mathbb{R}$ $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

(d) None of the above.

Table 5.2: Pre-teaching questionnaire results

	Question 1	Question 2	Question 3	Total
Correct in #	26	25	20	
%	78.8	75.6	60.6	71.7

Mistakes for Questionnaire 1 are listed in Table 5.3. The mistakes were classified using Table 5.1 on the page 56. Furthermore, in case a student made more than one mistake we have recorded only the one with the highest priority as illustrated in Figure 5.8. Due to the structure of the questionnaires we have one of the three possibilities:

- Correct answer and the correct justification

- Correct answer and incorrect proof
- Incorrect answer and incorrect (no) justification.

These explain the discrepancies between result tables and results per mistake tables, i.e. number of mistakes plus correct answers does not add up to the total number of questionnaires analyzed.

Table 5.3: Pre-teaching questionnaire results per mistake

Question	Type of proof	Mistake	Occurrence	
			in #	in %
Question 1	Direct proof	Arguing from example	7 in 28	25
		Jumping to conclusion	10 in 28	36
		Wrong (no) conclusion	8 in 28	29
		Solving for unknown	1 in 28	3.6
		Begging the question	2 in 28	7.1
Question 2	Vacuous proof	Intuitive proof	24 in 30	80
		Wrong (no) conclusion	6 in 30	20
Question 3	Trivial proof	Begging the question	1 in 20	5
		Intuitive proof	1 in 20	5
		Arguing from example	5 in 20	25
		Wrong (no) conclusion	13 in 20	65

On the other hand, if a student made a small computational errors, but other than that his proof would be correct we have counted that as a correct answer but recorded an error as well. This is another reason why the numbers in two tables, results and results per mistake, do not add up to total number of responses.

5.3 Post-teaching questionnaire

There were four problems in the second questionnaire.

(1) Prove that there exists an integer n such that $2n^2 - 21n + 40$ is prime.

(2) Prove that for all integers a and b if $a|b$ then $a^n|b^n$ for all $n \in \mathbb{N}$.

(3) Is this true or false? **Answer**: ________________

For all integers n, $6n+1$ is not divisible by 3.

Justify your answer.

(4) **Pythagorean theorem** In a right triangle with c representing the length of the hypotenuse, and a and b representing the lengths of the other two sides it holds that: $a^2 + b^2 = c^2$.

State the converse of the Pythagorean Theorem: (*we know that the converse of Pythagorean Theorem is also true*)

Can $a = 13$, $b = 84$, $c = 85$ be lengths of the sides of a right triangle?

Justify your answer.

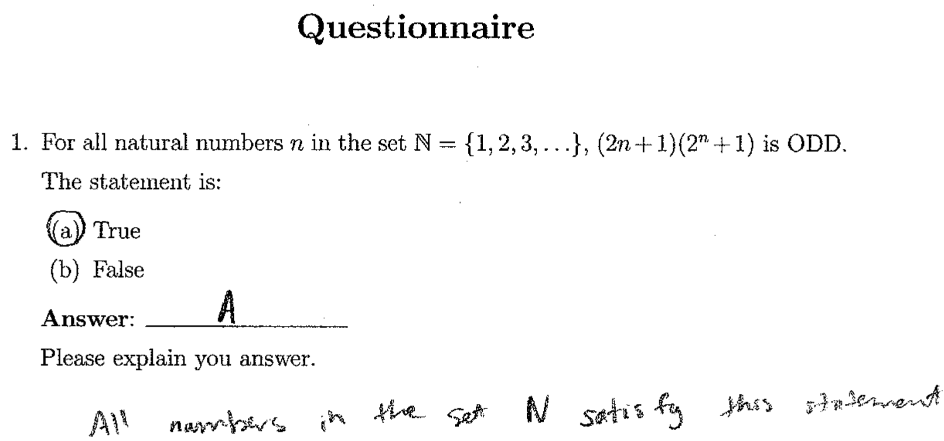
The post-teaching questionnaire was completed by 29 students and the results are shown in Table 5.4.

Table 5.4: Post-teaching questionnaire results

	Question 1	Question 2	Question 3	Question 4	Total
Correct in #	10	10	26	21	
%	34.4	34.4	89.6	72.4	57.8

The mistakes students made on the second questionnaire and their occurrence are recorded in Table 5.5. Also, we have noticed that students provided correct answers but still made proving mistakes, for example in the third question there were only 13 incorrect answers but 17 mistakes.

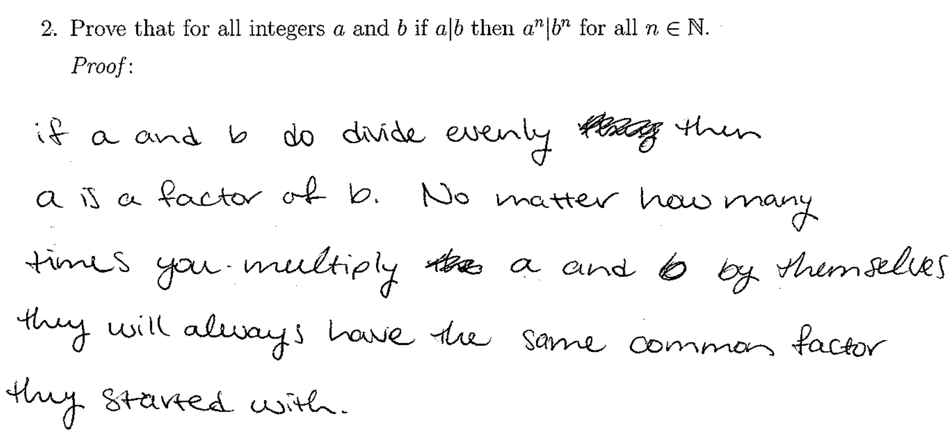
In the two figures 5.9 and 5.10 are examples of *begging the question* and *intuitive proof* mistakes, the first recorded in the pre-teaching, and the latter in the post-teaching questionnaire.



Questionnaire

1. For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$, $(2n+1)(2^n+1)$ is ODD.
   The statement is:
   (a) True
   (b) False
   Answer: A
   Please explain you answer.
   All numbers in the set N satisfy this statement

Figure 5.9: Example of *begging the question* mistake in pre-teaching questionnaire, by student " "



2. Prove that for all integers a and b if $a|b$ then $a^n|b^n$ for all $n \in \mathbb{N}$.
   Proof:
   if a and b do divide evenly then a is a factor of b. No matter how many times you multiply a and b by themselves they will always have the same common factor they started with.

Figure 5.10: Example of *intuitive proof* mistake in post-teaching questionnaire, by student "Jillian"

Table 5.5: Post-teaching questionnaire results per mistake

Question	Type of proof	Mistake	Occurrence	
			in #	in %
Question 1	Existential proof (constructive)	Computational mistakes	1 in 18	5.5
		Jumping to a conclusion	1 in 18	5.5
		Wrong (no) conclusion	16 in 18	89
Question 2	Direct proof	Intuitive proof	2 in 19	10.5
		Computational mistakes	2 in 19	10.5
		Arguing from example	2 in 19	10.5
		Wrong (no) conclusion	13 in 19	68.5
Question 3	Direct proof	Begging the question	5 in 17	29.4
		Intuitive proof	8 in 17	47
		Computational mistakes	1 in 17	5.9
		Arguing from example	2 in 17	11.8
		Wrong (no) conclusion	1 in 17	5.9
Question 4	Direct proof	Computational mistakes	2 in 13	15
		Wrong (no) conclusion	11 in 13	85

5.4 Quizzes

Immediately at the end of the first teaching session students were presented with the first quiz. The quiz was very short, it consisted of only one question and its purpose was to see if students were able to prove a problem almost identical to what they just saw on the blackboard. In total 31 students participated in the first quiz.

On the other hand, the second quiz was given to students at the beginning of the third session. It was also a one question test. The rationale for giving students a two day period before the second quiz was that they were introduced to a new topic, floor function, and they were told that the quiz was going to be about that topic. With the additional time we were hoping that students would understand and accept the new definition and therefore achieve better results on the quiz. There were 30 students engaged in the second quiz. Results for both quizzes are presented in Table 5.6.

The question in quiz 1 was:

(1) Prove: If n is an odd integer then $n^2 + 1$ is even.

and in the second quiz the question was:

(1) For n an odd integer prove that $\lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}$.

Table 5.6: Quiz 1 and quiz 2 results

	Quiz 1	Quiz 2
# of students	31	30
Correct in #	11	7
Correct in %	35.4	23.3

In the first quiz we have counted 24 mistakes categorized accordingly to Table 5.1 in six mistake types. While for the second quiz 23 out of 30 students were unable to provide a correct proof. In Tables 5.7 and 5.8 we have documented mistakes for both tests.

Table 5.7: Quiz 1 results per mistake

Type of the proof	Mistake	Occurrence in #	in %
Direct proof	Jumping to conclusion	8 in 24	33.3
	Argue from example	3 in 24	12.5
	Computational mistakes	1 in 24	4.3
	Solving for the unknown	5 in 24	21
	Wrong (no) conclusion	4 in 24	16.6
	Same identifier	3 in 24	12.5

Table 5.8: Quiz 2 results per mistake

Type of the proof	Mistake	Occurrence in #	in %
Direct proof	Jumping to conclusion	13 in 23	56.5
	Computational mistakes	1 in 23	4.4
	Wrong (no) conclusion	7 in 23	30.4
	Same identifier	2 in 23	8.7

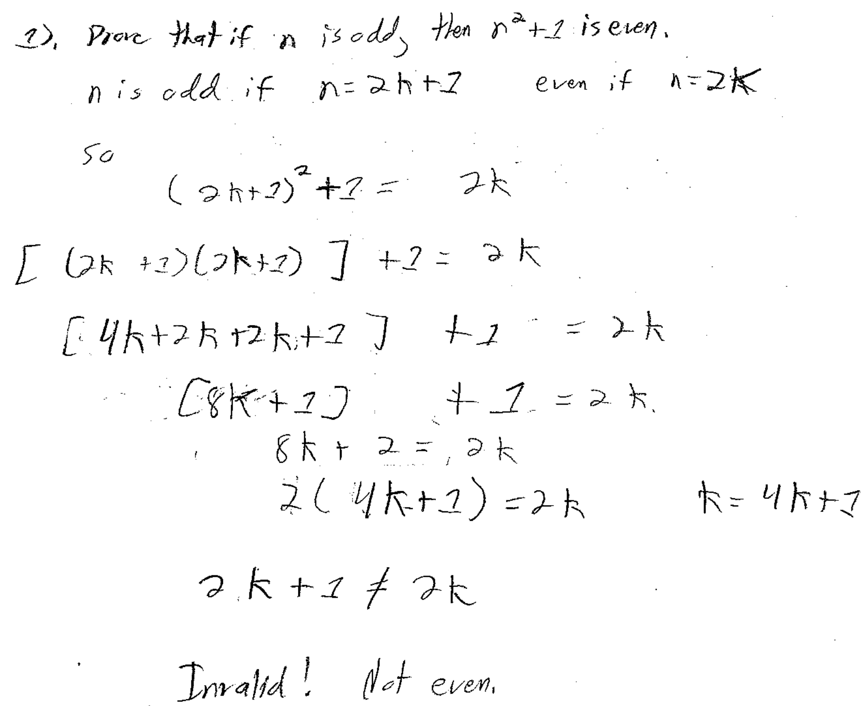


Figure 5.11: Example of *solving for the unknown* mistake in quiz 1, by student "mathmathmath"

In the figures 5.11 and 5.12 we can see the two mistakes made by the same student in the two quizzes. In Figure 5.11 "mathmathmath" student's mistake is categorized as *solve for the unknown* mistake. S/he states the problem in an equation form. But the last line in her/his proof is a contradiction line, meaning that the student believed s/he reached a contradiction and hence believes that

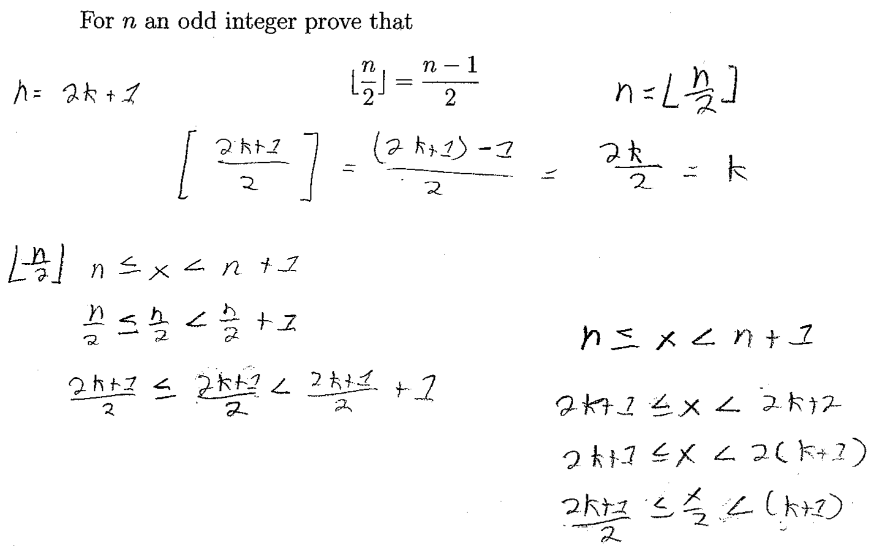


Figure 5.12: Example of *jumping to conclusion* mistake in quiz 2, by student "mathmathmath"

the statement to be proved is actually wrong. Also, student uses the same identifier for both general numbers ($2k + 1$ for an odd number and $2k$ for an even number). Thus, there are at least three mistakes involved: *solve for unknown*, *same identifier* and *wrong (no) conclusion.* Using the hierarchy of the mistakes as in Figure 5.8 the final classification of the mistake goes in the favor of *solve for unknown.*

Furthermore, in Figure 5.12 the same student makes another set of mistakes. The most obvious one is *jump to conclusion.* In the first line s/he assumes what needs to be proved. If s/he did not try to elaborate her/his conclusions we would classify the mistake to be *begging the question.* Unfortunately, her/his work following the first line has no value in validating the statement.

5.5 Midterm

The final data document in this case study is the midterm (see Appendix A pg. 114). The midterm, as described in Section 4.4.4, was an extensive test of students' performance and we have collected data from 31 midterm tests. Even though the midterm consisted of nine questions we are considering only five that are proof related. In these five questions there were subproblems and when these five problems were broken down to single questions we have ended with ten questions listed later in the section. To discourage cheating and to be sure of individual work Dr. G introduced two sets of questions. One set was labelled as "group E" and the other as "group O". Furthermore, the type of proof required (expected) was identical per question so we are combining results for both groups. The only exception are questions 3b) and 3c). These two questions are the same in both groups but they appear in inverted order, i.e. 3b) in E equals 3c) in O group. Accordingly we have combined answers so that the unique result can be obtained per question.

In Table 5.9 we have documented the results students achieved on the midterm exam. In total, data from 31 midterm exams were collected and results were reported as number and percentage of correct answers per question.

The underlying reasons for the huge discrepancies between certain questions are new or unrelated mathematical topics. We discuss each of these in Chapter 6. Furthermore, the midterm mistakes per each question together with their occurrence are documented in the tables: 5.10, 5.11 and 5.12. The tables are organized by the question topic. For example the first three questions are one question in the midterm with three subproblems.

To gain a better understanding of Table 5.10 we should notice that in the

Table 5.9: Midterm results

Question	# of correct	% of correct answers
1	29	94
2	23	74
3	14	45
4	17	55
5	25	81
6	16	52
7	9	29
8	5	16
9	17	55
10	10	32
Total		53

second question, or question 1b), students used two types of proof: direct proof and proof by cases. More precisely, nine out of 31 students made attempts to provide a direct proof while 22 students employed proof by cases.

In the following sections, the proof questions from the midterm are analyzed for mistakes. To make it easier to examine the mistakes, the 10 questions are analyzed in three groups, based on the problem subject. The questions appear first, then the tables showing mistakes.

(1) (E) Prove for n integer: If n is even, then $n^3 + 2$ is even.
(O) (If n is odd, then n^2 is odd.)

(2) (E) Prove for n integer: 2 does not divide $n^2 + (n+1)^2$.
(O) (2 divides $n^2 + (n+2)^2$.)

(3) (E) Prove for n integer: 4 divides $n^2 + (n+2)^2$ if and only if n is even.
(O) (4 divides $n^2 + (n+2)^2$ if and only if n is even.)

The fifth and sixth questions, 3b) and 3c) in the two versions of the original exam, are the same questions in both problem sets but in inverse order. Thus we

Table 5.10: Midterm results per mistake for questions 1 – 3

Question	Type of the proof	Mistake	Occurrence in #	Occurrence in %
1	Direct proof	Computational mistakes	2 in 2	100
2	Direct proof	Argue from example	1 in 3	33.3
		Wrong (no) conclusions	2 in 3	66.7
	Proof by cases	Same identifier	1 in 5	20
		Jumping to conclusion	4 in 5	80
3	Proof by cases	Argue from example	1 in 19	5.3
		Wrong (no) conclusions	1 in 19	5.3
		Computational mistakes	2 in 19	10.5
		Jumping to conclusion	15 in 19	78.9

are reporting combined results.

(4) (E) Prove or disprove, for a, b, c, d integers. If $a|b$ and $b|c$ and $c|d$ then $a|d$. (in both groups)

(5) (E) Prove or disprove, for a, b, c, d integers. If $2a|b$ then b is even.
(O) (If $a|2b$ then a is even.)

(6) (E) Prove or disprove, for a, b, c, d integers. If $a|2b$ then a is even.
(O) (If $2a|b$ then b is even.)

(7) (E) Prove or disprove, for a, b, c, d integers. If $a|b$ then $a^2|4b^4$.
(O) (If $a|b$ then $a^2|5b^3$).

Finally, the three last problems are individual problems on three different topics and their results are recorded in Table 5.12.

(8) (E) Prove or disprove that $\lfloor 4x - 4 \rfloor = \lfloor 4x \rfloor - 4$, where x is a real number.
(O) ($\lfloor 3x - 3 \rfloor = \lfloor 3x \rfloor - 3$)

Table 5.11: Midterm results per mistake for questions 4 – 7

Question	Type of the proof	Mistake	Occurrence in #	Occurrence in %
4	Direct proof	Jumping to conclusion	1 in 13	7.7
		Wrong conclusion	5 in 13	38.5
		Intuitive proof	3 in 13	23
		Argue from example	1 in 13	7.7
		Same identifier	3 in 13	23
5	Direct proof	Intuitive proof	2 in 6	33.3
		Wrong (no) conclusions	4 in 6	66.7
6	Proof by counterexample	Wrong (no) conclusions	15 in 16	93.7
		Computational mistakes	1 in 16	6.3
7	Direct proof	Wrong (no) conclusions	20 in 22	91
		Computational mistakes	1 in 22	4.5
		Argue from example	1 in 22	4.5

(9) (E) Prove that 3 divides $n^3 + 3n^2 + 5n$ for all integers n.

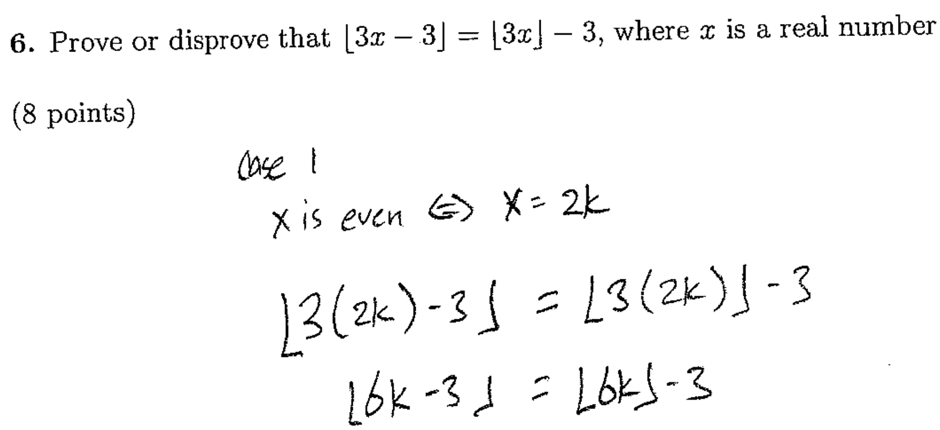
(O) (3 divides $n^3 + 3n^2 + 2n$)

(10) (E) Prove the Pythagorean Theorem. (in both groups)

Table 5.12: Midterm results per mistake for questions 8 – 10

Question	Type of the proof	Mistake	Occurrence in #	Occurrence in %
8	Direct proof	Wrong (no) conclusion	21 in 26	81
		Begging the question	1 in 26	3.7
		Argue from example	3 in 26	11.5
		Trying to solve for unknown	1 in 26	3.7
9	Proof by cases	Computational mistakes	2 in 13	15
		Wrong (no) conclusions	11 in 13	85
10	Direct proof	Argue from example	1 in 19	5.3
		Wrong conclusion	18 in 19	94.7

The floor function problem, problem eight in the report, has a high percentage of *wrong (no) conclusion* mistake type, but it is important to add that the large portion of the students who made that mistake actually were unable to manipulate with the definition of the floor function. More preciously, seven out of 21 students did not understand how to use the definition even though they stated the floor function definition clearly. Three more students were unable to state the definition itself. We can see in the figures 5.13 and 5.14 two examples of students' inability to state and use the definition. In Figure 5.13 student known as "3.1415" was unable to state the definition correctly. Thus, any further attempts yielded wrong conclusions. On the other hand, student "mnguy" started with an almost correct definition of the floor function but aside from stating the definition his proving attempt was unsuccessful. Since he does not use words and sentences to explain his work we had to assume that the first line represents the definition, but the rest of her/his work is mathematically false and even illogical.



6. Prove or disprove that $\lfloor 3x - 3 \rfloor = \lfloor 3x \rfloor - 3$, where x is a real number

(8 points)

Case 1

x is even ⟺ x = 2k

$\lfloor 3(2k) - 3 \rfloor = \lfloor 3(2k) \rfloor - 3$

$\lfloor 6k - 3 \rfloor = \lfloor 6k \rfloor - 3$

Figure 5.13: Example of *wrong (no) conclusion* mistake in midterm, 8th problem, by student "3.1415"

The last midterm problem, *Pythagorean Theorem*, is a well known problem

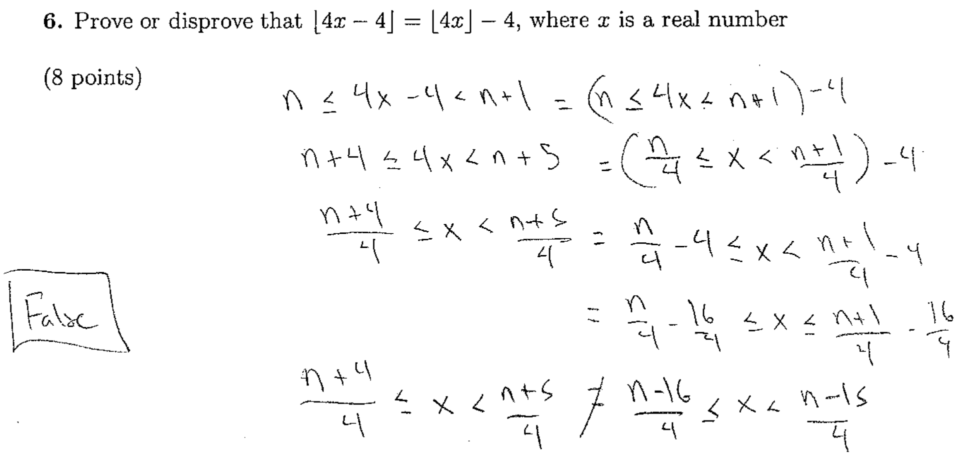


Figure 5.14: Example of *wrong (no) conclusion* mistake in midterm, 8th problem, by student “mnguy”

and numerous proofs are available. Earlier in the semester students were introduced to a few different proofs. The topic itself was not re-introduced during *methods of proof* sessions but a similar question was included in the post-questionnaire. Two thirds of students were unable to provide any type of proof. What is more interesting is that students were either able to mimic the complete proof as seen earlier or they did not even know where and how to start. There were no other mistakes in the proving process.

5.6 Summary

In this chapter we presented an extended table of common mistakes, Table 5.1, and their hierarchy, Figure 5.8, that allowed us to analyze students' work from the case study. Both, mistake categories and the mistake hierarchy, have been constructed by the researcher based on the observations of the lessons, collected data from the case study and students' textbook (Epp, 2003).

Results for each data instrument are presented in this chapter. We have presented results with two tables per each data instrument; the first table consists of numbers and percentages of correct answers while the second table is a list of mistakes per questions together with mistake occurrence in percentages.

Furthermore, there are scanned samples of students' unsuccessful proving attempts, such as Figure 5.11 representing a typical *solving for the unknown* mistake. For each mistake category we have provided at least one student example. Midterm results occupy the largest part of this chapter since the Midterm itself is the most extensive data instrument.

CHAPTER 6

DISCUSSION

In this chapter, we discuss the results presented in the previous Chapter 5. We provide analysis of each data instrument and conclude the chapter with comments on how students progressed and how their perceptions changed during the sessions.

6.1 Pre-teaching questionnaire

The pre-teaching questionnaire served as the introduction to students' understanding of mathematical proof before being introduced to different methods in the college curriculum. Since the methods of proof were not formally introduced we were hoping to gain an insight in to whether students understand when a proof is needed. From Table 5.4 we can see that students did very well on giving the correct answer, but from the following table 5.5 we realized that recognizing the conditions of the claim is the most difficult part of justifying their answer.

The pre-teaching questionnaire shows that students have good intuition, and based on previous mathematical experience they are able to provide correct answers. On the other hand, they did not demonstrate an understanding of the need for formal mathematical proof, and they based their mathematical beliefs on their intuitive understanding and previous mathematical experiences. Most students did not provide any justification, or they provided wrong arguments irrelevant to the claim; 34.6% of all mistakes were *wrong (no) conclusion* mistakes. The next most popular mistakes were: *intuitive proof*, (32%), and a*rguing from the example* (15.4%), both characteristic of earlier mathematical education. Examples for both mistakes are extracted from the first quiz: *Prove that if an integer* n *is odd then* $n^2 + 1$ *is even.* The three examples shown in figures 6.1, 6.2 and 6.3 are work from three students with pseudonyms "1244," "v193r" and "prince." As mentioned before, students chose the pseudonyms themselves and only through the researcher's database it is possible to connect an individual student to her/his work.

Example 6.1.1 (student 1244). [1]

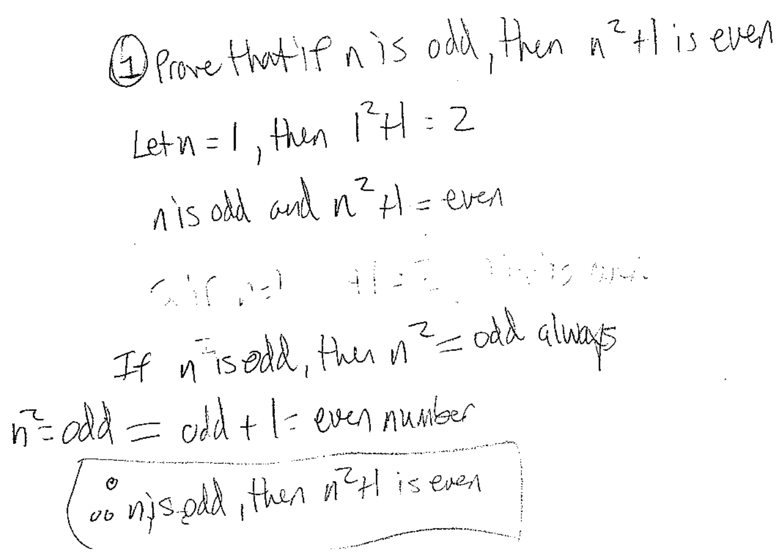


Figure 6.1: *Intuitive proof* mistake, made by student "1244" on the first quiz

[1] This is a scanned figure of actual student's work. Due to the scanning partially erased writing is slightly visible in the figure. This should be ignored while reading.

Example 6.1.2 (student v193r).

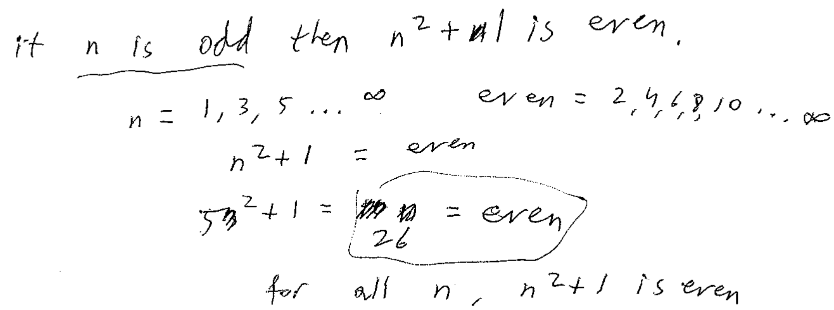


Figure 6.2: *Argue from example* mistake, made by student “v193r” on the first quiz

Example 6.1.3 (student prince). [2]

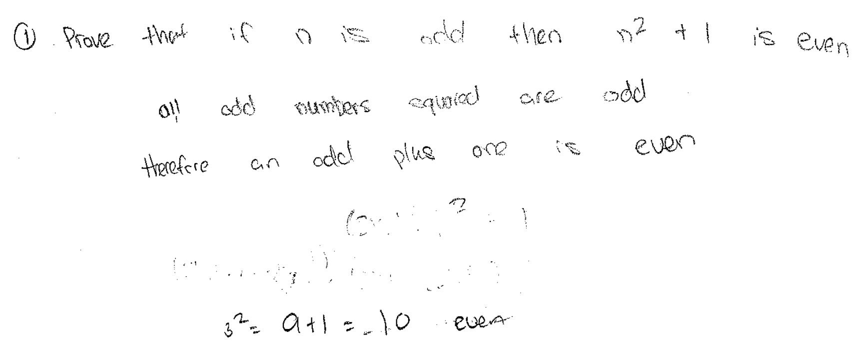


Figure 6.3: *Argue from example* mistake, made by student “prince” on the first quiz

[2] This is a scanned figure of actual student’s work. Due to the scanning partially erased writing is slightly visible in the figure. This should be ignored while reading.

Student 1244 takes the simplest example possible and he finds it convincing that if it works for the first choice in natural numbers it should work for every one. The second student, v193r, looks into all odd natural numbers (each number individually not as generally defined $n = 2k + 1$ for some $k \in \mathbb{Z}$) and picks 5 as a random example that provides the basis for his conclusion that it should work for every odd number. Finally the last student, prince as he calls himself, states a general conjecture that he does not prove but decides to justify with an example. It is interesting to note that student 1244 provided an intuitive proof using symbolic notation, while prince combines intuitive justification supported by an example and v193r gives us validation using a numerical example.

It is clear that all three examples are flawed and, though similar in error, each exhibits a different reasoning about generalization. None of the students in these examples uses a general definition of what being odd or even means and all three provide justification by one example. This way of justification is very common in middle school but should be completely abandoned in high school and rare in college. In the later data sources, we see that arguing from an example appears less often.

6.2 Quizzes

The two quizzes consisted of only one question and their significance was only to see if students were able to recognize and employ important proving steps as described in Section 3.3. On both, the most common mistake was *jumping to the conclusion*, 33.3% on the first and 56.5% on the second quiz, indicating that students did not hesitate to state a definition and employ one to argue about the claim. The difference between *jumping to the conclusion* and *wrong (no)* *conclusion*, which mostly occurs in the midterm and post-teaching questionnaire, lies in the students' confidence about how to start the proof. Such results indicate that if students are familiar with the subject they have no problem starting the proof while at the same time they feel almost presumptuous in believing that the claim is true, mostly based on the results seen in the teaching sessions and/or homework problems. As their teaching sessions progress and they learn about new topics they become less confident and *jumping to conclusion* is replaced by the *wrong (no) conclusion* mistake category.

6.3 Post-teaching questionnaire

An interesting result we got from the questionnaire 2 is that the majority of undergraduate college students in this case study failed to prove the first problem in the questionnaire: Prove that there exists an integer n such that $2n^2 - 21n + 40$ is prime, because they were unable to factor the binomial. The problem is very similar to a couple of problems they saw during the previous class session, and most of the students tried to use the same method. More students would succeed if their algebraic skills were a little more developed. Also, the problem is solvable using the quadratic formula but none of the students remembered to implement such basic knowledge to prove the statement.

Another problem that arose from the same questionnaire is that students did not know the meaning of “state the converse of the Theorem.” Surprisingly, almost all students answered the corresponding question correctly without realizing that the proof for that answer is in the converse they were supposed to state rather than in the given theorem.

Similar to the results from the pre-teaching questionnaire, students showed that on the intuitive level they can provide correct answers but are unable to prove or explain them mathematically. This conclusion is derived from the high percentage of *intuitive proof* mistakes (15%) which has the second highest rank in mistakes on the test, right after the *wrong (no) conclusion.* As opposed to the pre-teaching questionnaire where students based their intuition on the examples, on the second questionnaire *arguing from example* had no significant percentage, 6% only, and it was the lowest ranked mistake in the test. We are inclined to believe that students used previous experiences and what they remembered to be true from textbooks or high school mathematics rather then from numerical examples.

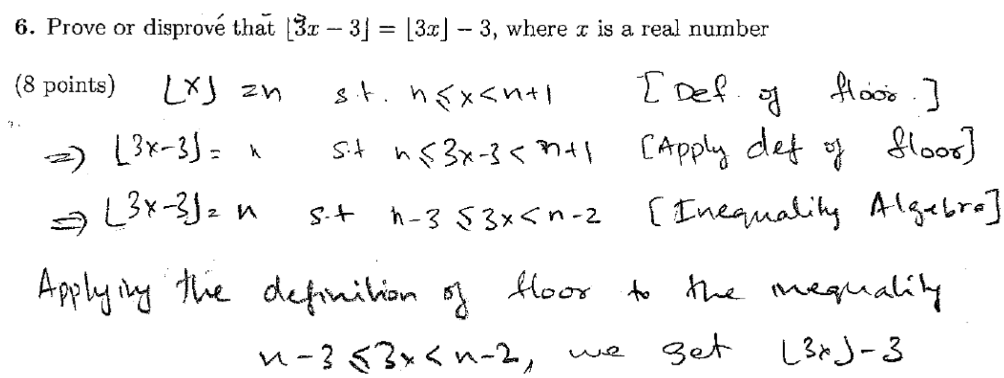
Such behavior implies that students do not realize the importance and power of mathematical proof in order to believe mathematical statements, but they did evolve from *arguing from example* which is the most basic level of understanding mathematical claims.

6.4 Midterm

The last and the most extensive data source is the Midterm exam where students were asked to provide proofs in different topics using different proving methods. The results, due to the mistakes and percentage of correct answers, were mostly as expected. Students obtained almost perfect scores on the two initial problems and the fifth problem, as can be seen in Table 5.9. Looking at the questions we should not be surprised about these results; the first two problems, see page 67, were very well known to the students from the first instruction session and they had many opportunities to see or even solve almost identical tasks. Also, the fifth problem, within the division topic, was very straightforward from the definition and almost all students proved it correctly from the definition and division properties.

Results on problems three, four, six and nine, where approximately half of the students failed to provide a formal mathematical proof, indicate that students are lacking the technical abilities to manipulate with the definitions and properties in order to reason deductively in more than two steps. Even though all of the problems were within the same topics as those three where they accomplished excellent results, these four problems required more sophisticated and complex use of definitions. To the experienced mathematician all eight problems would be the most simple and basic problems in the topic, but to the novice in this field these problems belong to the two different levels.

Finally, the remaining three problems, seven, eight and ten, show different difficulties students encountered when solving slightly more complicated mathematical assignments. The second of the two, problem number eight, dealt with the floor function, which we have already seen poses great difficulties to the students. The poor results indicate that students were still struggling with the definition and floor function properties. Also, a very high percentage of *wrong (no) conclusion* mistakes (81%) supports the idea that students need to affirm the new topic as the basic knowledge in order to be able to prove any further properties using the definitions and fundamental characteristics of the new mathematical material. On the other hand, the third of the problematic assignments, more precisely the poor results when proving the *Pythagorean Theorem*, combined with the high percentage of *wrong (no) conclusion* results, indicate that students got irritated when asked to think "out of the box," in other words to prove the theorem that was not covered in the most recent sessions but was introduced to students together with numerous proofs just before entering the methods of proof sessions. The difference between the different approaches to problems eight and ten can be seen in Figures 6.4 and 6.5. In the first one, most students, as the one in the example in Figure 6.4, at least stated the definition and tried, albeit unsuccessfully, to deduce the claim, while in the second problem most of the students did not even bother to explain their pictures, or even try to write down the intuitive proof, or to provide just partial work when they realized that they were unable to provide a full proof.


6. Prove or disprove that $\lfloor 3x - 3 \rfloor = \lfloor 3x \rfloor - 3$, where x is a real number

(8 points) $\lfloor x \rfloor = n$ s.t. $n \leq x < n+1$ [Def. of floor.]

$\Rightarrow \lfloor 3x-3 \rfloor = n$ s.t. $n \leq 3x-3 < n+1$ [Apply def of floor]

$\Rightarrow \lfloor 3x-3 \rfloor = n$ s.t. $n-3 \leq 3x < n-2$ [Inequality Algebra]

Applying the definition of floor to the inequality

$n-3 \leq 3x < n-2$, we get $\lfloor 3x \rfloor - 3$

Figure 6.4: Floor function solution example

Problem number seven in the Midterm exam was surprisingly low in number correct. Only 29% of students provided a correct mathematical proof. The topic of that assignment was divisibility and it required a direct proof; it should have been a fairly simple, straightforward proof following from the definition. The background of poor results on this problem might be in students' lack of confidence when asked to combine different mathematical topics, as in this case: divisibility and properties of exponential algebra. As we can see in Figure 6.6, the student had no problem constructing the direct proof for problem six, c), but at the same time he was unable to follow the same protocol in order

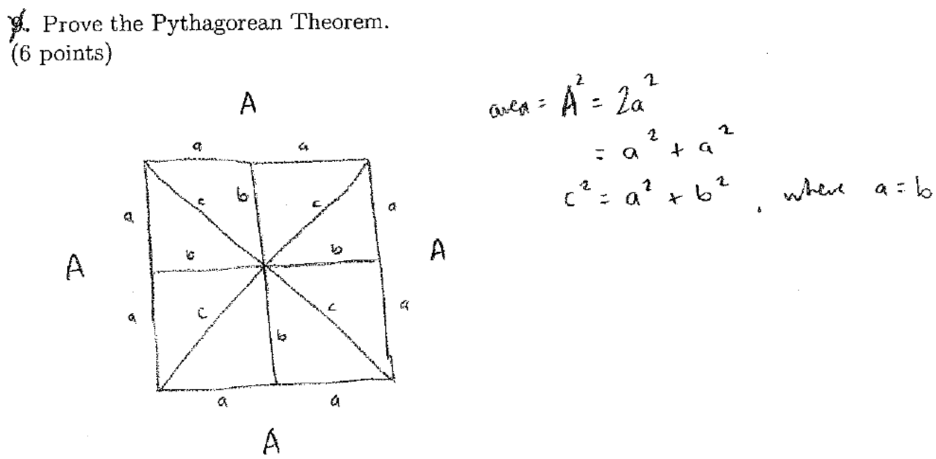


Figure 6.5: Pythagorean theorem solution example

to prove the divisibility property for just slightly complicated mathematical expressions.

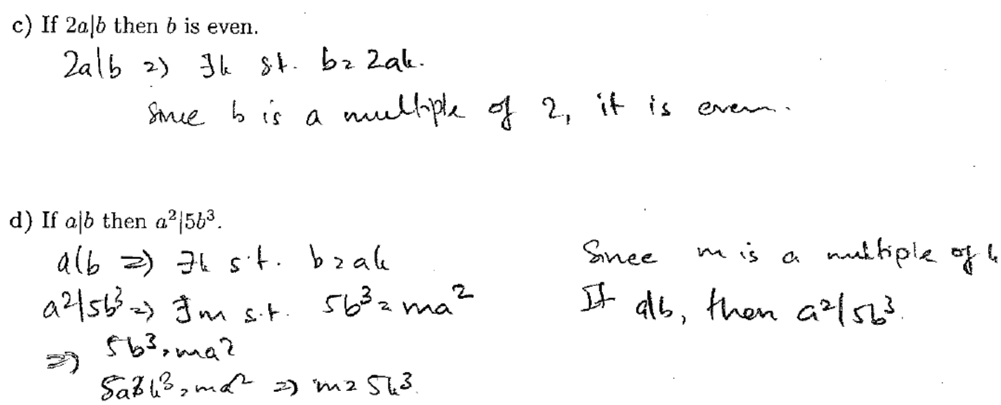


Figure 6.6: Example of problematic divisibility assignment solution

6.5 Cross data comparison

Having all five data instruments in mind we can see the progress students made in terms of mistakes. The first thing that comes to mind is to notice how much better students did on quizzes compared to the post-teaching questionnaire and midterm exam. One significant and obvious reason is that in quizzes they were asked to prove a claim that was discussed earlier in the teaching session, or the previous session, and a similar problem was given on the blackboard and/or for homework. Not only had students memorized the results and methods used to prove the quiz problems, but they were also more confident about the definitions and how to use them to start the proof itself. The topic of the first quiz problem, odd and even numbers, was common and they just needed to employ the definition, while on the other quiz they had to manipulate with the definition of the floor-function which was new to them. Having in mind the new topic we expected to have a slightly higher percentage of *wrong (no) conclusion* mistakes, since knowing and manipulating with the definition is integrated in that category. What is surprising is the extremely high percentage of the same mistake, namely 81%, on the midterm question similar to the quiz question (quiz #2). The increase in this mistake indicates that the new definition and the new topic were still unclear and confusing to the students. Clearly, students needed more time and practice with the new problems.

Each problem individually usually dictates the method of proof but sometimes students might have successfully opted for a different approach. That especially applies when the proof requires observations of different cases. The most typical proving method is the *direct proof*, and thereby it occupied the largest part of our case study. In other words, in 13 out of 20 questions given on all the assessments we expected to see direct proof, and in one problem students could choose between direct and proof by cases. Chronologically, the percentage of mistakes when constructing a direct proof decreased except in the case of the newly introduced topic, floor function. Also, when asked to prove a claim from number theory, which was the underlying content area for the unit, students preformed better as opposed to problems in geometry, e.g., Pythagorean Theorem. All the changes in the percentages of mistakes indicate that students had grown familiar with the method over time.

For other proving methods we have only one or two examples so it is difficult to discuss how well students accepted those methods. Table 6.1 lists desired proofs per tests and in the last column we have a total occurrence for each proving method. What can be said about students' work and other methods is that *proof by cases* seemed to be well accepted. The most common mistake that followed proving by cases is *jumping to conclusion.* In the context of proof by cases jumping to a conclusion usually stands for omitting certain cases. As the first step in the proving process students noted possible cases but after reaching a satisfactory conclusion in one case some decided to jump to the final claim.

Table 6.1: Type of desired proving method per test

Proof type	Test					Occurrence in %
	Quest 1	Quiz 1	Quiz 2	Quest 2	Midterm	
Direct	1	1	1	3	7	65
Vacuous	1	0	0	0	0	5
Trivial	1	0	0	0	0	5
Existential	0	0	0	1	0	5
By cases	0	0	0	0	3	15
By counterexample	0	0	0	0	1	5

From Table 5.3 we can see that an extremely high percentage of students had difficulties producing a vacuous proof; in other words they intuitively knew the correct answer but were unable to explain why there is nothing to be proved since the example stepped outside the claim's conditions.

Three other mistakes that appear in the study were: *computational mistakes*; *same identifier*; and *solving for the unknown. Computational mistakes* appear in very low percentages and the researcher finds no significant role of that mistake even though it is present in almost every problem. On the other hand, *same identifier*, appears only in four out of 19 problems, but the importance of such a mistake is in understanding the difficulties students encounter when constructing a proof. Even if a student understands the methods of proof and knows the steps, and even intuitively understands what needs to be proved and where to start, s/he still is not able to provide a correct mathematical proof due to the inability to translate his thoughts into formal mathematical language. Lastly, solving for the unknown appears only in two out of 19 instances. Since it appears in a such low percentage in both cases it would be unreasonable to discuss its significance.

In Table 6.2 we can see how often each mistake occurs per proving method, and in the last column are percentages of how often each mistake occurs in general throughout all the tests. As expected, *wrong (no) conclusion* mistake, has the highest occurrence in general, but that is not the case per each proving method. For example, we associate a much higher occurrence of *jumping to conclusion* with proof by cases method than *wrong (no) conclusion.*

6.6 Students' comments

During the second questionnaire students were very open to the researcher and they looked for hints or help in order to solve the problem. Since the researcher entered the study in a non-participant role, they were left to figure out

Table 6.2: Mistakes per proving method recorded in percentages

Mistake	Proof type						Occur. in %
	Direct	By cases	Vacuous	Trivial	Exis-tential	By counter-example	
Arguing from example	9.77	2.7	0	25	0	0	8.04
Same identifier	3.72	2.7	0	0	0	0	2.68
Jumping to conclusion	14.88	51.3	0	0	5.5	0	15.48
Begging the question	3.72	0	0	5	0	0	2.68
Intuitive proof	6.98	0	80	3.33	0	0	11.9
Solving for the unknown	3.26	0	0	0	0	0	2.08
Wrong (no) conclusion	53.02	32.4	20	65	88.9	93.75	52.38
Computational mistakes	4.65	10.8	0	0	5.88	5.88	4.76

the problem on their own. On the other hand, to gain a better understanding of students' difficulties they were asked to write in plain English about the problems they were facing, or frustrations they had with the assignment, maybe to put down speculations about how to solve the problem if they were unsure of the solution, or had no solution at all. We list all the comments sorted by problems. Problems can be found on page 59.

Problem 1 (1) This is probably wrong, isn't it?

(2) don't know.

(3) I forgot the Def of a prime and I would not know what to do with it even if I had it.

(4) I don't know where to continue.

(5) don't know

(6) Can't do this because factoring is hard!

(7) Can't seem to factor, seems so simple, but I can't find the factors.

Problem 2 (1) What is pipe? {pipe = symbol |}

(2) ... because it makes sense but I can't remember how to prove it though.

(3) don't remember the definitions

(4) I cannot find the right words to prove the statement.

(5) I need more time to master these problems.

(6) I have no clue where to start.

(7) ? $n \in \mathbb{N}$

(8) I don't know how to do it because I forgot.

Problem 3 (1) I can't remember how I solved these problems on HW. I need more practice.

Problem 4 (1) What is converse?

(2) Not sure what the converse is.

(3) No idea.

(4) Don't remember what converse is.

(5) I don't know the converse.

(6) don't know the converse.

(7) these are probably wrong. But I decided to do it anyway.

(8) Do not remember the converse.

(9) I forgot what a converse is.

(10) Don't know the converse.

(11) The converse of Pythagorean Theorem is (wrong formula). I forgot.

(12) Forgot what the converse is.

(13) ? Converse?

(14) ?huh?

(15) I can't recall what a converse of an algebraic expression is.

(16) I don't remember what the word "converse" means in this scenario so I'm not sure how to answer.

Comments that students left on their questionnaires just support the conclusions we have reached based on the mistake types and their occurrence. On the first problem students either proved (or almost proved) the statement or they where unable to start because of their lack of basic knowledge about binomials and how to factor one. On the other hand, problem number four confused most of them because almost half of the group did not know the meaning of the word "converse." Knowing the meaning and how to state the converse of a theorem does not depend on the proving methods and usually is taught separately much earlier in the high school curriculum. Furthermore, at the beginning of the course students were reintroduced to mathematical language, symbols and logic statements. The other half of the group did not provide a correct answer to that problem either, but they tried to state the converse, though unsuccessfully. Obviously, students need to be reminded more often about converses, how to state them and their meanings. Also, in addition to the first part of the problem where they were asked to state the converse, in the second part it was expected that students would implement their knowledge about the converse in order to draw a conclusion about the triplet given in the test. None of the students explained their answer using the converse, but they argued their answer incorrectly using the Pythagorean Theorem. Such practice indicates that students have none or very limited understanding of how the theorems can and should be used to support their answers. Algebraically they are capable of giving the right answer but they do not understand the theory behind the algebraic work.

6.7 Was there any progress?

Direct proof is the most employed proving method in the case study while other proving methods appeared only once or twice. As a consequence, direct proof is the only method that provides sufficient data for discussing students' progress. Overall, direct proof appears in 13 out of 19 problems. We have recorded six problems in the quizzes and questionnaires, and seven in the midterm, that were solvable using direct proof. The pre-teaching questionnaire and two quizzes were introduced before or during the teaching sessions while the post-teaching questionnaire and midterm were given to students after the instruction ended. Hence, the progress is discussed based on two categories: *pre/intermediate assessments* and *final assessment*. Both quizzes and the pre-teaching questionnaire constitute a *pre/intermediate* category and the post-teaching questionnaire and the midterm test form the *final* category.

Pre/intermediate In this category, we distinguish two groups of students; the first group consists of students who achieved one or fewer correct direct proofs ($\leq 50\%$) in *pre/intermediate* tests, while the rest form the second group who scored over 50%.

Final In this category we record how many correct direct proofs students accomplished in the post-teaching questionnaire and the midterm test.

Combined, both categories provide a ground for discussing students' progress in constructing a mathematical proof. We also need to note that some students missed one or more tests. In order to have more accurate results we exclude partial scores. Thus, only 24 results remained. Interestingly, half of the group (12 out of 24 students) scored 50% or less on *pre/intermediate* tests.

In Table 6.3 we can see the progress students made within *pre/intermediate* category. The progress is measured as a difference between percentages of correct answers before and during, and after instructions. In other words we have constructed the following formula to explain the progress:

$$\left(\frac{\text{final score}}{\#\ \textit{final}\text{ direct proofs}} - \frac{\text{pre/intermediate score}}{\#\ \textit{pre/intermediate}\text{ direct proofs}}\right) \times 100$$

for students who opted to use direct proof in the second midterm problem, where the number of *final* direct proof problems is 10, and the number of *pre/intermediate* direct proof problems is three. On the other hand, we used formula

$$\left(\frac{\text{final score}}{\#\ \textit{final}\text{ direct proofs}} - \frac{\text{pre/intermediate score}}{\#\ \textit{pre/intermediate}\text{ direct proofs}}\right) \times 100$$

for those who employed proof by cases in the second midterm problem. Thus, in this case there were nine *final* direct proof problems. If the obtained number is non-negative we say that the student showed progress while for a student whose number is negative we say s/he showed no progress.

Table 6.3: Students' progress in constructing direct proof; results for 24 students

	Pre/intermediate $\leq 50\%$		Pre/intermediate $> 50\%$	
	Progress	No progress	Progress	No progress
# students	12	0	8	4

Based on Table 6.3 we can say that in general instructions facilitated students' understanding and employment of direct proof as a proving tool; 83% of students improved their scores over time. But, there is a more significant difference between lower and higher scored *pre/intermediate* groups; 100% of low scored students advanced during teaching sessions vs 67% students with higher scores. In both cases the sample is too small to make final conclusions, but with a larger group size our results might be more conclusive.

6.8 Summary

In Chapter 5 we have discussed case study results based on students' answers on five data instruments: two questionnaires; two quizzes; and the midterm. Furthermore, a more detailed analysis of students' understanding of mathematical proof is based on mistakes students made, their occurrence and difficulties represented by each mistake category.

A comparison of students' achievement between different data instruments is provided in Section 6.5. We used that comparison to gain a better understanding of a relation between mistakes and proving methods. Furthermore, students' comments listed in Section 6.6 provided additional explanation of difficulties students encounter while constructing a mathematical proof. Finally, at the end of the chapter we have analyzed students' progress based on their test scores. Considering direct proof being the most employed proving method, students' progress was estimated only on problems requiring that method. To our satisfaction we can say that teaching had considerable impact on students' understanding and utilization of direct proof as a proving tool.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

In the previous chapter, we discussed mistakes and their occurrence on each data instrument as well as mistakes through all data sources combined. Some of the mistakes, such as *computational mistakes* and *solving for the unknown*, gave us no proper ground for discussion. On the other hand, *arguing from the example*, *wrong (no) conclusion* and *jumping to conclusion* provided better insight into how students understand mathematical proof. *Arguing from the example* was very common in the first data source, but it became less relevant later in the study indicating that students in a very short period of time realized that one or finitely many examples can not be taken as the proof but serve only to illustrate what is happening in individual cases and might happen in general. Looking at examples should be part of students' mathematical reasoning but they need to understand that while examples might serve as the foundation for generalization they never suffice as the formal proof.

Even though students should be able to use various types of reasoning and methods of proof by the end of high school (NCTM, 2000b), it seems that the majority of students hardly manage to manipulate with general mathematical conjectures and they leave proofs to textbooks, teachers and other mathematical authorities. The transition to college level mathematics, more precisely, to college level mathematical rigor, can be painful to some students and proof is one of the leading difficulties students have.

7.1 Answers to the research questions

At this point we can provide answers to the research questions proposed at the beginning of Chapter 4.

- *What type of proof do students accept as the most practical?*

  In most problems, students were expected to provide a direct proof, and the instructions were centered around the method, but we have noticed that even when students had an opportunity to prove the claim using another proving method most of them opted for the direct proof. Thus, we can say that the direct proof, as the proving method, is a primary type of proof. Also, most of the problems solved during teaching lessons were proved using direct proof, which might be the reason student were most comfortable using this method. Another reason to have the direct proof as the main proving method lays in the fact that the direct proof is a natural way of deducing what needs to be proved. One naturally starts with true statements and using logical reasoning comes to the desired conclusion.

- *What proving method do students find most complicated to use?*

  We are unable to answer this question with any certainty since our data sources are not extensive enough to cover all the methods. But we can notice from the first questionnaire that vacuous proof has no meaning to students. Vacuous proof is an argument that no proof is needed i.e. the statement to be shown is trivial. Students were asked how does the fact: "For $n = 0$ we have $(2 \cdot 0 + 1)(2^0 + 1) = 2$ and 2 is even" relates to the previous problem "For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$, $(2n + 1)(2^n + 1)$ is ODD." None of the students realized that $n = 0$ was not part of the statement to be proved, hence for $n = 0$ no proof is needed. Such realization would be a vacuous proof.

- *What common mistakes students do make*?

  The most common mistake as discussed in Chapter 6 and drawn from Table 6.2 are:

  (1) *Wrong (no) conclusions*

  There are many subcategories and we have discussed all in the previous chapters. The most important facts to emphasize in this category are students' inability to state and manipulate definitions needed to prove the statement, and lack of the basic mathematical knowledge such as the meaning of the symbols and mathematical expressions.

  (2) *Jumping to the conclusion*

  We have seen this mistake in two different methods: direct proof and proof by cases. The mistake has a different meaning and background in each type of proof. When it occurs in a direct proof, usually it comes from students' misunderstanding of what needs to be proved and inability to understand the process of deduction. On the other hand, when we have seen the mistake as part of a proof by cases, its significance was in the students' ignorance of all possible cases. Most students were satisfied with the positive conclusion in one of the cases but omitted the importance of other case(s).

- *Which mistakes exhibit a tendency to increase/decrease during and after the teaching sessions*?

  Two out of the six mistakes exhibited increasing/decreasing behavior during the study.

  (1) *Arguing from example*

  As mentioned earlier in Chapter 6, we have noticed that arguing from the example was a very common mistake before and at the beginning of the teaching sessions but students developed more abstract reasoning about examples quickly and abandoned such an approach. Some students still looked at the example(s) first just to be at ease with the assignment but proceeded with the general and formal proof as can be seen in Figure 7.1, (the proof in the example is almost correct, a student mixed identifiers in the last line.).

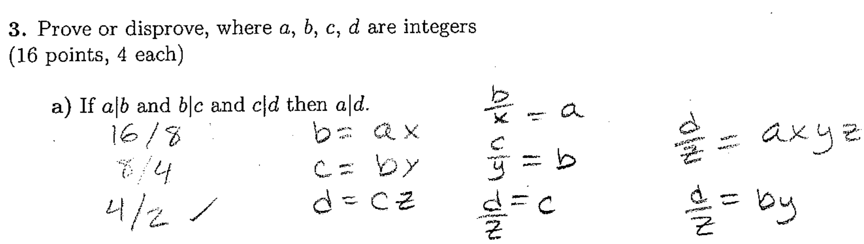


Figure 7.1: Numerical example followed by formal proof

  (2) *Wrong (no) conclusion*

  On the other hand *wrong (no) conclusion* showed increasing behavior toward the end of the study. We have discussed this increase in Chapter 6. If wrong conclusions were separated from no conclusions we might have gained better insight of where students failed to state the definitions and basic statements and when students had difficulties to manipulate with such statements and definitions. The reason we kept those two together was in the size of the sample. The study itself serves as the basis for future studies targeting more specific questions, listed on page 99, and in such we should be separating these two subcategories.

- *Which difficulties do students encounter when attempting to construct a valid mathematical proof in the early college curriculum*?

  The most important difficulty students face is the inability to manipulate with the definitions and basic claims that were proved earlier in the teaching sessions. Another troublesome point is students' insecurity in their mathematical knowledge from earlier education. The consequence of their insecurity is the inability to prove even a simple statement unless the proof can be conducted using the same principles and methods from the previous sessions and/or assignments. The best example of such behavior can be seen in Figure 7.2 about the post-teaching questionnaire problem solvable using the quadratic formula; only one out of 29 students remembered to employ the formula when they could not factor the trinomial. The third most prominent difficulty students faced during the study was the lack of basic mathematical knowledge when it comes to definitions, axioms and common knowledge. Students neglect the importance of definitions and axioms, taking them as theoretical knowledge that has no significant value in practice. They remember the definitions as the part of the proving process in individual, specific assignments and are unable to employ the same in a slightly different type of problem.

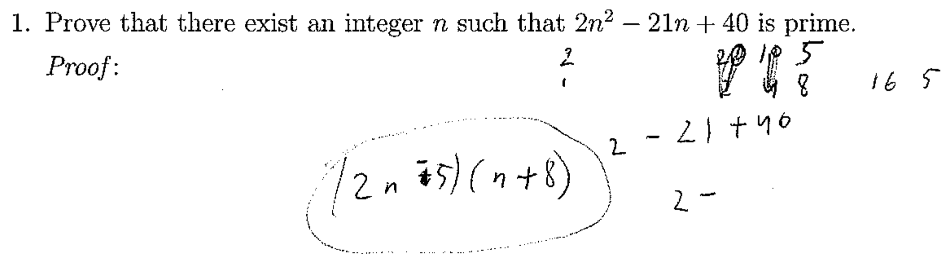


Figure 7.2: Student's inability to factor trinomial resulting in incomplete proof

7.2 Questions kindled by the study

Some of the answers to the research questions, as well as the observations based on the case study, stimulated other questions whose investigation might provide even better insight into the topic.

The questions that arise from observing behavior of arguing from example are:

"*Should the role of examples be discussed more often in high school?*",

"*Should teachers provide more problems where finitely many examples hold a certain property but not all numbers do?*"

"*How early in mathematical education can students accept examples only for what they are: examples, not a method of proof?*"

On the other hand, *wrong (no) conclusion* appears much more often in the problems where new topics, or assignments from different subjects, are introduced. The significance of such a phenomenon is that students need more time and practice to master the new topic in order to be able to prove related claims. Very often this category hides students' inability to move forward from the definition itself, either because they do not understand the definition or they lack the algebraic confidence to investigate the definition and conditions of the claim. The observed phenomenon raises the question: *Should the earlier mathematical education be more focused on mastering computational skills and practicing procedures and methods that allow students to do the technical parts of the problem solving faster?*. An even more important dilemma is: *How much time in K-12 mathematics curriculum should be devoted to practicing technical skills vs focusing on critical thinking and logical reasoning?*.

Another interesting observation came from the very low occurrence of *intuitive proof*. As the sessions progressed and students gained more knowledge about different proving methods they abandoned *intuitive proof*. Also, as the assignments got complicated students either succeeded in proving or they failed to build any logical and correct justification. There were no attempts to justify their reasoning in plain English and then translate such explanation into mathematical vocabulary.

In order to find the most efficient approach for teaching mathematics, the balance between practicing mathematical skills through numerous problems employing mathematical methods and insisting on logical reasoning should be found. Even if an individual is able to perform any mathematical operation and employ any mathematical procedure when given mathematical data, that is not insurance that one is able to construct a proof. To construct a formal mathematical proof of general statements one should be able to generalize and think in terms of abstract objects. On the other hand, if one is taught logical reasoning but lacks in mathematical knowledge, s/he is not able to start and continue even simple proofs. As we have seen in the study, the simple problem solvable using the quadratic equation became difficult to the students and almost no one even tried to employ such a simple but useful technique. Finally, from the difficulties we have noticed we can say that students do not realize the importance of the definitions and axioms. Even though they may have been taught the science of mathematics using definitions, only later in their education do teachers emphasize what is a definition and what is an axiom. Therefore, students start to think about definitions and axioms as a new part of mathematics instead of looking at them as the basis for any mathematical activity. They learn basic mathematical operations and take them as a solid grounding for arithmetic without realizing that such is possible due to the axioms and operation definitions in sets of numbers. Thus, another question that comes to my mind is: *At which educational level should students be introduced to official mathematical categorization using the formal language of definitions, axioms, theorems and such, in order to accept those as basic parts of mathematical procedures?*

7.3 Future work

Looking back into the study results we can list open questions that would be interesting and worthwhile investigating. For example, it would be interesting to see the results on the same, or similar, problems after approaching the methods of proof in a different way. There are two suggestions of how to alter the study to obtain more data.

- Encourage students to intuitively prove the claim and then translate such proof using mathematical symbols and language.
- After introducing a new topic, such as floor function in this case study,

solve or assign more homework problems combining the new topic and topics known to be difficult to students through high school, before asking them to construct proofs.

Another direction to investigate students' proving abilities and understanding the importance and function of proof is to investigate larger groups, in controlled samples where we could see whether there are differences between male and female students, younger (< 20 years of age) and older (> 20) students, or students taking a different number of mathematical courses. As is clear from Table 7.1 our study sample was too small to exercise inferential analysis of the data. Another issue with the group is that by gender the distribution of the group is in favor of male students (27 males vs seven female students). Another distribution category that exhibits such a gap is one-major students (29) vs two-major students (three). The complete distribution can be seen in Table 7.1.

As already mentioned, our sample size and distribution do not allow us to draw significant statistical conclusions, but they provide a hint of what might be worthwhile of further investigation. For example, in Table 7.2 we can see that there are some differences between female and male students. In general, we saw that, throughout all problems, female students had achieved better; out of all problems they made proving mistakes in 24.84% of problems vs male students who failed in 51.46% of the problems. Both groups had close percentages in most of the mistake categories except in *jumping to conclusion* and *solve for the unknown.* Male students made significantly more *jumping to conclusion* mistakes, where female students had a very low occurrence of this mistake. On the other hand, the *solve for the unknown* mistake occurred to a much higher percentage,

Table 7.1: Survey classification

Category	Sub-categories	# in the sample
Age	< 20	18
	≥ 20	16
Gender	Male	27
	Female	7
English proficiency	Native	22
	Non-native	12
# Majors	1	29
	2	3
Major 1	Mathematics	4
	Statistics	2
	Computer Science	12
	Software Engineering	3
	Computer Engineering	10
	Graphic Design	1
	Liberal Studies	1
	Math prep for Secondary Teaching	1
# mathematical courses	0	1
	1	9
	2	7
	3	7
	4	1
	5	6
	6	2
	7	1

7.89%, among female students (1.52% for males). It might be interesting to see if there would be the same or similar effect if our sample size and distribution were larger and symmetrical. Not only are the sample characteristics questionable in these observations, but also the low percentage of mistakes of this type might lead us to the wrong conclusions. *Wrong (no) conclusion* mistakes appear frequently in both categories, and there is a notable difference between females and males, 39.47% vs 51.52%. The grounds for such a gap might be one or more of the following:

- Female students are more diligent in learning definitions, thus more likely to start the proof.
- Female students are more confident to start the proof at least an intuitive one (18.42% intuitive proofs in females' vs 11.74% intuitive proofs in males' tests).
- Male students spend less time on the problems and miss the opportunity to come up with the solution even when not sure how to do it from the beginning.

All of the above presents valid research questions to which answers would give us an opportunity to better understand students' learning curve.

The third potential direction for future study might be in looking at the problems that can be proved using more that one method or approach. One might ask students to provide as many different proofs as they can think of in order to see which methods students prefer and to discover how to make other methods more popular or accessible to students. During such study students might be assisted with hints or help in order to motivate them in different

Table 7.2: Mistakes made by female vs male students

Mistake	% in female tests	% in male tests
Begging the question	2.63	3.41
Computational mistakes	7.89	4.17
Argue form Example	13.16	8.71
Intuitive proof	18.42	11.74
Jumping to conclusion	5.26	16.7
Same identifier	5.26	2.27
Solve for the unknown	7.89	1.52
Wrong (no conclusion)	39.47	51.52

directions. Furthermore, in this case study, one method, direct proof, was favored over others. In further investigations, problems should be altered to cover different methods equally in order to analyze students' progress more accurately.

7.4 Summary

In this chapter, five research questions were answered.

- *What type of proof do students accept as the most practical?* Direct proof.
- *What proof method do students find most complicated to use?* A vacuous proof seemed to be the most confusing method.
- *What common mistakes do students make?* Wrong (no) conclusion and jumping to the conclusion.
- *Which mistakes exhibit the tendency to increase/decrease during and after the teaching sessions?* Arguing from example showed a decreasing, while wrong (no) conclusion showed an increasing behavior.
- *Which difficulties do students encounter when attempting to construct a valid mathematical proof in the early college curriculum?* Inability to

manipulate with mathematical definitions, lack of self-confidence and inability to incorporate old (highschool) knowledge.

Furthermore, the study kindled further questions and we have suggested possible research directions toward addressing them.

BIBLIOGRAPHY

Boaler, J., & Humphreys, C. (2005). *Connecting mathematical ideas* (V. Merecki, Ed.). Portsmouth, NH: Heinemann.

Burger, W. F., & Culpepper, B. (1993, February). Restructuring geometry. In P. S. Wilson (Ed.), *Research ideas for the classroom: High school mathematics* (p. 140-154). Indianapolis, IN: MacMillan Reference Books.

Epp, S. S. (2003, December). The role of logic in teaching proof. *The American Mathematical Monthly*, *110*(10), 886-899.

Epp, S. S. (2004). *Discrete mathematics with applications, third edition.* Brooks/Cole Cengage Learning.

Glesne, C., & Peshkin, A. (1992). *Becoming qualitative researchers: An introduction.* NY: Longman Publishing Group.

Harel, G., & Sowder, L. (2007, April). Toward comprehensive perspectives on the learning and teaching of proof. In J. Frank K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (2Rev Ed edition ed., p. 805-842). Charlotte, NC: Information Age Pub Inc.

Hersh, R. (2009). What I would like my students to already know about proof. In M. L. B. D. A. Stylianou & E. J. Knuth (Eds.), *Teaching and learning proof across the grades* (p. 17-20). New York, NY: Routledge.

Huff, D. (1954). *How to lie with statistics.* 500 Fifth Ave, New York, N.Y.: W. W. Norton & Company, Inc.

Mason, J., Burton, L., & Stacey, K. (1982). *Thinking mathematically.* New York, NY: Addison-Wesley.

NCTM. (1989). *Curriculum and evaluation standards for school mathematics.* Reston, VA: NCTM.

NCTM. (1999). *Developing mathematical reasoning in grades K-12.* Reston, VA: NCTM.

NCTM. (2000a). *Principles and standards for school mathematics.* NCTM. Available from `http://standards.nctm.org/document/chapter3/reas.htm`

NCTM. (2000b). *Principles and standards for school mathematics.* Reston, VA: NCTM.

Polya, G. (1971). *How to solve it* (second edition ed.). Princeton University Press.

Schoenfeld, A. H. (2009). Series editor's foreword. In M. L. B. D. A. Stylianou & E. J. Knuth (Eds.), *Teaching and learning proof across the grades* (p. xii-xvi). New York, NY: Routledge.

Senk, S. L. (1989, May). Van hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education*, *20*(3), 309-321.

Wiliam, D. (1998). A framework for thinkging about research in mathematics and science education. In J. A. Malone, B. Atweh, & J. R. Northfield (Eds.), *Research and supervision in mathematics and science education* (p. 1-18). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

Zevenbergen, R. (1998). Etnography in the classroom. In J. A. Malone, B. Atweh, & J. R. Northfield (Eds.), *Research and supervision in mathematics and science education* (p. 19-38). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.

APPENDIX A

RESEARCH INSTRUMENTS

Survey

Name: ____________________ Student's Code: ____________________

Age: ___________________________

Please fill in the following fields.

(1) Gender: M F

(2) English Proficiency: Native speaker Non-native speaker

(3) Major 1: ____________________________________

(4) Major 2 (if applicable): ____________________________________

(5) Please list all college math courses you've taken. If at SJSU, please provide name and number of the course, including ones you are currently enrolled in. For college math courses taken on another college please provide only the name of the course.

(a) ______________________________

(b) ______________________________

(c) ______________________________

(d) ______________________________

(e) ______________________________

(f) ______________________________

A.1 Questionnaires

Student's Code: ____________________

Questionnaire 1

(1) For all natural numbers n in the set $\mathbb{N} = \{1, 2, 3, \ldots\}$,

$(2n + 1)(2^n + 1)$ is ODD.

The statement is:

(a) True

(b) False

Answer: ________________

Please explain you answer.

(2) For $n = 0$ we have $(2 \cdot 0 + 1)(2^0 + 1) = 1 \cdot 2 = 2$ and 2 is even.

How does this fact relate to the previous problem?

(a) It is irrelevant to the previous problem.

(b) It is a counterexample we can use to prove that the statement in 1. is not true.

(c) It is a special case of the previous problem.

Answer: ________________

Please explain your answer.

(3) For all positive real numbers a and b, the following is true:

$$\sqrt{a+b} < \sqrt{a} + \sqrt{b}.$$

Which of the following statements can be deduced:

(a) There exist $a, b > 0$ such that $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

(b) There are no $a, b > 0$ such that $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

(c) For all $a, b \in \mathbb{R}$ $\sqrt{a+b} = \sqrt{a} + \sqrt{b}$.

(d) None of the above.

Answer: ________________

Please explain your answer.

Student's Code:

___

Questionnaire 2

(1) Prove that there exist an integer n such that $2n^2 - 21n + 40$ is prime.

Proof:

(2) Prove that for all integers a and b if $a|b$ then $a^n|b^n$ for all $n \in \mathbb{N}$.

Proof:

(3) Is this true or false? **Answer**: ___

For all integers n, $6n + 1$ is not divisible by 3.

Justify your answer.

(4) **Pythagorean theorem**

In a right triangle with c representing the length of the hypotenuse, and a and b representing the lengths of the other two sides it holds that:

$$a^2 + b^2 = c^2.$$

State the converse of the Pythagorean Theorem: (*we know that the converse of Pythagorean Theorem is also true*)

___

___

Can $a = 13$, $b = 84$, $c = 85$ be lengths of the sides of a right triangle?

Justify your answer.

A.2 Quizzes

Name:________________________________

Quiz 1

Prove: If n is an odd integer then $n^2 + 1$ is even.

Name:________________________________

Quiz 2

For n an odd integer prove that

$$\lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}.$$

A.3 Midterm

Student's Code:

Midterm E

Show all work.

(1) Prove the following where n is an integer.

(a) If n is even, then $n^3 + 2$ is even.

(b) 2 does not divide $n^2 + (n + 1)^2$.

(c) 4 divides $n^2 + (n + 2)^2$ if and only if n is even.

(3) Prove or disprove, where a, b, c, d are integers.

(a) If $a|b$ and $b|c$ and $c|d$ then $a|d$.

(b) If $2a|b$ then b is even.

(c) If $a|2b$ then a is even.

(d) If $a|b$ then $a^2|4b^4$.

(6) Prove or disprove that $\lfloor 4x - 4 \rfloor = \lfloor 4x \rfloor - 4$, where x is a real number.

(7) Prove that 3 divides $n^3 + 3n^2 + 5n$ for all integers n.

(9) Prove the Pythagorean Theorem.

Student's Code:

Midterm O

Show all work.

(1) Prove the following where n is an integer.

(a) If n is odd, then n^2 is odd.

(b) 2 divides $n^2 + (n+2)^2$.

(c) 4 divides $n^2 + (n+2)^2$ if and only if n is even.

(3) Prove or disprove, where a, b, c, d are integers.

(a) If $a|b$ and $b|c$ and $c|d$ then $a|d$.

(b) If $a|2b$ then a is even.

(c) If $2a|b$ then b is even.

(d) If $a|b$ then $a^2|5b^3$.

(6) Prove or disprove that $\lfloor 3x - 3 \rfloor = \lfloor 3x \rfloor - 3$, where x is a real number.

(7) Prove that 3 divides $n^3 + 3n^2 + 2n$ for all integers n.

(9) Prove the Pythagorean Theorem.