Mälardalen University Doctoral Dissertation 346

Proof-related reasoning in upper secondary mathematics textbooks

Characteristics, comparisons, and conceptualizations

Andreas Bergwall

MÄLARDALEN UNIVERSITY
SWEDEN

Mälardalen University Press Dissertations
No. 346

PROOF-RELATED REASONING IN UPPER SECONDARY MATHEMATICS TEXTBOOKS

CHARACTERISTICS, COMPARISONS, AND CONCEPTUALIZATIONS

Andreas Bergwall

2021

MÄLARDALEN UNIVERSITY
SWEDEN

School of Education, Culture and Communication

Copyright © Andreas Bergwall, 2021
ISBN 978-91-7485-524-1
ISSN 1651-4238
Printed by E-Print AB, Stockholm, Sweden

Mälardalen University Press Dissertations
No. 346

PROOF-RELATED REASONING IN UPPER SECONDARY MATHEMATICS TEXTBOOKS
CHARACTERISTICS, COMPARISONS, AND CONCEPTUALIZATIONS

Andreas Bergwall

Akademisk avhandling

som för avläggande av filosofie doktorsexamen i matematik/tillämpad matematik vid Akademin för utbildning, kultur och kommunikation kommer att offentligen försvaras fredagen den 12 november 2021, 09.15 i Delta samt digitalt via ZOOM, Mälardalens högskola, Västerås.

Fakultetsopponent: Professor Kristina Juter, Högskolan Kristianstad

MÄLARDALEN UNIVERSITY
SWEDEN

Akademin för utbildning, kultur och kommunikation

Abstract

Proofs and proving are difficult to learn and difficult to teach. A common problem is that many students use specific examples as evidence for general statements. Difficulties with proofs are also part of the transition problems that exist between secondary and tertiary schooling in mathematics. As mathematics teaching often follows a textbook, the design of textbooks has been pointed out as one possible cause of the problems, and international textbook research suggests that proofs often have only a marginal place in textbooks.

This thesis focuses on proofs and proving in upper secondary mathematics textbooks. It also addresses theoretical and methodological questions about what marks an opportunity to develop proving competence, and which properties of such opportunities are relevant to investigate and characterize. The thesis is based on data from four Swedish and Finnish textbook series for upper secondary school, and focuses on sections on logarithms, primitive functions, definite integrals, and combinatorics. It examines how addressed mathematical principles are justified, and whether the textbooks' exercises offer opportunities to develop proof-related skills such as formulating and investigating hypotheses, developing and evaluating arguments, identifying and correcting errors, and finding counterexamples.

The results show that just over half of the mathematical principles addressed in the analyzed textbook material are justified, and that only half of the justifications are general proofs. Few exercises are proof-related (10%), and those that include reasoning about general cases even fewer. General proofs are more common in the Finnish books, but proof-related tasks are more common and of a more varied nature in the Swedish ones. The most common form of proofs are direct derivations of calculation formulas, while reasoning about existence and uniqueness is unusual, as are contrapositive proofs and proofs by contradiction.

Based on the results, explicit suggestions are offered as to what teaching can pay more attention to. For the analysis and design of proof-related activities, a framework consisting of four main categories is proposed: develop a statement, investigate a statement, develop an argument, and investigate an argument. Several properties that such activities may have, regardless of which category they belong to, are discussed. Finally, three areas for future research are suggested: how worked examples can support students' learning of proof, how textbooks can be designed to stimulate formulation as well as the formal proving of hypotheses, and mapping of differences regarding proof between upper secondary and university textbooks.

ISBN 978-91-7485-524-1
ISSN 1651-4238

To all ye faithful

Nothing pertaining to humanity becomes us so well as mathematics. There, and only there, do we touch the human mind at its peak.

Isaac Asimov

From Isaac Asimov's foreword in C. B. Boyer & U. C. Merzbach's *A history of mathematics,* 1991.

I may never find all the answers,
I may never understand why,
I may never prove what I know to be true,
but I know that I still have to try.

John Petrucci

From the song *The spirit carries on,* on Dream Theater's album *Metropolis pt 2: Scenes from a memory*, 1999.

Abstract

Proofs and proving are difficult to learn and difficult to teach. A common problem is that many students use specific examples as evidence for general statements. Difficulties with proofs are also part of the transition problems that exist between secondary and tertiary schooling in mathematics. As mathematics teaching often follows a textbook, the design of textbooks has been pointed out as one possible cause of the problems, and international textbook research suggests that proofs often have only a marginal place in textbooks.

This thesis focuses on proofs and proving in upper secondary mathematics textbooks. It also addresses theoretical and methodological questions about what marks an opportunity to develop proving competence, and which properties of such opportunities are relevant to investigate and characterize. The thesis is based on data from four Swedish and Finnish textbook series for upper secondary school, and focuses on sections on logarithms, primitive functions, definite integrals, and combinatorics. It examines how addressed mathematical principles are justified, and whether the textbooks' exercises offer opportunities to develop proof-related skills such as formulating and investigating hypotheses, developing and evaluating arguments, identifying and correcting errors, and finding counterexamples.

The results show that just over half of the mathematical principles addressed in the analyzed textbook material are justified, and that only half of the justifications are general proofs. Few exercises are proof-related (10%), and those that include reasoning about general cases even fewer. General proofs are more common in the Finnish books, but proof-related tasks are more common and of a more varied nature in the Swedish ones. The most common form of proofs are direct derivations of calculation formulas, while reasoning about existence and uniqueness is unusual, as are contrapositive proofs and proofs by contradiction.

Based on the results, explicit suggestions are offered as to what teaching can pay more attention to. For the analysis and design of proof-related activities, a framework consisting of four main categories is proposed: develop a statement, investigate a statement, develop an argument, and investigate an argument. Several properties that such activities may have, regardless of which category they belong to, are discussed. Finally, three areas for future research are suggested: how worked examples can support students' learning of proof, how textbooks can be designed to stimulate formulation as well as the formal proving of hypotheses, and mapping of differences regarding proof between upper secondary and university textbooks.

Sammanfattning

Matematisk bevisföring är svårt att lära sig och svårt att undervisa. En vanlig problematik är att många elever och studenter ser enstaka exempel som tillräckliga argument för generella påståenden. Svårigheter med bevisföring är också en del av den övergångsproblematik som finns mellan gymnasie- och universitetsstudier i matematik. Eftersom undervisningen ofta följer en lärobok så har böckernas utformning pekats ut som en möjlig orsak till problemen. Internationell läromedelsforskning antyder också att bevis ofta har en undanskymd plats i böckerna.

Den här avhandlingen bidrar med kunskaper om hur bevis hanteras i gymnasieböcker i matematik. Den tar också upp teoretiska och metodologiska frågor om vad som avses med att en bok erbjuder möjligheter att utveckla kompetens inom bevisområdet, och vilka egenskaper hos sådana möjligheter som är relevanta att undersöka och karaktärisera. Avhandlingen bygger på studier av fyra svenska och finska läromedelsserier för gymnasieskolans teoretiska program och fokuserar på avsnitt om logaritmer, primitiva funktioner, bestämda integraler och kombinatorik. Dels undersöks hur de matematiska principer som tas upp motiveras, dels om böckernas övningsuppgifter ger eleverna möjligheter att utveckla bevisrelaterade färdigheter såsom att formulera och undersöka hypoteser, argumentera för påståenden, undersöka giltigheten i presenterade argument, identifiera och korrigera felaktigheter i resonemang och konstruera motexempel.

Resultaten visar att drygt hälften av de matematiska principer som behandlas i det analyserade materialet motiveras, men att bara hälften av motiveringarna är generella bevis. Få uppgifter är bevisrelaterade (10 %) och de som inbegriper generella resonemang ännu färre. Generella bevis och resonemang är vanligare i de finska böckerna, men bevisrelaterade uppgifter är vanligare och av mer varierad karaktär i de svenska. Den vanligaste formen av bevis är direkta härledningar av beräkningsformler medan resonemang om existens och entydighet är ovanliga, liksom indirekta bevis och motsägelsebevis.

Med resultaten som grund ges konkreta förslag på vad undervisning kan ägna mer uppmärksamhet åt. För analys och utveckling av bevisrelaterade aktiviteter föreslås ett ramverk bestående av fyra huvudkategorier: utforma påståenden, undersöka påståenden, utforma argument och undersöka argument. En rad egenskaper som aktiviteterna kan ha oavsett vilken kategori de tillhör diskuteras. Avslutningsvis ges förslag på vidare forskning inom tre områden: hur lösta exempel kan stötta elevers lärande av bevis, hur läroböcker kan utformas för att stimulera till att formulera såväl som att bevisa hypoteser, samt kartläggning av hur gymnasie- och universitetsböcker skiljer sig åt i hanteringen av bevis.

Acknowledgements

Once, I was told that there are only two ways to exit a PhD program: dissertate or die. As it now happens, there's a good chance that the former is about to happen prior to the latter. But it's taken its time, having begun more than 25 years ago when I was enrolled in a PhD program in applied mathematics at Linköping University. Well, there are neither beginnings, nor endings, to the turning of The Wheel of Time, but it was *a* beginning. On that first try I didn't make it to a PhD degree, but fortunately, I got a second chance at Mälardalen University some ten years ago. Anyway, as so many years have passed since it all began, by now there are quite a few who deserve remembrance for the encouragement and support they've offered me throughout the years. Here I can only mention a few, however, and will restrict myself to those who have been of special importance during this second effort to complete my PhD.

My sincerest thanks go to my supervisors at Mälardalen University. First of all, Andreas Ryve, supervisor and friend, you've always trusted in my abilities, encouraged me, and without showing any disappointment, accepted all the circumstances that have delayed the work with this thesis. Your analytical sharpness and ability to pinpoint the essentials are enviable and have been crucial to my work. Kirsti Hemmi, you suggested the topic of this thesis and very kindly led me through the writing of the first (and best) of its papers. Kimmo Eriksson, you've kindly offered advice whenever I've requested it.

I'm also very thankful to Ewa Bergqvist at Umeå University and Jonas Bergman Ärlebäck at Linköping University for your thorough reading of and constructive feedback on earlier versions of the thesis.

I'm grateful for all the people I've come to know at Mälardalen University, all the interesting discussions we've had in the M-TERM group, and the patience you've shown me during my lengthy presentations. A special thanks to Maria and to Hendrik, for our close collaboration on projects for Skolverket and Skolforskningsinstitutet.

More recently I've been involved in a combined research and school development project financed by the municipality of Örebro. Elisabet, it's been a joy co-driving this project with you, and trying to match your tempo.

It's a privilege to have so many colleagues at the mathematics department at Örebro University, colleagues I consider dear friends and whom I admire and find so much joy and inspiration from working with. Thanks for the time you've spent reading and commenting on my work, and for all your encouragement. Malin, it made it so much easier that it was two of us taking on this endeavor together, even though I had to carry your bags. Per, having an extra professor at my side, reminding me to have a beer now and then, cannot be overestimated, but you'll never make me a dancer. Anna, you probably hold the high score in asking me when I'm going to finish my thesis. Marcus, always with a new pedagogical idea to test and the one to lean on at choir rehearsals. Anders, thoughtful, and someone to process your free church upbringing with – while skiing. Frida, without you all those exams would have killed me long ago. The other Andreas, younger, better looking, and more fit, so unfortunately no risk of confusion there. Niklas, the oracle that Google turns to, you took time I know you didn't have in order to help me free up the time I needed. I guess I could go on, but I choose to stop there. Hopefully I haven't forgotten anyone I've closely collaborated with on our mathematics education courses. Otherwise, please forgive me.

But nothing would mean anything if it weren't for those I turn to now. Here words will never be enough, so I'll keep it short. When it all began, back in 1995, you were toddlers. In the acknowledgements of my licentiate thesis in 1998, I wrote that you had taught me to appreciate the rare moments of peace and silence. Those moments are more frequent now. You are grownups, have studied at the university, and have your own kids. Isac, Elias, and Elin, you'll always be my greatest pride, as will your families and the marvelous grandchildren you've blessed me with so far: Otto, Vira, and Otis.

Finally, and forever, to the most beautiful woman I know, the love of my life: Du och jag, Lena.

Nora, October 2021
Andreas Bergwall

List of Papers

This thesis is based on the following five papers, which are referred to in the text by their Roman numerals. All of them relate to proof-related reasoning in Swedish and Finnish upper secondary school textbooks.

I Bergwall, A., & Hemmi, K. (2017). The state of proof in Finnish and Swedish mathematics textbooks—capturing differences in approaches to upper secondary integral calculus. *Mathematical thinking and learning, 19*(1), 1–18. https://doi.org/10.1080/10986065.2017.1258615

II Bergwall, A. (2015). On a generality framework for proving tasks. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Conference of the European Society for Research in Mathematics Education (CERME9, 4–8 February 2015)* (pp. 86–92). Prague, Czech Republic: Charles University in Prague, Faculty of Education and ERME.

III Bergwall, A. (2017). Conceptualizing reasoning-and-proving opportunities in textbook expositions: Cases from secondary calculus. In T. Dooley & G. Gueudet (Eds.), *Proceedings of the Tenth Congress of the European Society for Research in Mathematics Education (CERME10, February 1–5, 2017)* (pp. 91–98). Dublin, Ireland: DCU Institute of Education & ERME.

IV Bergwall, A. (2021). Proof-related reasoning in upper secondary school: characteristics of Swedish and Finnish textbooks. *International Journal of Mathematical Education in Science and Technology, 52*(5), 731–751. https://doi.org/10.1080/0020739X.2019.1704085

V Bergwall, A. (manuscript). Topic specific aspects of proof-related reasoning—cases from upper secondary mathematics textbooks.

Reprints are made with permission from the respective publishers.

The following two texts, reporting research I have been involved in during my graduate studies, are related to the present work and will be referred to, but the connections are not strong enough for them to be included in the thesis.

Paper VI is a comparative study between Sweden and Finland. Paper VII includes a systematic literature review of research on mathematics textbooks.

VI Knutsson, M., Hemmi, K., Bergwall, A., & Ryve, A. (2013). School-based mathematics teacher education in Sweden and Finland: characterizing mentor–prospective teacher discourse. In B. Ubuz, C. Haser, & M. A. Mariotti (Eds.), *Proceedings of the Eighth Congress of the European Society for Research in Mathematics Education (CERME 8, February 6–10, 2013)* (pp. 1905–1904). Ankara, Turkey: Middle East Technical University and ERME.

VII Ryve, A., Nilsson, P., Palm, T., Van Steenbrugge, H., Andersson, C., Bergwall, A., Boström, E., Larsson, M., & Vingsle, L. (2015). Kartläggning av forskning om formativ bedömning, klassrumsundervisning och läromedel i matematik [Mapping of research on formative assessment, classroom teaching, and curriculum programs in mathematics]. Stockholm: Vetenskapsrådets rapportserie.

Finally, the first half of my graduate studies in applied mathematics focused on mathematical modelling of compression molding of polymers. This resulted in one journal article, which is not accounted for in this thesis:

VIII Bergwall, A. (2002). A geometric evolution problem. *Quarterly of Applied Mathematics, 60*(1), 37–73. https://doi.org/10.1090/qam/1878258

Contents

1 Introduction

The research presented in this thesis lies at the intersection of two important fields of educational research in mathematics: the teaching and learning of proof and mathematics textbooks.

Proofs are central in mathematics: “The possibility of proof is what makes mathematics what it is, what distinguishes it from other varieties of human thought” (Hersh, 2009). There is also a wide consensus among scholars that proofs are important in learning mathematics: “Students cannot be said to have learned mathematics, or even *about* mathematics, unless they have learned what a proof is” (Hanna, 2000). However, just as central as proofs are to mathematics, students’ difficulty understanding and producing them as well as teachers’ challenges in teaching them are well-documented (e.g., Harel & Sowder, 2007). Among other problematic perceptions, students tend to justify general statements with specific cases, view counterexamples as exceptions, and believe conditional statements to be equivalent to their inverses and converses (e.g., Stylianides et al., 2017).

Research has also shown that proofs and proving have a marginal place in many classrooms, and mathematics textbooks have been pointed out as one possible reason for this (Stylianides et al., 2017). The extensive use of mathematics textbooks in classrooms around the world is well-known (e.g., Stein et al., 2007). Teachers rely on them in planning and conducting their teaching, and students spend a great deal of lesson time working with textbook material. It has been said that the textbook is “the only variable that on the one hand we can manipulate and on the other hand does affect student learning” (Begle, 1973).

Historically, textbook research focusing on proof has been rare (Hanna & de Bruyn, 1999). This branch of educational research is still small, but growing. The overall impression is that opportunities to learn proofs and proving from mathematics textbooks are small, especially in topics other than geometry (e.g., Otten, Gilbertson, et al., 2014; Thompson et al., 2012). However, these learning opportunities in a topic such as calculus, which dominates upper secondary school in most parts of Europe, have seldom been studied.

Textbook studies of opportunities for proof-related reasoning also actualize questions of how proof and proof-related reasoning can be conceptualized and operationalized for textbook analysis, how observed similarities and differences between different textbook materials can be understood and explained, and what the consequences for student learning and teaching practice might be.

1.1 Aim of the thesis

This thesis aims to contribute to an informed discussion about opportunities to learn proof-related reasoning offered in mathematics textbooks. The papers in the thesis have their separate research questions, but brought together they can be summarized as follows:

1. What characterizes opportunities to learn proof-related reasoning offered by Swedish and Finnish upper secondary mathematics textbooks?
2. What topic-specific characteristics of opportunities to learn proof-related reasoning are found in upper secondary mathematics textbooks?
3. How can opportunities to learn proof-related reasoning be conceptualized and analyzed?

The questions are interrelated. To answer them, four textbook series for Swedish and Finnish upper secondary school have been studied. The data includes expository sections and student exercises on logarithms, primitive functions, definite integrals, and combinatorics.

1.2 How to read the thesis

1.2.1 Connections between the papers

The thesis consists of five papers. Papers I and IV are published journal articles, while Papers II and III are conference papers. Paper V is a manuscript intended for publication as a journal article.

Paper I is foremost a comparative study of Swedish and Finnish textbooks. Opportunities for proof-related reasoning are studied in relation to two topics: primitive functions and definite integrals. The analyzed material includes expository sections and student exercises. The other papers are related to questions that arose during the work with Paper I. Two of them have a theoretical-analytical nature, while the other two are related to how the findings of Paper I can be generalized and understood.

Paper II explores how the “degree of generality” in student exercises can be conceptualized and analyzed in a more precise way, while Paper III explores how the “embedding” of justifications in expository sections can be analyzed to give more accurate descriptions of opportunities to learn proof-related reasoning. Paper IV is an extension of Paper I, and concerns the generalizability of the results reported in Paper I. For this purpose, Paper IV includes two new topics in the analysis: logarithms and combinatorics. Finally, Paper V builds on the same data as Papers I and IV, but focuses on differences between topics (instead of between Swedish and Finnish textbooks) and how these can be understood. The connections between the papers are illustrated in *Figure 1*.

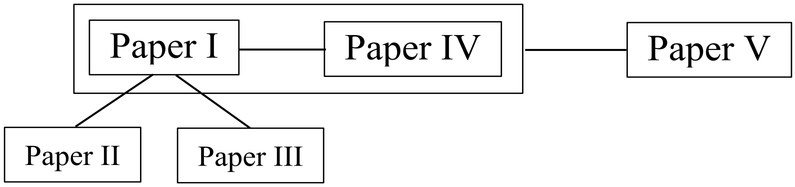

Figure 1. Connections between the papers.

1.2.2 Structure of the thesis

Chapter 2 (Literature review) gives an overview of relevant research on the teaching and learning of proof and textbooks. It concludes with some general findings about Swedish and Finnish mathematics education.

Chapter 3 (Analytical concepts and frameworks) first clarifies how opportunity to learn is interpreted in relation to proof-related reasoning in textbooks. Then follows a description of the frameworks for the type and nature of reasoning that are used in each of Papers I, IV, and V. A separate section is devoted to the specific frameworks and tools that are used only in Papers II, III, and V.

Chapter 4 (Methodology) has a shared focus between three different areas. The first is the organization of and steering documents for Swedish and Finnish upper secondary school. The second is the data set and reasons for the choice of textbook series and mathematics topics. Finally, the analytical procedures are described.

The papers are summarized in Chapter 5 (Summary of papers). While the focus is on their results, motives and research questions are also included.

Finally, the results of the five papers are discussed in combination in Chapter 6 (Discussion). The empirical findings are organized in three themes: general findings and context-specific findings relate to the first research question, while topic-specific findings relate to the second. A separate section is devoted to the development of the analytical frameworks and the conceptualization of proof-related reasoning, which relates to the third research question. With this as a basis, implications and suggestions for teaching practice and future research are discussed, as is the thesis's methodology.

2 Literature review

This chapter includes theoretical material with a focus on what a proof is, what the role of proof is in mathematics, and how proof and proof-related reasoning are conceptualized in the mathematics education literature. It also includes empirical findings on the teaching and learning of proof and from textbook research. Finally, some findings from comparative studies of mathematics education in Sweden and Finland are included.

2.1 Proofs in mathematics

Deductive arguments have a long history in mathematics and are at the core of what mathematics is. According to tradition, deductive elements were first introduced in mathematics by Thales of Miletus (ca. 624–548 B.C.), earning him the title "the first mathematician". Ancient sources tell that he provided a demonstration of the theorem now called the Thales theorem. However, there are no records of this demonstration, and it has been argued that Greek mathematics at the time were too primitive for such a contribution (Boyer & Merzbach, 1991). Nevertheless, the Greeks are considered to be those who brought to the Western world the notion that mathematical facts must be established deductively. This radically transformed the empirical mathematics of the Babylonians and Egyptians (Stylianou et al., 2009).

The earliest major Greek mathematical work to come down to us is Euclid's Elements. Composed about 300 B.C., it is the most influential textbook of all times with its 465 propositions on geometry, number theory, and (in a sense) algebra, all set up as a deductive system. It has set the standard for mathematical rigor that we observe to this day. A mathematics proposition is not accepted as a theorem unless someone can present a valid and convincing demonstration of how it follows from axioms, assumptions, and other proven propositions. Such a demonstration is a *proof*.

In mathematical logic, a proof is "a sequence of transformations of formal sentences, carried out according to the rules of the predicate logic", while for the common practicing mathematician it is "an argument that convinces qualified judges" (Hersh, 1993, p. 389). A more precise definition of proof upon which the mathematics community can agree is difficult to formulate. So are clear criteria for what should be accepted as "convincing". What counts as a proof may vary over time, between groups of mathematicians and between different fields of mathematics. The required level of formalism, rigor, and detail also depends on the knowledge and experience of the author as well as the reader of the proof, and on the purpose of the proof.

Conviction is an important purpose of proof, and the process of finding a proof is popularly described as "convince yourself, convince a friend, convince an enemy" (Stacey et al., 1982, p. 95). The existence of a proof is often described as a prerequisite for a mathematician to be convinced of the truth of a hypothesis. But many mathematicians have testified to an almost opposite relation: "When you have satisfied yourself that the theorem is true, you start proving it" (Pólya, 1954). Neither is it true that a formal proof alone is what is required for the mathematics community to accept a new theorem as true. This acceptance is rather related to a combination of an understanding of the theorem's concepts, logical antecedents and implications, its significance in terms of implications for other branches of mathematics, its apparent consistency with other accepted results, the author's reputation as an expert in the field, and (of course) the existence of a convincing argument (Hanna, 1989). Thus, citing the verification of truth as the only function of proof, and proof as the main source for conviction, is misleading.

The mathematical meaning of proof carries three senses: *verification or justification*, *illumination,* and *systematization* (Bell, 1976). de Villiers (1990) adds another two, listing five functions of proof: *verification*, *explanation* (which corresponds to illumination), *systematization*, *discovery*, and *communication*. The distinction between proof as a means for conviction (or verification or justification) and as a means for explanation (or illumination) is that the former is only concerned with truth while the latter relates to understanding. Whether or not a proof explains does not affect its validity, but rather whether it is aesthetically pleasing. If it provides insight into why a statement holds, it explains the *cause* of the statement. Overly formal proofs, contrapositive proofs, and proof by contradiction have been criticized for not having this explanatory power, as have overly lengthy proofs and computer proofs. To fulfill its function of conviction, a proof need only give the logical reasons for a theorem; that is, guarantee *that* the theorem is true, not *why*. It verifies, while an explanatory proof clarifies (cf. Hanna, 1990). Of course, a proof can convince and explain at the same time. Regarding the other three purposes, systematization refers to the fact that proofs connect axioms, definitions, and statements and organize them into a deductive system. From a mathematical point of view this might well be their primarily function. Proofs also pave the wave for new discoveries by use of similar methods and premises and are, of course, a means for the communication of mathematical knowledge. Professional mathematicians likely agree on a sixth function of proof and proving: *intellectual challenge* (de Villiers, 1999).

2.2 Proofs in mathematics education

2.2.1 Definitions of proof

Numerous definitions and descriptions of proof are used by scholars in mathematics education. In addition, concepts like *proof, argument*, and *justification* are used as synonyms by some but with different meanings by others (Staples et al., 2016). The activity of *proving* (or *arguing* or *justifying*) can refer to the presentation of a proof, but also to the long process of finding, evaluating, and revising the proof and the statement it is meant to prove (e.g., Lakatos, 1976). In this broader meaning, *proving* can include activities such as conjecturing and making generalizations (Stylianides et al., 2017). Concepts like *reasoning-and-proving* (Stylianides, 2007) and *proof-related reasoning* (Thompson et al., 2012) are also used to emphasize the inclusion of learning activities with elements central to developing a proving competence, such as making and investigating conjectures, developing and evaluating arguments, etc. In this thesis, the word "proving" and the phrase "reasoning-and-proving" (with or without hyphens) should be interpreted in this broad sense.

In general, a proof is thought of as a specific form of argument, with no further specification of what an argument is. NCTM (2000, p. 55) describes proofs as "arguments consisting of logically rigorous deductions of conclusions from hypotheses", which aligns with the verification function of proof. So does the definition of justification by Staples et al. (2012, p. 448): "an argument that demonstrates (or refutes) the truth of a claim that uses accepted statements and mathematical forms of reasoning". The emphasis on "accepted" statements and forms of reasoning is further strengthened in the following definition, which also points to the sequential, step-by-step form of proof (Stylianides, 2007, p. 291):

> *Proof* is a *mathematical argument*, a connected sequence of assertions for or against a mathematical claim, with the following characteristics: 1. It uses statements accepted by the classroom community (*set of accepted statements*) that are true and available without further justification; 2. It employs forms of reasoning (*modes of argumentation*) that are valid and known to, or within the conceptual reach of, the classroom community; and 3. It is communicated with forms of expression (*modes of argument representation*) that are appropriate and known to, or within the conceptual reach of, the classroom community.

This definition makes proof a meaningful concept for educational contexts, independent of educational level or student age, and highlights the social and contextual aspect of proof. With the addition that a proof should also live up to today's standards of the community of mathematicians, the definition of proof can be expressed as two criteria (Stylianides et al., 2016, p. 317):

> *Criterion 1*: An argument qualifying as a proof should use true statements, valid modes of reasoning, and appropriate modes of representation, where the terms "true," "valid," and "appropriate" are meant to be understood with reference to what is typically agreed upon nowadays in the field of mathematics, in the context of specific mathematical theories.
>
> *Criterion 2*: An argument qualifying as a proof should use statements, modes of reasoning, and modes of representation that are accepted by, known to, or within the conceptual reach of students in a given classroom community.

It is reasonable to assume that most arguments presented to students on an upper secondary level satisfying Criterion 2 also satisfy Criterion 1. Hence, in such a context, "proof" refers to "what any mathematician and mathematics teacher would likely call a proof" (Thompson et al., 2012, p. 259).

The social component of proof is emphasized in the following description of argumentation, which also includes the element of convincing: "discursive exchange among participants for the purpose of convincing others through the use of certain modes of thought" (Wood, 1999, p. 172). This definition includes no requirement that the argument be correct. Neither do the definitions of proofs, proving, and proof schemes by Harel and Sowder (2007). In their terminology, "*proving* is the process employed by an individual (or a community) to remove doubts about the truth of an assertion"; that is, the proof is what turns a conjecture into a fact. The process of proving consists of two subprocesses: ascertaining (convincing yourself) and persuading (convincing someone else). A *proof scheme* consists of what constitutes ascertaining and persuading for a person (or a community).

Harel and Sowder's definition of a proof scheme is subjective, with an individual as well as a social dimension. It assumes a kind of common understanding and agreement within the mathematics community as to what constitutes a valid proof; that is, one can (in principle) say that professional mathematicians have the same proof scheme. This proof scheme is what determines the meaning of "true", "valid", and "appropriate" in Criterion 1. Education aims to socialize students into this way of thinking: "[The notion of proof] evolves and develops up to the upper-level undergraduate and graduate school levels, where it merges with the notion as understood by researchers and other professional mathematicians" (Hersh, 2009, p. 17).

While all the definitions above emphasize the conviction/verification/justification function of proof and proving, this does not mean that educational researchers do not acknowledge other functions of proof. On the contrary, for the mathematics classroom, it has been advocated that the most important function of proof is to provide insights and understanding (e.g., Hanna, 1995; Hersh, 1993).

2.2.2 Proof-like arguments

The concept of proof scheme shifts the focus from what features an argument should have in order to be called a proof to what features it should have in order to be convincing for individuals and communities. Based on empirical data on students' (and teachers') reasoning, Harel and Sowder (1998) identify three main categories of proof schemes. *External proof schemes* refer to modes of reasoning based on properties external to the mathematics at hand. This could be by reference to authorities (a teacher), or through symbolic manipulation without reflection/connection to the situation at hand. *Empirical proof schemes* refer to modes of reasoning based on specific cases or visual (or other) impressions. Finally, *deductive proof schemes* include the modes of reasoning employed by mathematicians. All three categories have several subcategories (Harel & Sowder, 1998).

Similar categories of reasoning have been described by other scholars. Harel and Sowder's empirical and deductive proof schemes align well with Bell's (1976) *empirical response* and *deductive response*, respectively. Balacheff (1988a) distinguishes between *pragmatic proofs* and *conceptual proofs*, which essentially include the same distinction. Balacheff identifies several subclasses. Within the class of pragmatic proofs, he places *generic examples*. These are examples in which, typically, all one needs to do to obtain a general proof is to replace a specific number with a variable. This means that "one can see the general proof through it" (Movshovitz-Hadar, 1988, p. 19), because nothing specific to the example enters the proof. Harel and Sowder (2007) have generic examples as a subclass of deductive proof schemes.

Stylianides (2008) suggests an analytical framework for reasoning-and-proving, distinguishing between "providing non-proof arguments" and "providing proof arguments." Generic examples are placed in the "proof argument" category and empirical arguments in the "non-proof argument" category. Together, they form a mathematical component that he refers to as "providing support to mathematical claims". So far, his categorization of arguments is similar to Harel and Sowder's empirical and deductive proof schemes. Stylianides (2008) also broadens the idea of proving to include another mathematical component: "making mathematical generalizations". This component includes "identifying a pattern" and "making a conjecture". Stylianides's framework also includes a psychological and a pedagogical component. The psychological component focuses on the student's perception of the mathematical nature of the mathematical components, while the pedagogical component focuses on how the student's perceptions compare with the (true) mathematical nature and how this nature can be made transparent to the student.

Independent of Stylianides (2008), Thompson et al. (2012) suggest a similar framing of proof-related reasoning, categorizing reasoning activities along two dimensions: type and nature of reasoning. Type of reasoning captures the difference between proof and non-proof arguments and between empirical and deductive proof schemes, while nature of reasoning includes conjecturing, providing support to mathematical claims, and other elements of proof-related activities such as investigating conjectures, correcting arguments, and finding counterexamples. This framework will be described in detail in Section 3.2.

The idea that conjecturing and proving are closely related has been put forward by others. The term *cognitive unity of a theorem* has been introduced to refer to when arguments used during the formulation of a conjecture are reused in a subsequent formal proving stage (Garuti et al., 1998). The iterative process of reformulations of a statement and the continued efforts to find proofs for it is described in Lakatos (1976).

2.2.3 Proof as a field of educational research

In their review of recent decades' research on the teaching and learning of proof, Stylianides et al. (2017) identify three broad research perspectives on proving: *proving as problem-solving*, *proving as convincing*, and *proving as a socially embedded activity*. Here follow short descriptions of these perspectives and an overview of how recent decades' research on proof is distributed over the three perspectives, over different stages of school education, and over different mathematics topics. Key findings within the three perspectives that are relevant to this thesis are summarized in the next subsection.

Proving as problem-solving refers to research with a cognitive psychological perspective that aims to understand the skills and competencies students have or need in proof-related activities. Within this perspective, the goal of instruction is for students to be able to write arguments that researchers would consider proofs. The focus tends to be on the process of proving: can students answer proof-related questions correctly, are they able to produce correct proofs, can they determine whether arguments are valid proofs, etc. Theoretical constructs used within this perspective often distinguish between different parts of the proving process such as the formal-rhetorical part and the problem-centered part (Selden & Selden, 2013), or between proving processes based on different kinds of representations such as syntactic and semantic proof production (Weber & Alcock, 2004, 2009). The concept of cognitive unity of a theorem (Garuti et al., 1998) is also a construct within this perspective.

Proving as convincing takes a constructivist perspective and focuses on what arguments students (and teachers) find convincing, and how these arguments relate to standards of argumentation in the mathematics discipline. Within this perspective, the goal of instruction is for students to be convinced by the same types of arguments that convince mathematicians. It is within this perspective that the theoretical constructs by Bell (1976), Balacheff (1988a), and Harel and Sowder (1998) are used. The frameworks by Stylianides (2008) and Thompson et al. (2012) capture aspects of proving as convincing as well as aspects of proving as problem-solving.

Proving as a socially embedded activity is a perspective with a focus on how proof is practiced among mathematicians and in mathematics classrooms. This perspective includes mathematicians' and students' reasons for engaging in proving, what purposes proof and proving have for them, and how the role of proof is negotiated in the classroom community.

In a review of empirical research on proof and argumentation in K–12 mathematics, Campbell et al. (2020) use the three perspectives described above as an analytical framework for categorizing recent research. Dividing Grades K–12 into four bands (K–2, 3–5, 6–8, and 9–12), they found that proof as problem-solving was the dominant research perspective in all bands but Grades 3–5. 71% of the research studies with participants in Grades 9–12 had this perspective while 21% were categorized as proof as convincing, and only 7% had a focus that could be conceptualized as proof as a socially embedded practice.

Campbell et al. (2020) also mapped which mathematics topics were used to elicit argumentation, and what kind of argumentation was required of the students. Half of the studies used geometry tasks. In Grades K–5 arithmetic tasks were the most common, but researchers relied on geometry tasks in 64% of studies with participants in Grades 9–12. At this level, algebra was also well-represented (24%). However, no publications utilized calculus tasks. Of the tasks used in the research with Grades 9–12 students, 93% required general arguments (and not just arguments about a single or multiple cases).

Research on the teaching and learning of proof is most frequently focused on the secondary and tertiary levels. For example, of the 76 journal articles that satisfied the search criteria in Campbell et al. (2020), only five included participants from Grades K–2 and 14 from Grades 4–6, while 34 studies had participants from Grades 6–8 and 42 from Grades 9–12.

2.2.4 Findings on students' difficulties and misconceptions

Research has shown very clearly that students have various difficulties and misunderstandings related to proofs, and that teachers often have the same difficulties as their students (e.g., Harel & Sowder, 2007; Stylianides et al., 2016; Stylianides & Stylianides, 2017; Stylianides et al., 2017). As difficulties are especially well-documented on the secondary and tertiary levels, they are common even at university level (e.g., Hemmi, 2008; Weber, 2001). Many students experience difficulty in the transition to university mathematics, especially in relation to proof (Alcock & Simpson, 2002; Clark & Lovric, 2009; Gueudet, 2008; Hemmi, 2008; Hillel, 2001; Hong et al., 2009; Leviatan, 2008; Luk, 2005; Moore, 1994; Oikkonen, 2008, 2009; Selden, 2005; Selden & Selden, 2003; Thomas et al., 2012; Thomas & Klymchuk, 2012) and particularly in calculus (Alcock & Simpson, 2002; Alcock & Weber, 2005). Students also experience that proof is something that has never been discussed and that important features of proof have been invisible (Hemmi, 2008).

In their review, Stylianides et al. (2017) summarize 14 key findings in research on the teaching and learning of proof (shown below). While the findings that are the most relevant to this thesis are elaborated on, the references are only examples. For further detail, see Stylianides et al. (2017) and Harel and Sowder (2007).

The first five key findings are within the proving as problem-solving perspective:

1. Students from lower secondary school to the university level often have difficulty writing proof.
2. Students often lack many of the competencies needed for writing proof.
3. The consideration of diagrams and examples is potentially beneficial for students who are writing proofs.
4. Students often have difficulty translating informal arguments into proofs.
5. Students and teachers are often unable to distinguish between proofs and invalid arguments.

Several large-scale studies have found poor success rates when students are asked to prove things, even if only a single deduction beyond the hypothesis is required (e.g., Senk, 1985, 1989). Among common difficulties are those involving choosing a legitimate proof framework (Selden & Selden, 1995), a lack of understanding of the proof methods they use (Stylianides et al., 2007), counterexamples in proof (Ko & Knuth, 2013), implications (Durand-Guerrier, 2003), and negations (Antonini & Mariotti, 2008). Students also have difficulty distinguishing between axioms, definitions, and theorems (Vinner, 1977), and in understanding statements with complex logical structures (Zandieh et al., 2014). They frequently misuse quantifiers. In multiply quantified statements such as "for all …, there is …", students are not aware that the existential variable can change when the universal variable changes (e.g., Durand-Guerrier & Arsac, 2005). Students typically have greater difficulty with statements beginning with the existence quantifier (i.e., "there is … such that for all …") than with statements involving the reverse order between the quantifiers (Dubinsky & Yiparaki, 2000).

Minor studies and case studies report that using diagrams and examples can be beneficial for students' proof writing (e.g., Alcock & Weber, 2010), but the findings are inconclusive. There are also examples of the opposite, when diagrams provide students with false confidence in conjectures or confuse them regarding what can be assumed and what needs to be proven (e.g., Alcock & Simpson, 2004). Students have trouble producing a proof if the structural distance between the warrants in an informal argument, and those in a formal proof, is too wide (Pedemonte, 2007).

Several small-scale studies also show that students have difficulty distinguishing proofs from invalid arguments (e.g., Selden & Selden, 2003). For instance, even mathematics majors often accept proof of the converse of a conditional statement as proof of the original statement (Inglis & Alcock, 2012).

As the examples above indicate, research findings tend to focus on what students cannot do. Less is known about what kind of (basic) proofs they actually can produce (Stylianides et al., 2017).

Within the proving as convincing perspective, four key findings are described:

6. Students and teachers are often convinced by empirical arguments as proofs of generalizations.
7. Students and teachers are often unconvinced of the power of proof to prove.
8. Students and teachers are often convinced that a conditional statement is equivalent to its converse or inverse, and unconvinced that it is equivalent to its contrapositive.
9. Students and teachers are often convinced by superficial features of proof by mathematical induction.

One of the most well-documented features of students' proving is that they provide empirical arguments when asked for proof of a general statement, or evaluate empirical arguments as general proof (Almeida, 2001; Healy & Hoyles, 2000; Knuth et al., 2002; Martin & Harel, 1989; Morris, 2002; Sevimli, 2018; Sowder & Harel, 2003; Thompson, 1991). Counterexamples are not seen as sufficient for refuting a theorem, but are instead viewed as exceptions (Balacheff, 1988b; Galbraith, 1981). If students accept an argument as a valid proof, they might still feel a need to check specific cases (Porteous, 1990). It thus seems that, for some students, examples are powerful enough to justify a statement but not to refute it. However, there are also studies that indicate that students are aware that verification with specific examples does not qualify as proof (e.g., Healy & Hoyles, 2000; Hemmi, 2006) even though they find this more convincing. It could also be that students produce specific arguments when they are unable to write down a general proof. This means that there might be two different difficulties/misconceptions involved. One is that they do not know how to make use of examples in supporting or refuting mathematical claims, while the other is that the convincing power (from the student's point of view) of examples and general arguments does not match what the mathematics community requires of a convincing argument.

Regarding conditional statements, $P \Rightarrow Q$, several studies show that students are convinced of their validity based on proofs of their converses, $Q \Rightarrow P$ (Hoyles & Küchemann, 2002), but might reject arguments that are correct proofs of their contrapositives, $\neg Q \Rightarrow \neg P$ (Stylianides et al., 2004).

Finally, Stylianides et al. (2017) present five key findings within the proving as socially embedded activity perspective:

10. Students' perceptions are largely shaped by regularities they observe in their classrooms.
11. The format in which proofs are written can constrain the types of reasoning that take place in mathematics classrooms.
12. Mathematicians usually do not read proofs to gain certainty in theorems but to advance their mathematical agenda.
13. Students and secondary mathematics teachers often do not see proofs as providing explanation, and have difficulty understanding proofs.
14. Negotiating productive classroom norms can highlight students' responsibilities with respect to proof and thereby create learning opportunities.

Students' judgement of what counts as a valid proof is influenced by how similar an argument is to what they have seen in class (Hemmi, 2006). For instance, students can reject an argument including graphical representations simply because they have not seen such proofs in class even if they find them convincing (Weber, 2010), or choose a less convincing but algebraic argument when asked to select the argument they believe would score highest on a test (Healy & Hoyles, 1998). Other findings are that students do not see the purpose or need for proof (Coe & Ruthven, 1994; Tinto, 1990), and do not see it as a tool for explanation or communication (Healy & Hoyles, 2000).

2.3 Proofs in mathematics textbooks

Curriculum programs (that is, textbooks, teacher guides, and other related materials supplemented by textbook authors and publishers) are widely used in classrooms all over the world (e.g., Grouws et al., 2000; Mullis et al., 2012; Pepin & Haggarty, 2001). Internationally, approximately 75% of all Grades 4 and 8 students are taught mathematics with the aid of a textbook (Mullis et al., 2012), which makes textbooks the most commonly used curriculum resource. Not surprisingly, textbooks are regarded as crucial links between national curricula, teaching practice, and student learning (e.g., Cai et al., 2011; Pepin et al., 2013; Stein et al., 2007; Valverde et al., 2002).

Paper VII includes a systematics review of research on curriculum programs in mathematics published during the period 2008–2014. Four areas are in focus: studies that (1) describe design principles of educative curriculum programs and teachers' response to them; (2) describe how teachers prepare for teaching; (3) relate influencing factors that explain transformations between written, intended, and enacted curriculum, and student learning; and (4) describe the effectiveness of curriculum programs. These areas were selected based on future research recommendations by Stein et al. (2007) in their extensive review of curriculum research. One important conclusion in Stein et al. (2007) is that no curriculum is self-enacting. The findings reported in Paper VII give further detail on this. For instance, teacher resources and curriculum resources can uniquely and jointly contribute to the quality of instruction (Hill & Charalambous, 2012), but to make productive changes to the curriculum, teachers' goals need to be aligned with the curriculum's goals (Davis et al., 2011). Students working with new, ambitious, standards-based curriculum programs do not fare worse on traditional tests, but perform better on tests aligned with the ambitious philosophy (Gavin et al., 2013). However, in order to positively impact student performance, the learning environment needs to be in line with the curriculum program's philosophy. Otherwise, the use of such a curriculum program can result in lower student achievement (Tarr et al., 2008). Irrespective of curriculum program, teachers likely need extra support from it in implementing its more challenging tasks (Choppin, 2011).

Research shows that proofs have a marginal place in many classrooms. Textbooks have been highlighted as one possible factor behind this (Stylianides et al., 2017). A few decades ago, Hanna and de Bruyn (1999) pointed out that textbook research with an explicit focus on proving was rare. Since then, there has been increased interest in conducting and publishing this kind of research. However, the field is still young, and even if the body of research is growing it remains limited. While the goal is to come up with well-founded prescriptions for textbook design, much research remains in the stage of describing the current state of the art for proving in textbooks. Typically, textbook studies are empirical and compare textbooks (or textbook series) from different publishers or countries (Stylianides et al., 2016).

A search in Web of Science's core collections shows that textbook studies of proof remain a small subfield in educational research[1]. The search resulted in 23 articles with titles including "textbook" and at least one word related to proving, such as "prove", "proof", "justify", "argue", "explain", etc. Through a screening of titles and abstracts, 15 of them were found to be journal articles reporting findings from textbook analyses of proving (in a broad sense) from an educational perspective. Papers I and IV in this thesis are two of them. There is of course research published elsewhere (in journals not indexed in Web of Science, conference proceedings, edited books, etc.). Still, these 13 journal articles offer an indication of the characteristics of this field of research. The next three paragraphs give a brief overview based on them.

All 13 articles are published after 2009, and include analyses of textbooks from all parts of the world. While the United States dominates (Bieda et al.,

[1] The search was done on March 12, 2021, using the following search string: TOPIC:(math*) *AND* TITLE:(textbook*) *AND* TITLE:(proof* OR prove* OR proving* OR reason* OR argu* OR justif* OR demonstrat* OR motiv* OR explain* OR explanation*). Document type was restricted to article.

2014; McCrory & Stylianides, 2014; Otten, Gilbertson, et al., 2014; Otten, Males, et al., 2014; Stylianides, 2009; Thompson et al., 2012), there are also studies from Australia (Stacey & Vincent, 2009), Chile (Ortiz & Pastells, 2017), Malawi (Mwadzaangati, 2019), China (Zhang & Qi, 2019), Japan (Fujita & Jones, 2014), Israel (Dolev & Even, 2015), and Denmark (Jankvist & Misfeldt, 2019). Although none of these 13 studies compare textbooks from different countries, a number of them analyze and compare several textbooks (or textbook series) from the same country. For instance, the study by Thompson et al. (2012) includes 20 United States textbooks (reform-oriented as well as traditional), Stacey and Vincent (2009) analyze nine Australian textbooks, and 33 Danish books were read for the article by Jankvist and Misfeldt (2019). Only one study combines the textbook analysis with a study of how the textbook tasks are enacted in a classroom context (Mwadzaangati, 2019).

The analyzed textbooks cover all levels of schooling. While lower secondary school textbooks dominate (Dolev & Even, 2015; Fujita & Jones, 2014; Otten, Gilbertson, et al., 2014; Otten, Males, et al., 2014; Stacey & Vincent, 2009; Stylianides, 2009; Zhang & Qi, 2019), there are also studies of primary school textbooks (Bieda et al., 2014; Ortiz & Pastells, 2017), upper secondary school textbooks (Jankvist & Misfeldt, 2019; Mwadzaangati, 2019; Thompson et al., 2012), and textbooks for the education of elementary school teachers (McCrory & Stylianides, 2014).

Some studies include the analysis of complete textbooks for a whole school year, in which case several mathematics subtopics are included, while others select and analyze material related to specific topics. When topics are selected, geometry is the most frequent choice (Fujita & Jones, 2014; Mwadzaangati, 2019; Otten, Gilbertson, et al., 2014; Otten, Males, et al., 2014), sometimes in combination with algebra or number theory (Dolev & Even, 2015; Stacey & Vincent, 2009). There is also one study focusing on probability (Ortiz & Pastells, 2017) and one on exponents, logarithms, and polynomials (Thompson et al., 2012). When several topics are analyzed, the findings are presented so that comparisons are possible. However, comparing topics is not the primary focus in any of the studies.

Next, we turn to some typical research findings (not restricted to the 13 articles discussed above) on proof and proving in textbooks. On the one hand, textbooks can have a significant influence on classroom practice and students' opportunities to learn proof. On the other, analyses of mathematics textbooks indicate that proof has a limited place in textbooks, that common misconceptions are not sufficiently addressed, and that the sequencing and distribution of proving tasks within and across grade levels are inadequate (e.g., Stylianides et al., 2017). There is also limited support from the teacher guides and textbooks used in mathematics teacher education.

Textbook analyses typically indicate that textbooks offer few opportunities to learn proof. For example, argumentation and reasoning activities were not present at all in Croatian mathematics textbooks for Grades 6–8 (Glasnovic

Gracin, 2018), except in chapters on triangle similarity. In Australia, only 17% of Grade 8 textbooks offered deductive explanations for how to divide fractions, while all textbooks had deductive explanations on the area of a trapezium (Stacey & Vincent, 2009). In United States upper secondary textbooks on algebra and precalculus, approximately half of the addressed mathematical properties were presented without justifications, and when justifications were provided only half of them were general proof (Thompson et al., 2012).The figures were slightly higher in geometry, with 75% of the properties justified and 35% of them with general proofs (Otten, Gilbertson, et al., 2014). Among textbook tasks on geometry, 25% were proof-related in a wide sense (Otten, Gilbertson, et al., 2014). In algebra and precalculus only 6% of the tasks were proof-related and only half of them included arguments about general cases, but there was a tendency that proof-related tasks, tasks requiring reasoning about a general case, and tasks involving developing and evaluating arguments increased as students progressed through the curriculum (Thompson et al., 2012). Individual textbooks score higher, and there are examples of reform-oriented United States middle-school textbooks in which 40% of the tasks include reasoning-and-proving (Stylianides, 2009). However, few studies report such high measures.

2.4 Swedish and Finnish mathematics education

Sweden and Finland are neighboring countries with a common history, many close collaborations, and similar school systems. Textbooks are extensively used in mathematics classrooms in both countries, with large amounts of class time spent on textbook tasks (Boesen et al., 2014; Joutsenlahti & Vainionpää, 2010). In TIMSS 2011, 97% of the participating Swedish Grade 8 students had teachers who considered textbooks a main source for their teaching (Skolverket, 2012). Other studies have shown that students in Grades 7–9 spend almost half their lesson time working with textbook tasks (Skolinspektionen, 2009), and textbooks are assumed to be equally important in upper secondary school.

Finland has gained international interest for its students' performance on international evaluations such as TIMSS and PISA. Such tests often place East Asian and East European countries well above the United States and Western Europe. Finland is an exception to this, however, performing well (especially on PISA) and clearly outperforming Sweden.

Several of the papers in this thesis are related to research projects comparing different aspects of mathematics education in Sweden and Finland. The findings from these projects indicate substantial differences involving classroom teaching (Hemmi & Ryve, 2015a), conceptualizations and discourses in and about teacher education (Hemmi & Ryve, 2015b; Knutsson et al., 2013; Ryve et al., 2011), and the use of curriculum materials (Hemmi, Koljonen, et al., 2013; Hemmi et al., 2019). There are also indications that the character of school mathematics differs with respect to the status of proof and proof-related items. In elementary and lower secondary school, such items tend to be emphasized more in the Finnish than the Swedish context, while it is the other way around in upper secondary school (Hemmi, Lepik, et al., 2013).

3 Analytical concepts and frameworks

This chapter describes the analytical frameworks used in the five papers of the thesis, and defines relevant concepts. Opportunity to learn is discussed in Section 3.1 with a focus on proof-related reasoning in mathematics textbooks. For the analysis of opportunities to learn proof, all papers, and especially Papers I and IV, make extensive use of two analytical constructs adopted from Thompson et al. (2012): type of reasoning and nature of reasoning. Section 3.2 is devoted to a detailed presentation of their adaptation for the studies in this thesis. Section 3.3 includes the refined frameworks, approaches, and tools that are used only in Papers II, III, and V, respectively.

3.1 Textbooks and opportunities to learn

The research questions in this thesis concern opportunities to learn proof. *Opportunity to learn* is widely accepted as one of the most important predictors of student achievement (e.g., Hiebert & Grouws, 2007). The United States' National Research Council defines opportunity to learn as "circumstances that allow students to engage in and spend time on academic tasks" (National Research Council, 2001, p. 333). This should not be confused with "being taught" or "exposed to". Opportunities to learn are the results of, and shaped by, students' pre-knowledge, learning goals, the design and enactment of learning activities, etc. A learning opportunity can only occur when an activity is aptly scaffolded and placed within the learner's proximal zone of development (Vygotsky, 1980) so that learning is possible. With this definition, it is impossible to investigate opportunities to learn proof solely by analyzing mathematics textbooks. The textbook neither defines the curriculum nor determines student learning; it is an artefact that mediates learning.

Curriculum can be conceptualized as existing in – and undergoing transformations between – a series of temporal phases: the written curriculum, the intended curriculum, the enacted curriculum, and finally, student learning (Stein et al., 2007). The written curriculum includes national steering documents, policies, syllabi, etc. that describe, prescribe, or give recommendations on subject content, learning goals, teaching organization, and so on. Textbooks and teacher guides are part of the written curriculum, even though they are also textbook authors' interpretations of (for instance) national steering documents.

The intended curriculum refers to the teacher's plans and intentions. The teacher interprets the written curriculum and adapts it for instruction. This process is colored by the teacher's knowledge (subject matter as well as pedagogical), perceptions, and beliefs about mathematics teaching and learning, etc. The next temporal phase, the enacted curriculum, refers to the teaching that takes place. Whatever the teacher's intentions are, and how well-planned a lesson is, there are many factors that can affect how the teaching activities play out in the classroom and what the student learns from them. It is the enacted curriculum and the interplay between students, teacher, and peers, that give rise to learning opportunities.

In light of the above, the textbook is just one factor that can influence the enacted curriculum and the opportunities to learn. The textbook is therefore seen as offering *potential* opportunities to learn. To be able to assess such opportunities, the analyses presented in this thesis are based on the assumption that what is written in the book is a good description of how the curriculum has been enacted. This means that the textbook is interpreted as a correct account of the content and chronology of classroom instruction and student activities.

Textbook material intended for use by the student can be divided into two categories: *expository sections* and *student tasks*. Expository sections refer to the parts where the textbook plays a lecturing role. Here, the textbook authors present new concepts and terminology, formulate and justify mathematical statements, describe methods and algorithms, make connections to applications, give historical notes, etc. Also, the authors can exemplify and demonstrate solution techniques in worked examples.

Expository material is usually mixed with student tasks. Typically, expository sections end with exercise sets; that is, lists of tasks or activities of varying kinds meant for the student to complete on his/her own or together with peers. Chapters and sections can also start with introductory tasks, and at the end of chapters (or at the end of the book) there can be additional sets of mixed problems meant for enrichment and rehearsal. Wherever they appear and whatever their character is, they will be referred to as tasks or exercises, even if textbook authors also use labels such as problems, activities, or questions.

Textbook analysists have argued that an analysis of both kinds of textbook data is necessary to obtain a coherent picture of opportunities to learn proof-related reasoning (e.g., Li, 2000; Otten, Gilbertson, et al., 2014; Thompson et al., 2012). Papers I, IV, and V use both kinds of data. Expository sections are considered to offer opportunities to learn proof-related reasoning when they include a presentation and discussion of justifications of mathematics results. Student tasks are considered to offer opportunities to learn proof-related reasoning when they ask the student to make or investigate conjectures, develop or evaluate arguments, identify or correct mistakes, provide counterexamples, etc. The next sections include the relevant definitions and detailed information about the analytical frameworks that have been used. Procedural details are described in Section 4.3.

3.2 Frameworks for proof-related reasoning

Opportunities to learn offered in expository sections and in student tasks differ in at least one respect: In justifications in expository sections the textbook authors do the reasoning, while in tasks the student is expected to do it. This influences the analytical approach to the two kinds of textbook data.

3.2.1 Type of reasoning in textbook justifications

In expository sections, the focus is on the opportunities students are offered to learn the role of proof in mathematics and what makes a justification a valid proof. For this purpose, mathematics results addressed in expository sections are classified according to whether they are justified, and whether or not the justifications are proofs. Other aspects of proof that have been investigated are discussed in the methodology, as is the identification of the addressed mathematics results.

A *mathematical statement* is either true or false. In this thesis, no distinction is made between true and provable statements. True statements are those for which proofs have been presented that the mathematics community has accepted as valid. In mathematics, true statements are called theorems, propositions, lemmas, corollaries, laws, principles, etc., depending on their importance, content, look, and relation to other statements. In this thesis, they are called *mathematics results* (except in Paper I, where they are referred to as mathematics statements). Thompson et al. (2012) call them *properties*, but as mathematical objects have certain properties by definition and others by consequence, properties is a less suitable term.

A *justification* provides reasons for why a statement might be true; it explicates grounds and warrants for a claim. This means that a justification is an *argument* in the sense of Toulmin (2003). Therefore, the terms justification and argument are used interchangeably. The focus on truth means a focus on the verification role regardless of whether the justification also fulfills other roles, e.g. explanation (de Villiers, 1990). A *textbook justification* is a justification provided by the textbook authors.

Textbook justifications are classified as justification by *general proof* or by *specific case*. This classification is referred to as *type of reasoning* (Thompson et al., 2012). A general proof is a justification that lives up to the mathematics community's standards of proof, but with the necessary adjustments to make it understandable or within reach for the intended students. For instance, a proof in an upper secondary textbook cannot be expected to have the same level of formalism as a proof in a journal article. This means that proof is interpreted in the sense of Stylianides (2007). At this level of schooling, a reasonable interpretation is that a proof is "what any mathematician and mathematics teacher would likely call a proof" (Thompson et al., 2012, p. 259).

Mathematics results are often universal, stating something for an infinite class of objects. A proof must contain reasons for why the statement holds for all of them; that is, it must be valid for an arbitrary object in the class. To emphasize this aspect of proof, they are called general proof. However, not all proofs involve reasoning about general cases. A statement about a specific case need only provide a proof for that specific case. An existence result can be general in its formulation, but a proof only needs to provide one specific example to ensure the existence. In this thesis, such proofs are also placed in the general proof category. This means that a general proof is not the same thing as an argument about a general case. In addition, an argument about a general case need not be general enough to classify as a general proof. For example, the derivation

$$\ln a \cdot b = \ln e^{\ln a} \cdot e^{\ln b} = \ln e^{\ln a + \ln b} = \ln a + \ln b$$

is valid for arbitrary positive a and b. Thus, it is a general proof of the statement $\ln a \cdot b = \ln a + \ln b$, but not for the statement $\log_c a \cdot b = \log_c a + \log_c b$.

A typical non-proof argument is a justification of a universal statement by use of a specific example. This kind of justification is what Harel and Sowder (2007) include in empirical proof schemes, and what Bell (1976) calls an empirical response and Balacheff (1988a) a pragmatic proof. Harel and Sowder (2007) and Stylianides (2008) make exceptions for generic cases, which they include in the deductive proof scheme and count as a proof argument, respectively. In this thesis, though, such cases are not viewed as general proofs. There are also other ways in which a justification can fail to be a proof; it may be too informal, be incomplete, have logical flaws, etc. All justifications that are not valid proofs – for whatever reason – are included in the specific case category of justifications.

Summing up, a mathematics result can be presented in a mathematics textbook with or without a justification. A justification is either a (general) proof or not, in which case it is classified as a justification by specific case. For excluded justifications, there is another alternative: the authors can ask the student to complete the proof. This gives a total of four alternatives, summarized in Table 1. The descriptions, from Paper IV, differ slightly from those in Thompson et al. (2012): "property" has been replaced with "statement", and the subordinate clause in the description of the S category has been added. The exact formulations also differ between Papers I, IV, and V, but the categories have always been used in accordance with the formulations in Table 1.

Table 1. *Framework for type of reasoning in expository sections*

Code	Type of justification	Description
G	General proof	The statement is justified with a proof.
S	Specific case or other non-proof justification	The statement is justified using a deductive argument based on a specific case, or that has other flaws that make it a non-proof justification.
L	Left to the student	A justification of the statement is left to the student to complete, typically with a problem in the exercises for which a justification of some type is required.
N	No justification	No justification is provided, and no explicit mention is made of leaving the justification to the student.

3.2.2 Type of reasoning in textbook tasks

A textbook task can offer students opportunities to engage in proof-related reasoning activities. In such cases, the type of reasoning must be inferred from the formulation of the task as there is no explicit justification to assess. Type of reasoning therefore refers to the reasoning required for a correct solution of the task. If it requires argumentation about a general case the task is classified as *general*, but if it suffices to reason about a specific case, it is classified as *specific*. The framework is summarized in Table 2. Should a task ask the student to verify the validity of a general statement by use of a specific case, this task will be considered to be of type S, as the student has been asked to reason about a specific case. It is worth noting that when Otten, Gilbertson, et al. (2014) adapted the same framework from Thompson et al. (2012), they introduced a separate category for this particular situation.

Table 2. *Framework for type of reasoning in textbook tasks*

Code	Type of reasoning	Task with proof-related reasoning about
S	Specific	A specific case/specific cases
G	General	A general case

The distinction between a specific and a general case can be problematic. On the one hand, the logarithm law above, $\ln a \cdot b = \ln a + \ln b$, is general in the sense that it is valid for all positive real numbers a and b. On the other hand, it is specific in the sense that it only expresses a property for one specific logarithm function. If a task asks a student to prove this formula, this task can be considered specific or general depending on what perspective one chooses. In Paper IV (and V), textbook material on the introduction of logarithms is analyzed. In these sections, logarithms are not studied as functions but as numbers. In such a context, the logarithm law above is seen as a general result, and hence the reasoning required to prove it is of type G. In Paper II, on the other hand, textbook material on primitive functions is analyzed. There, logarithms occur as functions. A task asking a student to prove that $\ln x$ is a primitive function to $1/x$ would therefore be classified as specific. The general principle at play here is that a task involving functions can only be classified as type G if there is some variable, parameter, or degree of freedom beyond the independent variable. For instance, if the student is asked to prove that $\ln ax$ is a primitive function to $1/x$ for all $a \neq 0$, then this would have been a task of type G. Issues like this are discussed more thoroughly in Subsection 3.3.1.

3.2.3 Nature of reasoning in textbook tasks

An essential part of the work of a mathematician is establishing new mathematical knowledge, and finding a formal proof can be the final step in this process. It also includes investigating specific cases, formulating and testing hypotheses and conjectures, evaluating and refining arguments, correcting errors, constructing counterexamples, outlining as well as filling in details in an argument, etc. All these activities have their own characteristics and are part of "proving" in a broad sense. Thompson et al. (2012) introduce a framework with seven such activities, referring to them as *natures of reasoning*: (1) make and (2) investigate a conjecture, (3) develop and (4) evaluate an argument, (5) correct or identify a mistake, (6) find a counterexample, and (7) outline an argument. Together, these constitute *proof-related reasoning* (abbreviated PR). With minor revisions, this framework has been adapted for the analysis of student tasks and worked examples in expository sections.

Table 3. *Framework for nature of reasoning in student tasks and worked examples*

Code	Nature of reasoning	Task in which student is asked to
M	Make a conjecture	Make a conjecture, formulate a true mathematical statement, or find the precise conditions for a certain statement to be true
I	Investigate a conjecture	Investigate whether a given conjecture or statement is true or false
D	Develop an argument	Justify or explain why a certain statement holds
E	Evaluate an argument	Evaluate whether a certain justification or solution is correct
C	Correct or identify a mistake	Find and/or correct an error in an argument or solution
X	Counterexample	Find a counterexample to of a false mathematical statement
P	Outline a proof	Outline an argument without the details of a full proof
O	Other	Use some other element of proof-related reasoning
N	Not proof-related	Do something else

Table 3 presents the framework as it is described in Paper IV. In the original framework the P category is called "Principles of proof", but otherwise the names are the same. The exact descriptions vary slightly between Papers I, IV, and V. The second half of the description of the M category was added during the analysis for Paper IV when such tasks were encountered in the textbooks. The O category was also introduced during the work with Paper IV, when proof-related tasks that did not fit in the other categories were found. In Paper V the X category was renamed "Example" to include tasks in which the student is asked to supply supportive examples. However, no such tasks were found in the analyzed textbook data.

The purpose of the different natures of reasoning is not to distinguish proof-related from not proof-related reasoning, but to identify and characterize essential activities connected to proving and to do this in a way that is useful in the analysis of opportunities to learn proof. The list of natures is not necessarily exhaustive. Nevertheless, tasks in which the student is asked to do other things, such as just calculate something or solve an equation, will be referred to as not proof-related.

Paper IV includes 18 examples with authentic textbook tasks on logarithms and combinatorics and how they have been classified according to type and nature of reasoning. Here follows a short description of some general distinctions between the different natures of reasoning.

The difference between M and I tasks is that in M tasks the student must formulate a mathematical statement, typically after studying specific cases, while in I tasks the statement is given. The difference between the I and D tasks is that in I tasks it is not given whether the statement is true or false, but in D tasks the statement is always true. Hence, in I tasks the student does not know whether to prove or disprove the statement. In E tasks, the student is faced with an argument. Here, it is not a matter of deciding whether a statement is true or false but instead whether or not the argument is a valid proof for the statement. E tasks also include tasks in which the student is asked to evaluate a solution, but not tasks in which the student is only asked to evaluate an answer. C tasks differ from E tasks in that it is given that there is a flaw in the argument/solution and the student is asked to find it.

In X tasks it is given that a statement is false, and that it should be refuted by finding a counterexample. An I task might well include a false conjecture that can be refuted by finding a counterexample, but in such cases this information is not included in the formulation of the task. Similarly, there could be a D task that can be solved by providing a counterexample, but a D task does not include explicit prompts to approach the task in such a way.

Finally, a P task is similar to D, but instead of asking for a complete and detailed proof, the student only has to provide a proof outline or a sketch of the main steps.

3.3 Extensions and refinements

Papers II and III elaborate on analytical difficulties encountered during the work on Paper I. Paper II focuses on the distinction between specific and general tasks and pilots a refined framework for this, presented in Subsection 3.3.1. Paper III focuses on the analysis of expository sections, and problems caused by using textbook justifications as a unit of analysis. The approach of Paper III is described in Subsection 3.3.2. Finally, Paper V combines results from Papers I and IV to obtain descriptions of topic-specific characteristics for proof-related reasoning. The tools used for this are presented in Subsection 3.3.3.

3.3.1 Levels of generality

As discussed at the end of Subsection 3.2.2, the distinction between specific and general tasks can be problematic, especially in tasks involving functions. Consider tasks that ask for proofs of the following statements:

1. $\int_1^4 \sqrt{x}dx = \frac{14}{3}$
2. $\int_1^a \frac{1}{x^2}dx$ never exceeds π
3. $F(x) = \frac{a^x}{\ln a}$ is a primitive to $f(x) = a^x$
4. $\int_{-a}^a f(x)dx = 2\int_0^a f(x)dx$ if $f(x)$ is even

The first identity is not universal, and its proof only requires a direct calculation and use of the second fundamental theorem of calculus (see Appendix). The calculation involves algebraic manipulations of expressions with one real variable x, but all identities are between numbers. In the context of calculus, this is a task involving reasoning about one specific function, $x \mapsto \sqrt{x}$, and there is no other "generality" involved.

A proof of the second statement is also restricted to reasoning about a specific function, $x \mapsto 1/x^2$. However, the statement is universal in the sense that it is valid for all $a > 0$. In this sense, the proof involves reasoning about a general case.

The third statement is also valid for all $a > 0$. Here, though, a proof will involve reasoning with an infinite, one-parameter family of functions. Hence, the generality represented by the parameter a affects this task differently.

The fourth statement includes two different kinds of generalities as it is true for all $a \in \mathbb{R}$ and all even (integrable) functions f. Hence, it combines the kinds of generalities of the second and third statements. In addition, the class of even functions is much “bigger” than the one-parameter class of functions in the third example. Even, integrable functions form an infinite dimensional subspace of all integrable functions. So, although “evenness” is easily expressed algebraically as $f(-x) = f(x)$, there is no simple closed formula for representing all even functions at once.

According to the general principle applied in Papers I and IV, the first task would be classified as specific and the other three as general. But only the third and fourth examples offer opportunities for reasoning about an infinite class of functions, and the functions involved in the fourth are more “arbitrary” than those in the third. With inspiration from how the concept of dimension is used in connection to vector spaces and manifolds, a three-level framework for generality is tested in Paper II: Reasoning about a finite number of specific functions (as in the second task) is called *non-general*. Reasoning about parameter families of functions (as in the third task), with a finite number of parameters, is called *finitely general*. Within this category, the number of parameters can be used to measure how “general” the reasoning is. Reasoning about “arbitrary” sets of functions (as in the fourth task), where the whole class cannot be represented by a finite, closed formula, is called *infinitely general*. As this classification only applies to reasoning about functions, the framework is called *the function generality framework*; this is summarized in Table 4. Note that (as in the second task above) a task that offers reasoning about a general case can be classified as non-general according to the function generality framework.

Table 4. *The function generality framework*

Generality level	Task that includes reasoning about
Non-general	Specific function(s)
Finitely general	Parameter families of functions
Infinitely general	Classes of functions that cannot be represented by a closed, finite formula

From an “opportunity to learn” perspective, the difference between non-general and finitely general tasks is that in the latter case the student is offered opportunities for learning to distinguish independent variables from other variables and parameters, and to handle parameters when manipulating function expressions. However, there are algebraic expressions available for manipulation in both cases. In the case of infinitely general tasks, the student has to find and use suitable representations of the relevant properties (like “evenness” in the example above). Thus, the three levels of generality correspond to different learning opportunities.

Similar ideas have been used by others, but for splitting the class of specific tasks into two subcategories. For instance, Stylianides and Ball (2008) distinguish between reasoning about a single case, a finite number of cases, and an infinite number of cases. Similarly, Bass (2011) discusses the distinction between three levels of cardinality of a set of objects about which a claim is made: one, finitely many, or infinitely many. In this thesis, however, tasks that only include reasoning about a finite number of specific cases will always be classified as specific.

3.3.2 Embeddings of justifications

When justifications are classified according to type of reasoning (Table 1), the formulation of the statement that is justified must be considered. The unit of analysis, therefore, is the statement and its justification(s). For the analysis of visibility aspects of proof, Papers I and IV also include an analysis of the labeling of addressed mathematics results and justifications, and whether the justification is placed before or after the statement. Textbooks can also include worked examples and student activities in the presentation of results and justifications, provide several different types of justifications, use a variety of representations, etc. They can be more or less formal in their expositions, make more or less clear connections to other parts of the mathematics content, and be more or less clear on structural and logical aspects of mathematics results and proofs. The list can be made longer. This widened perspective on how mathematics results are addressed and justified in textbooks is referred to in Paper III as the *embedding* of justifications in expository sections. Its details require a more thorough investigation than simply checking the labeling and placement of justifications and whether or not they are general. Moreover, an overly simplified quantitative analysis risks missing important information about opportunities for proof-related reasoning and can give misleading results.

To cast light on the limitations of approaches like those of Papers I and IV, and to get data for developing a more refined analysis, the following approach is tested in Paper III. Condensed descriptions of how textbooks address a specific mathematics result are established. Each description follows the chronology of the textbook and includes the *statement* of the addressed mathematics result, all provided *justifications* of the result, and *definitions* of relevant mathematics concepts. It also includes material placed immediately *before*, *after*, and *between* definitions, statement, and justifications. The description includes all the details needed to make an analysis of the type and nature of the reasoning (according to the descriptions in Section 3.2), describe analytical difficulties related to this classification, and capture other issues of relevance for drawing conclusions about opportunities to engage in proof-related reasoning. The descriptions are then used as a basis for conducting and presenting the *analysis* of the types of reasoning, the description of *analytical difficulties*, and the discussion of *other issues* of relevance. Finally, a short *summary* pointing out aspects of proof-related reasoning that could be better incorporated in the analytical framework is established. The italicized words are rubrics in the description, and in this sense are used as a kind of analytical framework.

This approach is subjective in the sense that it highlights analytical difficulties as experienced by the individual researcher. On the other hand, analytical difficulties are indications of a need to develop or refine the analytical frameworks and procedures, and to make clearer analytical distinctions. In addition, when results from analyses of different textbooks are compared and found to be incomplete in some respect, it is reasonable to use this as a starting point for a discussion on how to further improve the validity of the frameworks and the reliability of the analytical procedures.

3.3.3 Characterizing topics

To measure whether a textbook offers opportunities for learning that mathematical statements can be justified through logical arguments, and that these are an essential part of mathematics, the relative proportion of addressed mathematics results that are justified can be used. To measure whether a textbook offers opportunities to learn what a general proof is and what makes it differ from other types of justifications, the relative proportion of justifications that are also general proofs can be used. The findings reported in Papers I and IV show that a textbook section can score high on one, both, or neither of these measures. To illustrate this, and to identify patterns across textbook series, the two-dimensional diagram in *Figure 2* is used as an analytical tool in Paper V. The upper rightmost corner represents textbook material in which most addressed mathematics results are justified, and the justifications are general proofs. The lower leftmost corner represents the opposite: textbook material in which few addressed mathematics results are justified and the provided justifications are not general proofs.

A similar scheme is used for summarizing results concerning textbook tasks. In this case, the relative proportion of tasks that are proof-related is measured on the horizontal axis, and the relative proportion of proof-related tasks that include reasoning about a general case is measured on the vertical axis.

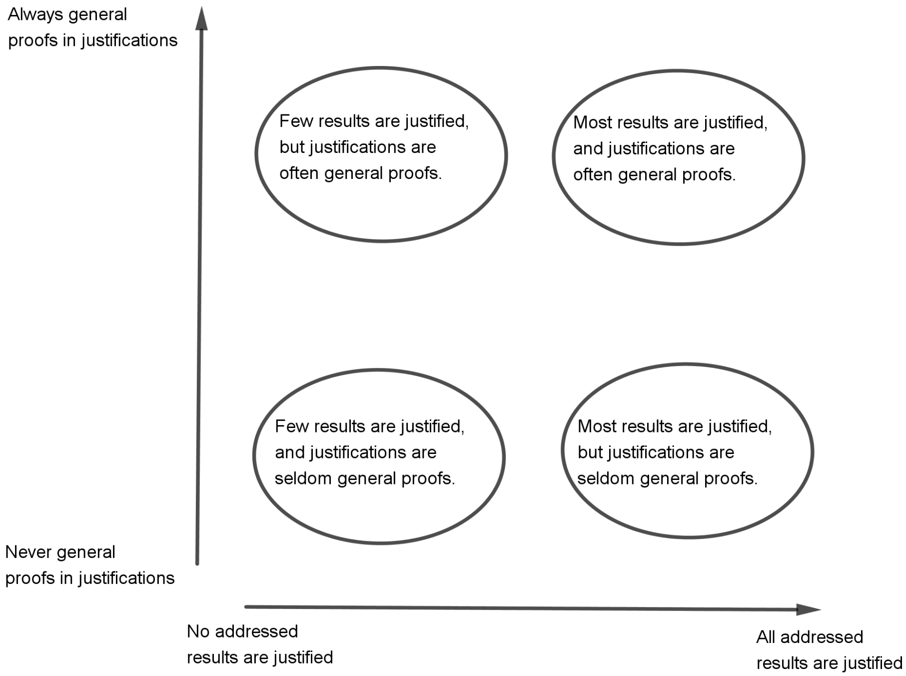

Figure 2. Justifications in expository sections.

4 Method

All the papers in this thesis include analyses of Swedish and Finnish upper secondary mathematics textbooks. This chapter focuses on the educational contexts of Swedish and Finnish upper secondary education, the selection of topics and textbook material, and the analytical procedures.

4.1 The Swedish and Finnish upper secondary school systems

Swedish and Finnish upper secondary schools are course-based, with different sets of courses for different programs. National steering documents prescribe content and learning outcomes in general terms. In this thesis, all descriptions of Sweden's upper secondary mathematics education are according to their national curriculum from 2011 (Skolverket, 2011). Since then, there have only been minor revisions to the mathematics curriculum. All descriptions of Finland's upper secondary mathematics education are according to their national curriculum from 2005 (Finnish National Board of Education, 2004). In Finland, some recent changes have affected the chronology between the mathematics content areas. However, these revisions had not been fully implemented, and revised textbooks were not available, when this thesis's textbook analyses were conducted.

4.1.1 Mathematics courses for science and technology students

As students preparing for higher studies in mathematics, science, and technology are the most likely to (and have the most need to) have opportunities for learning proof, the investigations in this thesis focus on textbooks for this group. In Sweden, upper secondary students studying science and technology are offered five mathematics courses and an optional sixth course with content the schools determine themselves. The first three courses exist in different versions for different programs, while the others are for science and technology students only. The courses are called 1c, 2c, 3c, 4, and 5. The typical prerequisites for higher education in science and technology are Courses 1c–4. In Finland, students preparing for higher studies in science and technology follow the advanced syllabus in mathematics, also called the "long course", which includes 13 parts. Parts 1–10 are mandatory, while 11–13 are specialization courses. For easier reference, the Finnish courses are referred to as parts.

4.1.2 Course content

All Swedish courses cover a range of different mathematics topics, while the Finnish parts focus on one or a couple of closely related topics. The parts are named after their content. This thesis focuses on logarithms, primitive functions, definite integrals, and combinatorics (see Subsection 4.2.1). In the Swedish setting, logarithms are introduced in Course 2c, which also treats topics such as linear equations, polynomials, exponential functions, geometry, and statistical methods. Logarithm functions and their derivatives are not treated until Course 4. In Finland, logarithms are covered in Part 8, which is titled Radical and Logarithm Functions. This part also includes general concepts such as composite and inverse functions, power and exponential functions, and derivatives of power, root, exponential, and logarithm functions.

Differential and integral calculus is introduced in Course 3c in Sweden; here, students encounter primitive functions and definite integrals for the first time. These topics are further developed in Course 4. Course 3c also includes algebraic equations and trigonometry. Trigonometry is also part of Course 4, which also treats complex numbers. In Finland, Part 10 is called Integral Calculus and is devoted to this topic only. The non-mandatory Part 13 is a continuation course in differential and integral calculus. Derivatives have their own course: Part 7, The Derivative.

Swedish Course 5 focuses on differential equations and discrete mathematics, which includes combinatorics. In the Finnish setting, combinatorics is treated in connection to probability theory in Part 6, Probability and Statistics.

4.1.3 Objectives related to reasoning and proof

On a general level, the two countries have similar objectives related to reasoning and proving. In the Swedish curriculum, the general objectives express that students should be given opportunities to develop their abilities to follow, conduct, and assess mathematical reasoning. Meanwhile, the overall objects of instruction in the Finnish steering documents are slightly more detailed. They include that students should "learn to appreciate precision of presentation and clarity of argumentation", "learn to perceive mathematical knowledge as a logical system", and "become accustomed to making assumptions, examining their validity, justifying their reasoning and assessing the validity of their arguments and the generalizability of the results" (Finnish National Board of Education, 2004). However, proofs as core content are emphasized more in the regulations of the Swedish courses.

In the Swedish setting, proof is explicit content in Courses 1c, 3c, 4, and 5 (see Skolverket, 2011):

- Course 1c includes “mathematical argumentation with elementary logic, including implication and equivalence” and “illustration of the concepts of definition, theorem and proof, for instance with the Pythagorean theorem”.
- Course 3c includes “proofs and application of the cosine, sine, and area theorems for arbitrary triangles”.
- Course 4 includes “application and proof of de Moivre’s formula” and “different proving methods in mathematics with examples from arithmetic, algebra, or geometry”, as well as “derivation and application of differentiation rules for trigonometric, logarithmic, exponential, and composite functions, and for products and quotients of functions”.
- Course 5 includes “proof by induction with concrete examples from, for instance, number theory”.

In Finland, the only instance in which proof is explicitly mentioned as core content is in Part 11, Number Theory and Logic. This course includes “formalization of statements; truth values of statements; open statements; quantifiers; direct, contrapositive and indirect proofs”.

It is worth noting that there are no explicit writings about proof in connection to logarithms, primitive functions, definite integrals, or combinatorics in either Sweden or Finland.

A detailed analysis of proof in Swedish and Finnish (and Estonian) steering documents for elementary and secondary school is found in Hemmi, Lepik, et al. (2013). One observation the authors make is that in Finland, the compulsory school curriculum discusses a number of proof-related topics, whereas (as described above) the upper secondary school curriculum mentions proof and proving only in an optional course in number theory and logic. The Swedish steering documents reflect the opposite, suggesting an attempt to elevate the status of proof in upper secondary schooling.

4.2 Data sample

The choice of textbook material for the analyses in the papers of this thesis is the result of three major considerations: mathematics topics, textbook series, and the selection of material within these topics and textbook series. This means that the data sample has been chosen to represent different countries, different textbook series (or publishers/authors), and different mathematics topics, but also different stages of upper secondary schooling.

4.2.1 Mathematics topics

At secondary level, explicit attention to proof has often been confined to courses in Euclidean geometry. When researchers choose topics for empirical research on the teaching and learning of proof, geometry is also the most frequent choice (Campbell et al., 2020). Consequently, other topics are less investigated. As there is broad consensus in the field that reasoning and proving should be emphasized in all areas of school mathematics (e.g. Stylianides et al., 2017), there is good reason to choose topics other than geometry for empirical research on reasoning and proof. This argument was used by Thompson et al. (2012), and also explains why the investigations in this thesis have focused on other topics.

The first mathematics topics chosen for this thesis were primitive functions and definite integrals. Together, they will be referred to as integral calculus. There were at least four different reasons for this choice of topics. First, calculus (with its two main branches, differential and integral calculus) has a strong position in the upper secondary mathematics curricula of many European countries. In the early twentieth century it was typically incorporated in school curricula to bridge the gap between school and university mathematics. Except during the "new math" era the approach has been informal, with a focus on applications and procedural skills (Törner et al., 2014). Second, calculus is a major part of introductory university courses in mathematics, and many university students have difficulty with this topic. Experts have raised concerns about the increasing gap between secondary and tertiary calculus teaching when it comes to formalism and focus on proof (e.g. Job & Schneider, 2014; Moreno-Armella, 2014; Roh, 2008). Third, a formal presentation of the theory of definite integrals requires the $\epsilon - \delta$ formalism of mathematical analysis, which is far beyond the scope of most upper secondary curricula. Thus, it is interesting to see how textbook authors have dealt with theoretical issues on upper secondary level. Finally, textbook studies focusing on proof rarely concentrate on calculus, especially not integral calculus; nor does empirical research. In the review by Campbell et al. (2020), no publications utilized calculus tasks.

Papers I–III are based on textbook material on integral calculus only, and for Papers IV and V, material on logarithms and combinatorics were added. The topics were chosen to broaden the scope, and to provide data and results on as diverse a set of topics as possible. Logarithms and combinatorics complement primitive functions and definite integrals in many ways. While integrals are a calculus-oriented topic, logarithms and combinatorics are algebra-oriented (even though logarithmic functions are studied in calculus as well). Students typically encounter logarithms earlier on in upper secondary schooling than they do integral calculus, which means that logarithms represent an earlier stage of upper secondary school. The same cannot be said about combinatorics. In Sweden combinatorics is studied after integral calculus, in Finland before. On the other hand, this offers an opportunity to study differences in how the topic is treated depending on its chronological place in the curriculum. Compared to definite integrals, logarithms and combinatorics are also theoretically less advanced topics, even though logarithms are known to be difficult for many students.

There are also reasons for choosing logarithms and combinatorics that are specific to one or the other. One of the most important reasons for choosing logarithms is that it is one of the topics investigated in Thompson et al. (2012). Hence, including logarithms in the investigation offers better possibilities to compare results involving Swedish and Finnish textbooks with similar results for United States textbooks. The other topics in Thompson et al. (2012) are exponents and polynomials. Compared to exponents, logarithms are more problematic for students. Also, there is a certain similarity between how logarithms relate to exponents, and primitive functions to derivatives, through an inverse operation: $b = \ln a$ if $e^b = a$, $F(x) = \int f(x)dx$ if $F' = f$. This structural similarity might lead to similarities regarding opportunities for proof-related reasoning worth studying. Finally, regarding combinatorics, this topic is perhaps the least theoretically advanced of all four. It requires no pre-knowledge except about the addition and multiplication of numbers. The theory is built on two simple principles only: the addition principle and the multiplication principle. It is also a topic within discrete mathematics and the most likely of the four to include proof by induction. Finally, combinatorics is more easily related to everyday situations than logarithms and integrals are.

As shown in the descriptions above, topics have not been chosen based on an idea about where opportunities for learning proof are the most common, or the most rare; rather, it has been a question of diversity and variation. The four chosen topics differ in several ways, and therefore provide a variety of cases.

4.2.2 Textbook series

Swedish and Finnish textbook publishers usually provide one textbook per course. In the Swedish case this means five textbooks to cover Courses 1c, 2c, 3c, 4, and 5, and in the Finnish case up to 13 textbooks to cover the mandatory Parts 1–10 and the supplementary Parts 11–13. The textbooks are produced on a free market without state control or certification. Schools make their own decisions regarding which (if any) textbooks to use, and teachers often have a say in this. Students do not have to pay for their books but can borrow them from their school, unless the school provides them for free.

In Sweden, the textbook series Matematik 5000 (and its predecessors Matematik 1000, 2000, 3000, and 4000) has had a dominating position on the market for several decades. In a 2011 survey of enrollees in engineering programs at Örebro University, more than 80% said they had used textbooks from these series during their years at upper secondary school. Although publishers are unwilling to reveal their market shares, based on informal contacts with teachers the impression is that Matematik 5000 continues to be the most popular textbook series on science and technology programs. Its main competitor is Matematik Origo; this textbook does not have the long history of Matematik 5000, but was released after the national curriculum revision in 2011. If Matematik 5000 represents a traditional (by Swedish measures) approach to mathematics, Matematik Origo can be said to be more reform-oriented. In its foreword, its authors express an intention to highlight problem-solving, understanding, and communication.

In Finland, one of the most popular textbook series for the long course is Pyramidi. The publisher offers this material in a Swedish translation, titled Ellips, for the country's Swedish-speaking minority (approximately 5% of the population). In informal contacts with the publisher, they have confirmed that Pyramidi and another Finnish textbook series, Pitkä matematiikka, are the most common textbook series for students following the advanced syllabus.

In view of the above, the textbook series Matematik 5000, Matematik Origo, Pyramidi/Ellips, and Pitkä Matematiikka comprise a representative selection of textbooks from the Swedish and Finnish educational contexts. They reflect the curriculum material a vast majority of Swedish and Finnish upper secondary students preparing for higher studies in mathematics have used in recent years. Therefore, these materials were a natural choice for the investigation in Paper I (which was the starting point for the research presented in this thesis). As the author of this thesis neither speaks nor reads Finnish but is a native Swede, the Swedish version Ellips was chosen instead of its Finnish original, Pyramidi. The second author of Paper I, who is native Finnish and fluent in Swedish and Finnish, conducted the analysis of Pitkä Matematiikka for Paper I. For the other papers of this thesis, this textbook series was excluded and all research was based on textbook material from Matematik 5000, Matematik Origo, and Ellips. This is of course a drawback when it comes to the possibility to generalize results to the Finnish setting. However, the original Finnish version of Ellips, Pyramidi, is one of the most used Finnish textbooks. Focusing on Swedish-speaking students in Sweden and Finland, Matematik 5000, Matematik Origo, and Ellips represent what most students have used in upper secondary courses in mathematics in recent decades.

Throughout the thesis the four respective textbook series are referred to using the codes Sw1, Sw2, Fi1, and Fi2, while different codes are used in the various papers (see Table 5). Bibliographic details are found in the reference list.

For Paper II, student exercises from a Swedish university compendium in single variable calculus were included (Matematikcentrum, 2010). This book is referred to as SwU. The corresponding textbook (Persson & Böiers, 2010) has been used for several decades in many Swedish universities' introductory courses in mathematical analysis. When new authors revised the text (Månsson & Nordbeck, 2011), most of the student exercises remained the same (Matematikcentrum, 2011). University material was included to determine whether the analytical framework discussed in Paper II could also be useful for tasks on this level.

Table 5. *Analyzed textbook series*

Code	Textbook series	Publisher
Sw1/SW1/S1/M5000	Matematik 5000	Natur & Kultur
Sw2/SW2/S2/Origo	Matematik Origo	Sanoma utbildning
Fi1/FI1/F1/Ellips	Ellips	Schildts Förlags Ab
Fi2/FI2/F2	Pitkä matematiikka	WSOY Oppimateriaalit Oy

4.2.3 Textbook data

All the analyzed upper secondary textbook series provide one book per course. As the four topics (logarithms, primitive functions, definite integrals, and combinatorics) are treated in different courses, all the corresponding textbooks from the chosen series have been analyzed. The only exception is Fi2, which has only been included in the analysis of integral calculus, and only in the work with Paper I. The included textbooks from the four series and their connection to the four analyzed topics are summarized in Table 6. As there is material on primitive functions and definite integrals in two books each from Sw1 and Sw2, it is sometimes necessary to be clear about which textbook a certain material belongs to. Otherwise, references will only be to textbook series.

Table 6. *Analyzed textbooks*

Series	Topic	Book
Sw1 (M5000)	Logarithms	2c
	Primitive functions	3c & 4
	Definite integrals	3c & 4
	Combinatorics	5
Sw2 (Origo)	Logarithms	2c
	Primitive functions	3c & 4
	Definite integrals	3c & 4
	Combinatorics	5
Fi1 (Ellips)	Logarithms	8
	Primitive functions	10
	Definite integrals	10
	Combinatorics	6
Fi2 (Pitkä)	Primitive functions	10
	Definite integrals	10

All analyzed textbooks have a similar structure. They are divided into chapters, sections, and sometimes subsections, and within each section (or subsection) are expository sections and exercise sets. Expository sections are typically followed by student exercise sets. Sometimes, chapters or sections start with a few introductory tasks (this was more common in the Swedish textbooks than in the Finnish ones). At the end of chapters, or at the end of the book, there are often additional sets of mixed problems meant for enrichment and rehearsal. Both expository sections and student exercises have been analyzed, as both kinds of textbook data are necessary to obtain a coherent picture of opportunities to learn proof-related reasoning (e.g., Li, 2000; Otten, Gilbertson, et al., 2014; Thompson et al., 2012).

As most analyzed textbooks also cover other topics (besides logarithms, primitive functions, definite integrals, and combinatorics), relevant chapters, sections, and subsections had to be singled out. This was achieved by reading the content pages and skimming through the complete textbook material. Such sections, and their corresponding exercise sets, have usually been included in their entirety. Related complimentary tasks placed within sets of mixed problems at the end of chapters (or at the end of the book) have also been included in the analysis. In general, the stance has been to include rather than exclude material. However, some choices had to be made regarding what should count as “related” to the four topics. Such considerations are summarized below.

In relation to logarithms, sections on powers and exponents that precede the introduction of logarithms have not been included, even if they are within the same chapters as logarithms. Sections on the derivative of the logarithm functions have not been included either, as it is only in the Finnish context that derivatives are introduced prior to logarithms. In relation to definite integrals, scientific applications with integral formulas for mass, energy, etc., have not been included, and neither have formulas for continuous probability distributions. There are two reasons for this. First, only the Swedish materials feature subsections on such material in their integral chapters (in the Finnish context, continuous probability distributions are treated in another, non-mandatory, continuation course). Second, the Swedish materials are unclear as to whether such integral formulas define the physical/probabilistic concepts or whether they are derivable properties. Finally, the Finnish textbooks provide some complementary material in short theory sections at the end of the books. This material was excluded because it was not part of the main text.

Within the selected expository sections, all main mathematics results have been included in the analysis. A main mathematics result is one that the textbook authors highlight, e.g., with a colored background or a frame, or by labelling it as a theorem, principle, rule, etc. Statements in the inline text are considered main results if they are accompanied by worked examples illustrating their use. In short, main mathematics results are the addressed theorems.

The choice to include all main mathematics results in the analysis differs slightly from the method of Thompson et al. (2012). They first assembled a list of mathematics results (within the topics they studied) that were likely to be addressed in the textbooks. Then, they looked for these results in the textbooks and analyzed how they were addressed and justified. The inclusion of all addressed main results is motivated by a wish to eliminate the risk of bias in the conceptions regarding what ought to be treated (Clarke, 2013). In addition, the aim of this thesis is not to study textbook coverage of certain content but opportunities for proof-related reasoning vis-a-vis the content covered.

In addition to main mathematics results, all worked examples in expository sections have been analyzed, as have all student exercises. The total numbers of analyzed mathematics results, worked examples, and student exercises are summarized in Table 7.

Table 7. *Size of data sample*

Textbook series	Mathematics results	Worked examples	Student tasks
Matematik 5000	29	78	760
Matematik Origo	27	83	876
Ellips	42	145	747
Pitkä	28	57	572

In Papers I and IV, the analysis was restricted to integral calculus. Paper II focused on proving tasks only; that is, tasks explicitly asking the student to "prove" or "show" something. Paper III focused on expository sections and included the complete definition-theorem-proof chain of the representation formula $F(x) + C$ for the primitive functions to $F'(x)$. The analysis also incorporated material presented immediately before, between, and after the definition, theorem, and proof.

The analyses have taken the perspective of the student in the sense that they are based on what students encounter in their textbooks. Teacher guides can explain textbook authors' intentions and include recommendations for classroom enactment of textbook material. Such information can be valuable in judging students' opportunities to learn from the textbook material. However, the role of teacher guides depends on how, and to what extent, they are used. While elementary school teachers are known to use teacher guides, there is no such widespread tradition among Swedish and Finnish upper secondary mathematics teachers. Publishers do not always supply teacher guides. For instance, there are no teacher guides for Fi1. When the Swedish series Sw1 and Sw2 were published in 2011 in connection to the implementation of a new national curriculum, these series only included student textbooks. Teacher guides were published several years later, and are now only available as digital resources. For these reasons, teacher guides have not been included in the analyses.

4.2.4 Remarks on generalizability

As is evident from the descriptions above, neither the topics nor the textbook series or the analyzed sections within them were chosen randomly. The topics were chosen to provide a diverse set of interesting cases, and the textbook series were chosen to represent what most science and technology students use. When all relevant textbook sections had been identified, there was no further selection involved – all addressed main mathematics results, all worked examples, and all student exercises were analyzed. This has important implications when it comes to the generalizability of the findings. The statistical significance of the findings in relation to all topics and all available textbook series cannot be measured using statistical tests. Such generalizations can only be guesses based on assumptions about the nature of the topics, textbook authors' didactical preferences, or contemporary norms for textbook design. On the other hand, all material on the chosen topics from the chosen series was analyzed. Thus, the findings can be used to draw conclusions about what most Swedish and Finnish science and technology students encounter in their textbooks when they study logarithms, primitive functions, definite integrals, and combinatorics at upper secondary level.

4.3 Analytic procedures and considerations

Most of the textbook analyses underpinning this thesis were conducted during the work with Papers I and IV. Below is a detailed account of the more practical aspects of these analyses, followed by subsections focused on the differing and unique aspects of the analyses in the other papers. Common to all papers is that the textbook material has been studied from a student perspective. This is elaborated on in the last subsection.

4.3.1 Papers I and IV

On a global level, the textbook analysis adhered to an iterative procedure. The first topic to be analyzed was integral calculus (i.e., primitive functions and definite integrals). First, a preliminary analysis of all materials on integral calculus in Sw1 and Sw2 was conducted by the first author of Paper I. All instances of analytical difficulties were then discussed with the second author of Paper I, and more detailed principles for the classification of tasks and justifications were determined. Then, Sw1 and Sw2 were re-analyzed. At this stage, textbook material from Fi1 was also analyzed. This work was paralleled by the second author, who analyzed Fi2. The second author translated expository sections in Fi2 to Swedish to enable the double-checking of coding. In addition, all the tasks in Fi2 evaluated as proof-related or ambiguous (i.e., it was unclear whether they should be viewed as proof-related) were translated and discussed. Through this procedure, we agreed on the coding and ensured that our analyses were based on the same principles.

For Paper IV, textbook material on logarithms and combinatorics from Sw1, Sw2, and Fi1 was analyzed. This analysis was guided by the principles agreed upon during the work with Paper I. To further ensure validity and reliability, the analysis of the tasks proceeded as follows. First, the analysis was restricted to "ordinary" tasks; i.e., introductory tasks and special activities at the end of the book sections were skipped. This was to determine whether and how the coding principles needed any adaptation for use on the new topics. A sample of tasks representing different natures and types of reasoning, as well as tasks found to be difficult to classify, were chosen and discussed with colleagues at a seminar. Based on this discussion, final analytical principles were settled upon, and were then applied during a final coding iteration of the complete data set to logarithms and combinatorics.

In the analysis of expository sections, the main unit of analysis has been a main result, along with its justification(s) (if such can be found near the statement itself rather than in a separate theory section at the end of the book). The main results' justifications were then classified in terms of type of reasoning, as shown in Table 1 in Subsection 3.2.1. Notes were also taken on the labelling of the main results and their justifications, generality, and logical structure in the formulation of the main results, proof techniques in justifications, and whether the justifications were placed before or after the statements. Such details were collected to be able to discuss opportunities to learn proof in terms of transparency and visibility.

In the Swedish context, integral calculus is introduced in Course 3c and continued in Course 4. This has the consequence that some mathematics results are addressed twice in the same textbook series. However, when such recurrence takes place in Book 4, the mathematics results are always expressed as reminders without justifications or expressed with exactly the same justifications as in Book 3c. Therefore, these statements were not coded twice but only according to their justifications when they first occur in Book 3c. This means that within each textbook series, no mathematics result appears more than once in the data.

There are a few occasions when textbooks offer more than one type of justification for a mathematics result. For instance, a specific case is discussed prior to the formulation of a general mathematics result, after which a general proof is presented. In Paper I double coding was avoided, and such mathematics results were coded as justified with a general proof. In Paper IV, however, such mathematics results received double codes. When findings from Papers I and IV were compared, this difference was taken into consideration.

In addition to main mathematics results and their justifications, all worked examples in expository sections have been analyzed. This analysis has followed the same principles and procedures as the analysis of student exercises.

In the analysis of exercise sets, the unit of analysis has been the complete textbook formulation of an exercise with its own identifier. If an exercise has been divided into parts, typically denoted (a), (b), (c), and so on, each part has been counted as a separate task. All tasks have been classified according to type and nature of reasoning, as described in Subsections 3.2.2 and 3.2.3. In principle, an exercise can receive several codes if it asks the student to accomplish several things. However, in no case did any exercise receive double codes for type of reasoning. In a few cases, double codes for nature of reasoning were assigned. Details on double coding can be found in the separate papers.

The two types of reasoning can occur in combination with all natures of reasoning. However, one can assume that a typical task in the counterexample category consists of refuting a general statement with a specific example. Thompson et al. (2012) chose not to designate a type of reasoning for such tasks (nor tasks in the outline a proof category). In Papers I and IV, it was decided that tasks asking for counterexamples would be counted as general (as it must be a general statement that they refute). One can argue that, if the student is to provide a specific example, it would be more stringent to view such tasks as specific. However, as no tasks asking for counterexamples were found in the analyzed textbooks, this decision was of no consequence to the findings.

4.3.2 Papers II, III, and V

In Paper II, the student exercises in integral calculus analyzed for Paper I were further analyzed. However, no exercises from Fi2 were included. Instead, exercises from a university text, SwU, were included. Proving tasks (exercises explicitly asking the student to "prove" or "show" something) were sorted out. All such tasks in Sw1, Sw2, and Fi1 had already been categorized as proof-related and classified as specific or general (according to the principles for type of reasoning described in Table 2 in Subsection 3.2.2). Proving tasks in SwU were identified and coded accordingly. In the next step, all general proving tasks were classified according to the function generality framework described in Table 4 in Subsection 3.3.1. This required a detailed collection of information about the function classes involved in the tasks.

Some extra distinctions, or analytical clarifications, had to be made during the analytical process. First, there were cases in which proving tasks expressed relations between two classes of functions. In such cases, the classification was based on the most general (in the sense of the function generality framework) of the two classes. Second, the classification was based on mathematical characteristics that the student could be expected to know about and make use of. For instance, a task involving parabolas (formulated in general terms without explicit algebraic expressions) was categorized as finitely general, as parabolas are graphs of second-degree polynomials and this fact could be used to solve the problem. In another example, the student was asked to prove *Rolle's theorem*: if $f'(x) = 0$ everywhere in an interval, then f is constant there. Here it is part of the conclusion that the class of functions under consideration is a one-parameter family. Obviously, this cannot be elicited in the proof; hence, the task was classified as infinitely general.

Paper III takes as its starting point the analytical difficulties encountered during the work with Paper I. Condensed descriptions of how Sw1, Sw2, and Fi1 introduce and define primitive functions, formulate the representation formula $F(x) + C$, and justify this formula, were prepared using the rubrics and structure described in Subsection 3.3.2. The classification of type of reasoning, and the possibilities to draw conclusions about opportunities to learn proof-related reasoning based on such a classification, were then problematized in relation to these accounts.

Paper V combines the results of Papers I and IV and includes a topic-by-topic analysis of the findings in order to characterize and contrast the four studied topics. An important tool is the diagram in *Figure 2* in Subsection 3.3.3. These characteristics are then discussed from the point of view of purely mathematical details of the mathematics content. Aside from this, the analysis includes no new steps or considerations beyond what has been accounted for in the section above. The discussion, however, requires some fundamental knowledge of the mathematics involved. A selection of important details is summarized in the Appendix.

4.3.3 Textbook reader perspective

All textbook analyses have been conducted from the perspective of a student reading the book and doing its exercises. A textbook reader can be conceptualized in different ways. Weinberg and Wiesner (2011) distinguish between the empirical reader, the implied reader, and the intended reader. The empirical reader refers to the actual reader of the text. This perspective has not been utilized in this thesis; students' ways of reading, comprehending, and responding to textbook material have not been investigated. The implied reader refers to the competencies the empirical reader needs in order to correctly interpret the text. In principle, the implied reader can be inferred from an analysis of the text itself: "the mathematical community has well-established conventions for writing, and mathematicians share a common understanding of the concepts that textbooks describe; consequently, the different attributes identified by expert readers all contribute to a coherent description of a textbook's implied reader" (Weinberg & Wiesner, 2011, p. 52). Finally, the intended reader is the idea of the reader formed in the mind of the author.

The attributes of the intended reader cannot be revealed by studying the textbook only, but if textbook authors themselves are expert readers and successful in constructing the textbook for their intended reader, the intended and implied readers coincide. This is the perspective taken in this thesis. There has not been a thorough analysis of the attributes of the textbooks' implied students, or of the textbook authors' intended students. Analyses have been conducted from the perspective that the intended student is one that follows the textbook strictly, works with all the material, interprets it correctly, and comprehends it. That is, the intended and implied students are the same. Tasks have been coded based on the reasoning required for a correct solution. The answer sections have been used to better understand the textbook authors' intentions and expectations. Proof has been interpreted in the sense of Stylianides (2007), as described in Subsection 2.2.1. Occasionally, it has been necessary to check previous textbook sections to determine whether or not a textbook justification can be considered a proof.

Stylianides (2014) discusses four different perspectives for textbook analysis: student, mathematical, teacher, and textbook author. Here, the student and teacher perspectives refer to an *a priori* defined student category, or a teacher planning to enact textbook material in a specific classroom context. Neither of these perspectives has been applied in this thesis; the mathematical and the textbook author perspectives come closer. Justifications in expository sections have been coded based on whether or not they are general proofs. Tasks have been coded based on the type of reasoning required for a mathematically correct solution from a student who correctly interprets the task.

5 Summary of papers

In this chapter, the five papers are summarized. The focus is on their results, but the summaries also include their motives, research questions, and methods.

5.1 Paper I

Paper I is guided by the following research question: What is the nature and extent of the reasoning and proving opportunities offered by secondary-level integral calculus textbooks in Finland and Sweden? There were several reasons for this direction: the extensive use of textbooks in mathematics classrooms around the world, the textbooks' assumed effect on student learning, students' well-documented difficulties with proofs, the central position of calculus in many countries' upper secondary school curricula, and the fact that proofs and calculus have been pointed out as problematic in the transition between secondary and tertiary education. In addition, the neighboring countries Sweden and Finland offer interesting cases as they have similar educational systems, although Finland has outperformed Sweden in international surveys like PISA.

Textbook material on primitive functions and definite integrals from the two most widely used textbook series in each country was analyzed. A total of 93 statements and justifications in expository sections, 235 worked examples, and 1,962 student exercises were classified according to type and nature of reasoning, following an analytical approach adopted (and adapted) from Thompson et al. (2012).

The results show that the percent of justified statements in expository sections varies considerably between the textbooks, from 35% to 89%, but that Finnish textbooks are more likely than their Swedish counterparts to provide justifications. When there are justifications, in the Finnish books they are almost always general proofs, whereas the Swedish books mostly base their justifications on specific cases. The structure of the mathematical theory is made more visible in the Finnish books by use of clear labeling of statements and proofs, while the word "proof" is virtually absent in the Swedish books. In the Swedish books the justifications are placed before the statements, while the Finnish ones place them after as often as before. The Finnish books more often have proof-related worked examples.

The percent of proof-related tasks in the exercise sets is around 10%, and varies from 7% to 18% when textbook series are compared. The percentage is higher in the Swedish books than in the Finnish ones, 14% compared to 8%. Overall, “develop an argument” is the dominating nature of reasoning in proof-related task, and more so in the Finnish textbooks (62% and 76% respectively) than in the Swedish ones (35% and 46% respectively). Other natures of reasoning are hence represented more in the Swedish books, and in one of them “investigate a conjecture” is the most common. The Swedish books have a few “identify or correct a mistake” tasks, while the Finnish ones do not.

The conclusion is that proofs are more visible in the Finnish textbooks, and that the Finnish textbooks more often offer opportunities for learning general proof and tend to focus on proofs as verification. Proof-related tasks in Swedish textbooks reflect a higher variation in proof-related activities and put more emphasis on conjecturing and inductive reasoning, but offer fewer opportunities for transition from an empirical to a deductive proof scheme. In terms of praxeologies, the Finnish textbooks’ treatment is closer to the deductive praxeology of university calculus, while that in the Swedish books falls within the pragmatic praxeology.

Paper I includes a comparison with research findings about United States textbooks that shows that Finnish textbooks offer justifications more often, and more frequently in the form of general proofs. Regarding proof-related tasks, Swedish books are like an average United States textbook when it comes to percentage of general tasks and variation in natures of reasoning.

5.2 Paper II

In Paper II, three research questions are pursued: How can “degree of generality” in proving tasks involving functions be framed? What analytical difficulties arise when proving tasks are classified according to function generality? What can classification according to function generality reveal about textbooks that a “specific-or-general” classification cannot? There are two main motives for this study. The first involves students’ well-documented difficulties understanding the difference between a general argument and what can be concluded from a specific case. The other is that the work with Paper I revealed analytical difficulties in relation to the classification of textbook tasks as either specific or general, and that general tasks could be general in different ways.

Paper II suggests and tests a framework for the classification of tasks’ level of generality based on what class of functions a task asks for reasoning about: reasoning about specific functions is non-general, while reasoning about parametric families of functions is finitely general, and reasoning about even larger classes of functions is infinitely general. The classification's pedagogical relevance is motivated by the different demands these types of reasoning place on the student.

The framework is tested on proving tasks (that is, tasks with the explicit prompt to prove or show something) in integral calculus in three upper secondary textbooks (two Swedish and one Finnish) and one Swedish textbook for a university calculus course. A total of 80 tasks are analyzed.

The implementation of the analysis shows that proving tasks are mostly easy to classify according to function generality. Exceptions to this are tasks in which it is unclear which function class is focused on, or whether the student can be assumed to be familiar with usable parametric representations. Another difficult situation is when the property to prove is a statement about the size of a class of functions, for instance a uniqueness property. There are situations in which a proof requires reasoning with classes of functions that are more general than the statement itself.

The results of the analysis reveal that there is not necessarily a correlation between the number of general proving tasks and the opportunities for students to engage in reasoning about arbitrary functions. Two of the studied textbooks were similar, in the sense that approximately ¾ of their proving tasks were general (according to the principles applied in Paper I). In one of them (a Swedish upper secondary textbook), several of the general proving tasks were of non-general function generality, and only one was infinitely general. The other (a Finnish upper secondary textbook) had no non-general tasks, and a large part were of infinite generality. This means that an analysis of function generality can help determine the extent to which a textbook associates the words "showing" and "proving" with general statements and offers opportunities to reason about arbitrary functions, or sets of functions, that do not lend themselves to parametric representations.

5.3 Paper III

The articulated aim in Paper III is to examine aspects of proof-related reasoning that risk being missed when a framework like the one by Thompson et al. (2012) is used. The paper also aims to contribute to a more refined conceptualization of opportunities to learn proof-related reasoning in mathematics textbooks. The background is that several research studies have used similar analytical approaches, with the advantage that findings are easily comparable. The disadvantage is the risk that important aspects are constantly missed.

The investigation in Paper III is based on two ideas. The first is that analytical difficulties are indicators of potential weaknesses in the framework that is used. The second is that by comparing different textbooks' coverage of the same mathematics content, differences can be discovered that are not captured by the framework. Such differences indicate areas where the framework can be developed.

An in-depth analysis of textbook passages from three different upper secondary textbooks (two Swedish and one Finnish) was conducted. The passages included the textbooks' introduction of primitive functions and the proof of their uniqueness (up to an additive constant). Detailed accounts of analytical issues, and of differences not captured by the framework, were created.

The findings are condensed into four suggestions for inclusion in an analysis of justifications in expository sections:

- Definitions and theorems and levels of generality in statements and justifications. There are different levels of generality in statements and justifications, and the validity of a justification can only be judged in relation to the statement it justifies and to relevant definitions.
- The forms of representation that are used and their purposes. The use of words, symbols, and diagrams affects clarity, precision, and understanding of justifications.
- Structural aspects of the mathematical theory. The role of proof in mathematics can be more or less explicit depending on how the logical structure within and between definitions, statements, and justifications is emphasized.
- The ordering of statements, justifications, student activities, and worked examples. Student activities and informal justifications placed before statements offer opportunities for conjecturing, while general proofs placed after a statement emphasize verification.

All the suggestions focus on object properties of proof-related reasoning and hence on opportunities for a reification of the proof concept.

5.4 Paper IV

Paper IV is guided by a research question similar to that in Paper I: What characterizes opportunities to learn proof-related reasoning offered by Swedish and Finnish upper secondary textbooks? The underlying reasons were also similar: students' well-documented difficulties with proof and textbooks' central role in classroom practice. In addition, Paper IV aims to clarify which results from Paper I are generalizable to algebraically oriented topics.

Textbook material on logarithms and combinatorics from three of the textbook series (two Swedish and one Finnish) analyzed in Paper I was included. A total of 33 statements' justifications in expository sections, 128 worked examples, and 992 student exercises were analyzed and classified according to type and nature of reasoning, following the same principles as in Paper I.

Most findings about logarithms and combinatorics are in line with those concerning integral calculus reported in Paper I. Expository sections typically address general results and label them as laws or principles. The Finnish books have a more detailed and formal exposition than the Swedish ones, and head more directly for general formulations. On the other hand, the Swedish books usually offer activities through which general results can be conjectured. While it is slightly more common for a main result to be justified in the Swedish books (in integral calculus, justifications are more frequent in the Finnish ones), in the Finnish books justifications are almost always general proofs whereas those in the Swedish books are most often based on specific cases. Justifications often precede the statements. Proofs are more visible in the Finnish books, which use phrases like "we prove", but proofs are never labeled as such (which they are in the Finnish books on integral calculus). Only the Swedish books contain worked examples that are proof-related (in integral calculus, proof-related worked examples are more common in the Finnish than the Swedish books).

The Swedish books offer more exercises on logarithms and combinatorics than the Finnish ones do, and a higher percentage of the exercises are proof-related. But while only 35% of the proof-related tasks are general in one of the Swedish textbook series and 60% in the other, almost all proof-related tasks in the Finnish books are general. In the Finnish books, all but one of these tasks are proving tasks. In the Swedish books, arguments are asked for with wording other than "prove that" or "show that", for instance as "explain why" or "motivate why", and about a third of the proof-related tasks ask the student to make or investigate a conjecture. There are also proof-related tasks involving other natures of reasoning. However, none of the analyzed books contain tasks that ask for a counterexample or an outline of a proof.

Two "new" natures of reasoning were identified during the analysis: explaining the thinking behind a presented argument, and matching data in a contextual description to the premises of a specific theorem. Such activities are suggested as input to a discussion about revisions of the analytical framework itself, and of what competencies students need to practice and develop in relation to proofs.

The results on logarithms and combinatorics are also combined with those on integral calculus from Paper I. Together, they convey a picture that opportunities to learn proof-related reasoning are generally few, but are oriented more toward deductive reasoning in the Finnish textbooks and more toward empirical reasoning and conjecturing in the Swedish ones.

Finally, the results are compared with findings about logarithms in United States textbooks reported in Thompson et al. (2012). The Finnish and Swedish textbooks tend to be more oriented toward general justifications than an average United States textbook. Proof-related student tasks are more often general in Swedish and Finnish books, and justifications in expository sections are always general. In the typical United States textbook, only half of the justifications are general. The percentage of proof-related tasks is higher in a typical

United States textbook than in the Finnish books, but lower than in the Swedish ones. Swedish and United States textbooks are similar regarding the variation in nature of reasoning in proof-related tasks, and are not as focused on proving tasks as the Finnish books are.

5.5 Paper V

Paper V aims to answer the following research question: What topic-specific characteristics of opportunities to learn proof-related reasoning are seen in upper secondary school textbooks? An underlying assumption is that different topics offer different kinds of opportunities for learning proof that need to be described, explained, and utilized. It has been advocated that reasoning and proof be a central part of all topics in the school curriculum, but research has indicated that curriculum materials offer teachers limited guidance in this.

The quantitative results regarding types and natures of reasoning reported in Papers I and IV are used as a starting point. The data sample consists of textbook material from three textbook series (two Swedish and one Finnish) and four mathematics topics (logarithms, primitive functions, definite integrals, and combinatorics), and includes 98 justifications of statements in expository sections and 2,272 student exercises. Topic-specific patterns in expository sections are uncovered and characterized by studying the relation between the proportion of addressed results that are justified and the percentage of justifications that are general proofs. Similarly, patterns in exercise sets are described in reference to how the proportion of tasks with proof-related reasoning relates to how such tasks are distributed over different types and natures of reasoning.

The percentage of proof-related tasks is low in all topics in all textbooks, and never exceeds 25%. Most proof-related tasks are of the "develop an argument" nature – in most cases there are more tasks in this category than in the other categories together. Given this, the findings in Paper V can be summarized as follows.

Logarithms are characterized by few justified mathematics results; but when justifications exist, they are general proofs. Proof-related tasks are less frequent than in the other topics, but relatively many of them involve general cases, at least in two of the three investigated textbooks. Logarithm sections also offer relatively many "make a conjecture" tasks.

The treatment of primitive functions differs among the textbook series. Justifications are relatively common in two of them, but in only one is it the general arguments that dominate. Proof-related tasks are more common than in sections on logarithms. There are almost as high percentages of "investigate a conjecture" tasks as there are of "developing an argument" tasks.

Definite integrals is a topic in which addressed mathematics results are seldom justified; even less so than in sections on primitive functions. In addition, existing justifications tend to be based on specific cases. Approximately, proof-related tasks on definite integrals are as common as on primitive functions, and hence more common than in sections on logarithms. Compared to primitive functions, proof-related tasks are more often general. "Investigate a conjecture" tasks are relatively common.

Combinatorics, finally, is characterized by a high percentage of justified statements, but general arguments are only emphasized in one of the textbooks. Proof-related tasks are more common than in sections on logarithms, but in relation to the other investigated topics there are no clear patterns except that "develop an argument" is the most dominant in combinatorics. There are also relatively many "investigate a conjecture" tasks.

The paper also suggests explanations for the topic-specific findings that originate in the mathematics itself, and uses mathematical details of the four topics to point to aspects of proof-related reasoning that could be emphasized more:

Logarithm rules look very similar and are proven in similar ways, and it is easier to construct a general proof than to find a generic case. Hence, it is a reasonable choice to give general proofs for some formulas and omit justifications for the rest of them. Instead, students can be asked to conjecture and prove the general formulas. This opportunity can be made more visible in the textbooks. The investigated textbooks also miss the opportunity to discuss existence and uniqueness, which are mathematical features that are frequently discussed in university mathematics.

As logarithm rules correspond to power laws, calculation rules for primitive functions correspond to differentiation rules. All of them can be proven using the same technique: differentiation. This may explain why justifications are often omitted in sections on primitive functions. If results are presented for families of functions, generic cases can easily be constructed, which is a plausible explanation for why two of the textbooks base their justifications on specific cases. A possible conclusion is that there is an opportunity here to let the student make conjectures from generic cases and prove them. This opportunity is not fully utilized by the textbooks. Another example is the representation formula $F(x) + C$ for primitive functions. This is a rare example of how an infinite class of objects can be represented by a closed formula. In addition, it is a result that can be formulated as an equivalence and proven by studying necessity and sufficiency one at a time. This is another opportunity that is not utilized by more than one of the analyzed textbooks.

The theoretical complexity of definite integrals is a plausible explanation for why justifications are few and general proofs are rare. With only one exception, the results addressed in expository sections are calculation principles. Only one or two are presented with general proofs. However, two of the books contain an example in which a principle from physics is used as an argument for a mathematical theorem. Therefore, even if this topic offers few opportunities for learning general proof, there can be opportunities for connecting mathematical principles to applications, and for explorative activities. The investigation of student tasks also indicates that this topic offers relatively many opportunities for conjecturing.

Combinatorics has a rather simple theoretical foundation (the addition and multiplication principles), which may explain the high number of justified statements. Here generic cases come naturally, as do opportunities for transitions between different representations. One can argue that such transitions are central to completing a proof, which is in contrast to the other topics in which algebraic manipulations are often the core of the proofs. There is also an opportunity for the inclusion of completely non-algebraic arguments concerning existence if content such as the Dirichlet box principle is included in the curriculum.

6 Discussion

The work with this thesis has been guided by three research questions that also encompass the research questions of its separate papers:

1. What characterizes opportunities to learn proof-related reasoning offered by Swedish and Finnish upper secondary mathematics textbooks?
2. What topic-specific characteristics of opportunities to learn proof-related reasoning are seen in upper secondary mathematics textbooks?
3. How can opportunities to learn proof-related reasoning be conceptualized and analyzed?

In this final chapter of the thesis, the first section (6.1) will focus on the empirical findings related to the first and second research questions. They are discussed against the background of student difficulties and misconceptions that are well-documented in the literature. The third research question is discussed in the second section (6.2). This discussion is based on the adaptations and refinements to the analytical frameworks that have been made in the various papers. Then follows a section (6.3) on implications for teaching, with suggestions for improving opportunities to learn proof-related reasoning. The chapter concludes with a methodological discussion (6.4) and suggestions for future research (6.5).

6.1 Contributions from empirical findings

The summary of empirical findings is structured “top-down”. First come the general characteristics seen in the accumulated data and throughout most of the analyzed data material. Then, the discussion zooms in on specific features of countries and topics. This structuring serves two purposes. The first is to bring some order to what the character of opportunities to learn proof-related reasoning *is*. The other is to suggest possible explanations of what it is that *determines* this character. Findings presented as context-specific can be the result of national steering documents or educational traditions of that country, while findings presented as topic-specific can be a consequence of purely mathematical aspects of that topic.

Many findings are based on absolute numbers or relative frequencies of justified statements, general proofs, proof-related tasks, general tasks, etc., in the textbooks. The impression might be that the higher the number, the better the book; this would be a misinterpretation. There is a limit to what a textbook can cover and how many tasks a student has time for. If all theorems are formulated in detail in the textbook, then there is little room for students' inquiry into the exact circumstances under which a statement is true. If all statements are justified with general proofs, then there is no room for the students to provide the proofs. If all tasks are proof-related, then there are few opportunities to develop other competencies. If all proof-related tasks are general, then students will not have opportunities to learn the different roles of specific and general arguments. If all proof-related tasks are related to developing arguments, then there is no room for the evaluation of arguments. The list can be made longer. It is impossible to say what the ideal percentage of proven statements and proof-related tasks is, or what the perfect mix of types and natures of reasoning is. But low absolute numbers indicate few opportunities to learn, and low relative frequencies indicate little emphasis on this kind of reasoning, unless students are provided other learning opportunities.

6.1.1 General findings

Throughout the analyzed material there are certain findings that seem to be independent of textbook series and topic. Here, these are summarized as three general findings.

The first is that *proof and proof-related reasoning do not have a prominent position in the textbooks*. In one of the textbooks, 23% of the logarithm tasks are proof-related. This is the highest observed frequency in the analyzed textbook material; on average, approximately 10% of the tasks are proof-related. As the definition of proof-related is broad and includes almost all kinds of reasoning, this means that the opportunities for proof-related reasoning are few. Tasks asking the student to "prove" or "show" are but a subset, so opportunities for constructing proofs are even more rare. In addition, there are few proof-related worked examples that can serve as role models for the students.

The figures for Swedish and Finnish textbooks are higher than those reported for United States non-geometry textbooks, in which approximately 5% of the tasks were proof-related (Thompson et al., 2012). On the other hand, there was great variation between United States textbook series. In geometry, for instance, 20–38% of the tasks were proof-related (Otten, Gilbertson, et al., 2014). When comparing the figures from these studies one must also bear in mind that, even though the research groups have used the same analytical frameworks, they may have made different analytical distinctions.

In expository sections, approximately 60% of addressed mathematics results are justified. The figures in the United States textbooks were the same. In the Finnish case the justifications are mostly proofs, while in the Swedish case they usually are not. It is difficult to use these figures to draw conclusions about opportunities to learn proof. A mathematics result presented without proof can be an opportunity for the student to find the proof. However, this opportunity is untapped as there are only a few examples of the textbooks explicitly asking the student to provide an omitted proof.

The second general finding is closely related to the first: *there are few opportunities for reasoning with general cases.* Only half of the proof-related tasks in the Swedish books are general. In the Finnish books three quarters are general, but the Finnish books have a lower frequency of proof-related tasks. These figures are similar to those reported for United States textbooks (Thompson et al., 2012). This means that such an important feature of mathematics, that pure reasoning can offer absolute certainty about universal statements, is an exception in the textbooks' exercises. The observed pattern is also that the higher proportion of proof-related tasks, the lower the share of tasks that include reasoning about general cases.

The third general finding is that *there is little variation in the structure of statements, proving methods, and natures of reasoning activities*. Addressed mathematics results are usually general, universal conditional statements, typically in the form of a formula or calculation principle: logarithm rules, elementary primitive functions, rules of integration, and formulas for permutations and combinations. There are a few examples of existence results, but none of uniqueness results. There is only one example in the analyzed material where it is emphasized that a statement is an equivalence. Most proofs are direct derivations. There are no examples of contrapositive proofs or proof by contradiction. Sometimes the textbooks use specific cases to justify a result and then follow up by presenting a general proof, but in no cases are two different valid proofs presented for the same statement.

The most common proof-related task involves developing an argument: to prove, to show, or to explain something. Typically, it is the justification of a formula that the task asks for. To some extent there are tasks focusing on making or investigating conjectures, and occasionally on the evaluation of arguments, or finding or correcting errors. But such tasks are few compared to those involving developing arguments. Students are never asked to provide a counterexample or outline a proof.

Common to the general findings is that areas where opportunities for learning are rare correspond well to areas where student difficulties are well-documented. Students have difficulty understanding proofs (e.g., Healy & Hoyles, 2000), writing proof (e.g., Senk, 1985, 1989), and formalizing informal arguments (e.g., Stylianides et al., 2017). The analyzed textbooks offer few opportunities to read general proofs or to learn what a proof is and the importance of proofs in mathematics. There are also few opportunities to engage in proof-related activities in textbook tasks. Tasks in which students can outline or fill in the details of a proof are nonexistent.

More specifically, common difficulties include the role and meaning of counterexamples (Balacheff, 1988b; Ko & Knuth, 2013), implications (Durand-Guerrier, 2003), negations (Antonini & Mariotti, 2008), and quantifiers (e.g., Durand-Guerrier & Arsac, 2005). Students have difficulty with statements with a complex structure (Zandieh et al., 2014), and frequently confuse conditional statements with their converses and inverses (Hoyles & Küchemann, 2002; Stylianides et al., 2004). In the analyzed textbooks, opportunities to read proofs typically involve direct proofs of general statements in the form of computational rules. Other kinds of statements (such as existence and uniqueness properties) and other kinds of justifications (such as contrapositive proof or proof by contradiction) are rare. There are few opportunities to learn the differences between necessary and sufficient conditions. Proof-related tasks are also typically devoted to the derivation of formulas, and there are no tasks asking for counterexamples.

It is well-known that empirical arguments often convince students about generalizations (e.g., Almeida, 2001; Healy & Hoyles, 2000; Morris, 2002; Sevimli, 2018; Sowder & Harel, 2003), and that students have difficulty distinguishing invalid arguments from valid proof (Inglis & Alcock, 2012; Selden & Selden, 2003). In the analyzed textbooks, many mathematical principles are justified with specific cases, and a considerable part of the proof-related tasks only involve reasoning with specific cases. There are few opportunities to evaluate or correct arguments.

Given this, the general findings from the textbook analyses are not surprising. On the other hand, one can argue that the knowledge about students, combined with an increased focus on reasoning and proof in steering documents and policies, would eventually influence curricular resources such as textbooks toward more emphasis on proof-related reasoning.

6.1.2 Context-specific findings

When Swedish and Finnish textbooks are compared, there are some characteristics more strongly connected to the Finnish books than the Swedish ones and vice versa. Three such context-specific findings are highlighted below.

The first is that *proofs and structural aspects of mathematics are more visible in the Finnish textbooks than the Swedish ones*. In the Finnish textbooks the labeling in expository sections makes clearer distinctions between definitions and theorems, and between arguments that are and are not proofs. In the Finnish books, general proofs are more frequent. In the Swedish books, justifications are mostly based on specific cases; but even when they are general proofs, they are not labelled or referred to as such.

The second context-specific finding is that *the Finnish textbooks place greater emphasis on deductive reasoning and reasoning about general cases than the Swedish ones do*. Justifications in expository sections in the Finnish textbooks are mostly general proofs. Proof-related tasks are few, but more frequently involve reasoning about general cases than in the Swedish textbooks, and are mostly proving tasks. In the Finnish textbooks, proving tasks can even be “infinitely general” in the sense that they involve large classes of functions.

In the Swedish expository sections, most justifications are based on specific cases. Proof-related tasks are more frequent in the Swedish textbooks, but involve reasoning about general cases to a lesser degree. Proving tasks are not as dominant, and even proving tasks can be about specific cases only.

The third context-specific finding is that *there is more variation in natures of reasoning and greater emphasis on conjecturing in the Swedish textbooks than in the Finnish ones*. Tasks about providing justification for a given claim are dominant in both countries' textbooks, but less so in the Swedish materials. Proof-related tasks that involve conjecturing are few but more common in Swedish textbooks than Finnish ones, as are tasks in which students are to evaluate or correct arguments. There are also examples of tasks in the Swedish materials in which the student is to match given data to the premises of a theorem, or suggest plausible thinking behind a presented argument.

The differences between the Swedish and Finnish textbooks can be summarized as follows: the Swedish materials are slightly more oriented toward conjecturing, evaluation, and the use of empirical arguments, while the Finnish ones are slightly more oriented toward general deductive arguments and formalism. These differences are difficult to explain from a steering document point of view. Proofs are not emphasized more in the Finnish steering documents than the Swedish ones; it is rather the other way around (Hemmi, Lepik, et al., 2013). Perhaps one explanation for this could be that all analyzed topics are in the second half of the Finnish upper secondary curriculum, but closer to the middle of the Swedish one. As was found in Thompson et al. (2012), general proofs, proof-related tasks about general cases, and tasks focusing on deductive arguments tend to be more common toward the end of upper secondary curricula than the beginning. Another tentative explanation could be that there is a more formal and deductive tradition in Finnish mathematics education, while Swedish mathematics education is more affected by reform ideas from the United States (like those expressed in the NCTM Standards) and is therefore more oriented toward inquiry and conjecturing.

One should be careful in declaring which countries' textbooks are the best. Given students' difficulties with understanding and writing proof (e.g.,Stylianides et al., 2017), distinguishing between axioms, definitions, and theorems (Vinner, 1977), and understanding the logical structure of statements and arguments (e.g., Zandieh et al., 2014), and the fact that students find proof invisible (Hemmi, 2008), the character of the Finnish textbooks would be preferable. On the other hand, students' difficulties with evaluating arguments (e.g., Selden & Selden, 2003), and the important role of inquiry and conjecturing in relation to learning proof (e.g., Garuti et al., 1998), speak in favor of the tendencies of the Swedish books.

6.1.3 Topic-specific findings

Some characteristics of opportunities to learn proof-related reasoning seem to be connected to the mathematical topic, rather than to textbook series or country.

Logarithmic properties are typically not justified in the textbooks, but when they are, they are justified with general proofs. This is in line with the findings in Thompson et al. (2012). A plausible explanation for this could be that proofs of elementary logarithmic laws are very similar. The laws also correspond to laws for powers and exponents, which students can be assumed to be familiar with. Consequently, logarithms offer a possibility to engage the student in conjecturing and proving activities. When compared with the other analyzed topics, one can also see that proof-related tasks on logarithms more often involve conjecturing. In expository sections students are not explicitly told to provide omitted proofs, but such proofs are included in the student tasks anyway.

The definition of logarithms offers an opportunity to discuss fundamental mathematical questions about existence and uniqueness, questions that are rare in the upper secondary mathematics curricula. This opportunity, however, is not utilized by the textbooks. The conclusion, therefore, is that by relatively small means, proof-related reasoning can be elevated in the teaching of logarithms.

Primitive functions offer relatively good opportunities for proof-related tasks, but seldom about general cases. As with logarithms, the close correspondence with differentiation rules should be able to be used for proof-related reasoning. This is seen to some extent in the analyzed material. Differences between the textbook series also point to several areas where proof-related reasoning can be emphasized. For instance, the formula $F(x) + C$ for the primitive functions to $f(x) = F'(x)$ is one of very few examples of how a certain class of functions can be represented algebraically, and it also provides an opportunity to discuss necessary and sufficient conditions and how equivalences are typically proven. Also, parameter families of functions (such as $\sin(kx)$) offer easy examples for conjecturing and the justification of formulas for primitive functions, either by general arguments or by discussion of generic cases.

Textbook sections on definite integrals include relatively few justifications, and when they are provided they are frequently based on specific cases. Given the theoretical complexity of this topic, this is not surprising. On the other hand, even with a focus on the use of computational principles, students have opportunities to read and apply principles that are valid (and expressed for) large classes of functions. One of the analyzed textbook series (Fi2 in Paper I) explicitly takes as an assumption (an axiom) that the area bounded by the graph of a continuous function is well-defined and satisfies certain conditions. With this as a starting point, it is possible to define definite integrals and derive fundamental properties rather rigorously on a level comprehensible to upper secondary students. There are also interesting examples of principles from physics being used to justify mathematical theorems. While such lines of reasoning are not valid proof, they may have other values in a classroom setting. The close connections to applications also offer possibilities for experimental work, inquiry, and conjecturing.

Combinatorics, finally, is the only topic in which essentially all addressed mathematics results are justified, albeit not with general proofs. This topic typically offers possibilities for general reasoning with or without proof by induction, or by use of generic examples. Combinatorics differs from the other investigated topics in that it does not require high levels of pre-knowledge. Hence, proof-related activities need not be hindered by, for instance, a lack of algebraical abilities. Indeed, the Dirichlet box principle is an example of a trivial principle that offers opportunities for proof-related reasoning with almost no algebra involved.

6.2 Contributions to frameworks and conceptualizations

During the work with this thesis and its papers, there have been several occasions when the analytical frameworks used have had to be adapted. Analytical difficulties as well as findings have pointed to a need for adaptations of differing kinds: sharper distinctions between categories, more refined frameworks with new subcategories, broader frameworks with new categories, and new approaches and new kinds of analyses. Such ideas have been touched upon in all papers, and focused on in Papers II and III. This subsection brings all these considerations together, resulting in a rather comprehensive framing of proof-related reasoning for textbook analysis. A textbook analysis need (or should) not employ all its parts. Doing research requires a focus on certain specific aspects while others are left out. However, by presenting all parts together one gets a picture of the complexities involved and all the different analytical questions the textbook analysist may consider investigating. Hopefully, this can serve as a starting point for other researchers planning textbook studies of opportunities to learn proof and proof-related reasoning, and who have chosen approaches similar to that of Thompson et al. (2012) and Otten, Gilbertson, et al. (2014), as well as those used in the papers of this thesis. Also, this summary contributes to a more refined conceptualization of what opportunities for proof-related reasoning encompass in relation to textbooks.

6.2.1 Proof-related reasoning in expository sections

The categorization of justifications in expository sections, according to Table 1 (p. 37), combines the question of who takes (or is given) responsibility for providing justifications, and what type of justifications the textbooks provide. There can be reason to study the first of these two aspects in more detail. There is a difference between leaving a mathematics result unjustified without further comment, and saying that "one can prove that", or "Euler proved that". On the one hand the latter alternatives suggest that the existence of a proof is important, while on the other they still leave the student to trust an unspecified person or an unknown authority (cf. authoritarian proof schemes (Harel & Sowder, 2007)).

The second aspect, whether or not a provided justification is a general proof, includes several considerations. One is the generality of the justification and how it relates to the generality of the justified statement. In Paper II, reasons for a more refined characterization than just "general or specific" are offered in relation to student tasks, but can be applied to textbook justifications as well. The validity of a proof is also a question about formalism and levels of detail. Paper III argues for an analysis that considers how justifications are "embedded" in expository sections. In the various papers, other aspects such as labeling, ordering, and explicit connections to other parts of the expositions are analyzed. The logical structure of statements and justifications, and the proving methods and strategies, have also been studied. Investigations of such details have their value, as it is well-known that many students find proofs invisible, have difficulty distinguishing theorems from definitions and axioms, have problems understanding statements with complex logical structures, are unconvinced by contrapositive proofs, etc.

Table 8 summarizes all these ideas, extending and refining the type of analysis of expository sections that has been conducted and discussed in the papers of this thesis. There are likely many other relevant questions when investigating opportunities for proof-related reasoning in textbook expositions. However, even if the list in Table 8 is incomplete, it shows the multitude of questions that textbook analysists, textbook authors, and mathematics teachers need to have in mind, and the many elements that constitute proving competence.

Table 8. *Analytical questions for analysis of expository sections*

General questions	Examples of specific questions
What kind of mathematics results are addressed?	Specific or general statements? General in what sense? Universal or existential statements? Multiply quantified statements? Uniqueness statements? Statements with sufficient conditions, necessary conditions, or both?
Can mathematics results be conjectured?	Are students given opportunities to conjecture addressed theorems and results? On what grounds?
Who is responsible for providing justifications?	No one? The student? The textbook author? The teacher? Some other authority or unspecified person?
What kind of justifications are provided in the textbook?	Reasoning with specific or general cases? General in what sense? Generic cases? Do the justifications qualify as proofs? Why/why not? Level of formalism and detail? Intuitive arguments? What forms of representation are used? Algebraic, verbal, graphic, or other? What proving methods and strategies are used? Mathematical induction? Separation of necessity and sufficiency? Division into cases? What logical rules of inference are used? Direct proofs? Proof by contradiction? By contrapositive? How are examples used? Supportive examples? Counterexamples?
What purposes do the justifications serve?	Verification? Explanation? Systematization? Discovery? Communication? Intellectual challenge?
How is the mathematical structure made visible?	How are statements and justifications labeled and talked about? Are axioms, definitions, and theorems separated? Is it clear when justifications are proofs and not? What is the order between justifications and statements?
Meta-perspective	Are there explicit connections to other parts of theory? Are proving methods and ideas discussed and compared? Are there several proofs of the same statement?

6.2.2 Proof-related reasoning in tasks

Proving as an activity, meant to produce proofs or develop the necessary competencies for proof construction, has been conceptualized as a set of natures of reasoning (Table 3, p. 38) that represent separate proof-related activities. Some are broad while others are narrow. For instance, "develop an argument" refers to almost all kinds of tasks in which the student is asked to justify a given claim, while "counterexample" refers to the very specific situation of the student being asked to provide a counterexample to a false claim. One can argue that to provide a counterexample is also to develop an argument, albeit of a very specific kind. "Investigate a conjecture" can also involve finding a counterexample. Instead of viewing these natures of reasoning as separate, one can see them as related to different aspects of the same proof-related activity. This can be made clearer by splitting and sorting the natures of reasoning in another way, by viewing some of them as subsets of others, and by introducing new categories and subcategories. The result is the rather comprehensive, yet still tentative, framework presented in Table 9 and Table 10. In order to keep this collection of ideas separate from the natures and types of reasoning employed in the various papers of this thesis, they are now referred to as *reasoning activities* (Table 9) and *properties* of reasoning activities (Table 10).

The original framework by Thompson et al. (2012) had seven distinct natures of reasoning. In what follows we will discuss these one at a time, how they have been interpreted and adapted for the present study, and how they are parts of the more comprehensive framework.

The first category was "make a conjecture" (M tasks). Typically, students are presented with a finite pattern and must reason inductively to conjecture how the pattern continues. Thus, the student must find a mathematical principle for how the pattern develops. However, there are other kinds of tasks in which the student is asked to specify a mathematical principle. For instance, in Paper I some tasks were found in which the student must find a geometric formula for a volume by use of integral calculus. Such tasks were also included in the M category. Thus, this category includes finding a (true) mathematical statement by derivation, as well as making a conjecture by "guessing". Another typical feature of M tasks is that the premises are given; that is, the student's task is to specify the conclusion of a conditional statement. In the analysis reported in Paper IV, however, textbook tasks were encountered in which the student must specify the premises. In principle, there could also be situations in which one must find the logical relation between two properties. Hence, M tasks can focus on (at least) three different parts of a mathematical statement, all of which are central in relation to conjecturing (and hence in the process of finding a proof, e.g. Lakatos (1976)). In Table 9, this category is therefore renamed "develop a statement". To specify premises, conclusions, or logical relations are listed as examples of subactivites.

The second category was “investigate a conjecture” (I tasks). This can also refer to tasks in which students study a pattern of some sort, but now the textbook suggests a principle or conclusion and the student must decide whether or not it is true. In this thesis, however, the I category has been used in a broader sense, in analogy with the use of the M category. This means that whenever the student has been asked to decide whether or not a given statement is true, this has been considered an I task. Therefore, in Table 9 this category of activities is named “investigate a statement”. It is implicit that the student is assumed to argue for his/her opinion.

The third category, “develop an argument” (D tasks), was the broadest category and the one in which most of the identified proof-related tasks were placed. Below is an argument for an even broader interpretation that includes arguments by counterexample, correction of arguments, and proof sketching. In Table 9, this category keeps the name “develop an argument”.

The fourth category was “evaluate an argument” (E tasks). This category has been used for all tasks in which an argument is presented to the student and the student must decide whether or not the argument is correct. For an argument with a flaw, this need not include identifying *what* is wrong but simply determining *that* something must be wrong. Note that it should not be given whether or not the argument has a flaw. In Table 9, however, the identification of errors has been included in this category and it has been given the name “investigate an argument”.

The four natures of reasoning discussed so far can be regarded as the main reasoning activities under which most other proof-related reasoning activities can be sorted. These four can also involve highly similar reasoning activities. What separates them from one another is how the task is formulated, and what information is included. The discovery of the exact formulation of a statement can require writing down a full proof. Hence, “develop a statement” and “develop an argument” can include similar activities. In “investigate a statement” activities, the student is not told whether to argue for the truth or falsity of the statement. As some evidence for why the statement is true (or false) must be presented, such reasoning activities can be similar to “develop an argument” activities. There can also be considerable differences within these four categories of reasoning activities. Usually, a different kind of argument is needed in order to determine that a statement is true, compared to determining that it is false. Evaluations of correct and incorrect arguments may also be very different. In the right column of Table 9, some fundamentally different situations related to the four main reasoning activities are presented.

Table 9. *Reasoning activities in textbook tasks*

Main activities	Examples of subactivities
Develop a statement	Specify premises
	Specify conclusions
	Specify logical relations
Investigate a statement	Investigate true statements
	Investigate false statements
Develop an argument	Argue for true statements
	Argue against false statements
Investigate an argument	Evaluate valid arguments
	Evaluate invalid arguments

The three remaining natures of reasoning were "correct or identify a mistake" (C tasks), "counterexample" (X tasks), and "outline a proof" (P tasks). In Paper IV an "other" category was also introduced (O tasks). Instead of viewing these as separate from the first four natures of reasoning, it is now suggested that they be used for further characterization of properties of the four main activities.

"Correct or identify a mistake" (C tasks) comprises two kinds of activities. Correcting a mistake entails developing a correct argument starting from an incorrect one, while identifying a mistake is a form of investigation of an (invalid) argument. The difference from tasks in the original E category is that here it is given that there is an error (but not where the error lies). Thus, E tasks and tasks involving identifying a mistake both entail investigating an argument. Hence, C tasks can be either "develop an argument" or "investigate an argument" activities. Making corrections are relevant in relation to several reasoning activities. For instance, a "develop a statement" activity can take an incorrect statement as its starting point. Therefore, "make corrections", is now placed in the list of properties in Table 10.

As discussed above, "counterexample" (X tasks) refers to tasks in which the argument against a false claim should be of a very specific kind: providing a counterexample. Thus, X tasks can be placed within the "develop an argument" category. Of course, counterexamples may be the best way to reveal that a statement is false; that is, counterexamples may occur in relation to an "investigate a statement" activity. An "investigate an argument" activity can also include an argument that presents a (potential) counterexample. Thus, counterexamples can occur in all four main kinds of reasoning activities. For this reason, "counterexample" is now listed as a property in Table 10, within the category called "logic". Examples in support of a statement can also be part of all four reasoning activities. Therefore, "supportive example" is another property included in Table 10. Both uses of examples are important when learning the correct role of examples in mathematical arguments.

“Outline a proof” (P tasks) involves sketching a proof, to develop an argument of a less formal form, without all the details of a full proof. Here, one can also have the “opposite” situation: filling in details in a proof outline. This is also a form of argument development. Thus, P tasks are viewed as “develop an argument” activities, while “proof outline”, “fill in details”, and “full proof” are listed as properties within the category “detail and formalism” in Table 10.

Two kinds of tasks studied in Paper IV were placed in the “other” category (O tasks). The first involved explaining the thinking behind an argument, which is yet another kind of “develop an argument” activity. The second kind involved tasks in which a real-life situation was to be mapped to the premises of the Dirichlet box principle. When wanting to apply a theorem in an argument, one must check that the premises are fulfilled. Hence, this can also be considered an element of argument development. “Explain underlying ideas” and “formalize” are included as properties in Table 10.

To summarize, the framework for natures of reasoning as employed in the textbook analysis reported in this thesis can be condensed into four main kinds of reasoning activities that capture the essentials of the original M, I, D, and E categories. The properties connected to the C, X, P, and O categories are other aspects of reasoning and argumentation. As arguments can be developed as well as investigated, and used when investigating as well as developing mathematical statements, it might be better to study these aspects independently. This is why they have been listed as properties of reasoning activities in a separate table (Table 10).

In addition, most characteristics of statements and justifications detailed in Table 8 also apply to tasks. Statements and arguments in tasks can be specific or general. General statements/arguments about functions can involve specific functions, parameter families of functions, or “arbitrary” functions. The statements that are investigated (or argued for) can be universal or existential, and can express uniqueness or equivalence. An argument that is investigated (or asked for) can be in the form of a positive proof, a contrapositive proof, or a proof by contradiction. It can be communicated using different forms of representations. It might be necessary (or preferable) to split an argument into cases, or to study necessary and sufficient conditions separately. A textbook can also provide clear directions, or hints, for how to solve the task, or give no clues at all. Table 10 is a summary of all these ideas. Note that the category called “generality” includes what has been referred to throughout the thesis as type of reasoning, but in the refined sense proposed in Paper II.

Table 10. *Properties of reasoning activities*

Property	Examples of subproperties
Generality	Specific case(s)
	General “one-parametric” case
	General “finite-parametric” case
	General “non-parametric” case
Detail and formalism	Full proof
	Proof outline
	Fill in details
	Formalize
	Make corrections
	Explain underlying ideas
Structure of statement	Universal
	Existential
	Uniqueness
	Implication
	Equivalence
	Formula
Logic	Positive proof
	Contrapositive proof
	Proof by contradiction
	Counterexample
	Supportive example
Forms of representations	Algebraic
	Verbal
	Graphic
Methods and strategies	Case division
	Necessity and sufficiency separately
	Mathematical induction
Directions	With directions/hints
	Without directions/hints

6.3 Implications for practice

While opportunities to learn proof-related reasoning offered by a textbook are potential opportunities, such opportunities are not guarantors of student learning. Nor does a missing opportunity imply that students will not learn. What real opportunities students are offered depends on how the curriculum is enacted. Teachers can pick and choose from the textbook content, decide on a different focus than the textbook has, emphasize other aspects of reasoning, supplement the textbook materials with other curricular resources, and so on. A standard computational task can offer opportunities for reasoning in a classroom where it is an established norm to always reflect on how a task is solved, why methods and algorithms work, what the general underlying principles are, etc. However, if teachers rely heavily on the textbook there is a great risk that what is not emphasized and made explicit in the textbook will not be taught. In this case, aspects of proof and proof-related reasoning that are not emphasized in the textbook risk being marginalized in classroom practice, and affecting student learning negatively, unless student learning is supported by other means. Consequently, it is important that teachers be aware of the limitations of their textbook resources and that (if needed) they be offered other kinds of support for teaching important aspects of proof-related reasoning. Taking this perspective on the findings in this thesis, they can be used to highlight areas where such support is needed.

It is also clear that the findings reported in this thesis correspond well to what research has shown about students' difficulties with proof. Students have difficulty constructing proofs with arguments about general cases, and hold various misconceptions regarding the role of examples in relation to proving and refuting mathematical statements. This thesis shows that proofs are often invisible in textbooks, that textbooks offer few opportunities for reasoning about general cases, and that students are never asked to provide counterexamples. Research has shown that students have difficulty with the logical rules of inference, and do not know the relations between a conditional statement and its converse, inverse, and contrapositive. This thesis shows that textbooks seldom emphasize the logical structure of statements, and seldom provide other kinds of arguments than direct derivations of formulas.

Against this background, the recommendations in the following subsections should be read as suggestions concerning aspects of proof-related reasoning that one should consider giving more attention to in teaching and in curriculum resources. They are therefore formulated as questions. How they should (or could) be addressed in teaching and textbooks is a matter for future research to investigate.

6.3.1 General suggestions

Considering what Swedish and Finnish textbooks generally offer in terms of opportunities for proof-related reasoning, and what students generally have difficulty with, it is recommended that curriculum designers, textbook authors, teachers, and other users of curricular resources consider the following when determining the design of curricular resources and teaching:

- Does the exposition make clear distinctions between axioms, definitions, and theorems?
- Is the logical structure of addressed mathematics results made clear?
- Does the exposition provide a rich variation of results regarding structure, including existential results and uniqueness results, and necessary and sufficient conditions?

- Are the purposes of presented arguments clear, and is it explicit whether or not they are valid proofs?
- Are the underlying, logical ideas of proofs clear, and are proofs discussed from a meta-perspective?
- Are students given role models for proofs of various kinds – including proof by contrapositive, by contradiction, and by case division – in expositions and worked examples?
- Are there situations in which it could be instructive to discuss different proofs of the same result?
- Can students be given the responsibility for conjecturing and proving some of the central mathematics results?
- Are there enough proof-related tasks, especially those that require reasoning about general cases and reasoning with classes of objects, that cannot be represented with a closed algebraical formula?
- Are there tasks in which the reasoning does not require algebraic manipulations?
- Are there tasks in which the student can practice the complete proving process of investigating special cases, formulate and test hypotheses, and finally prove a general, conjectured principle?
- Are there tasks in which the student can evaluate, correct, or fill in the details of given arguments?
- Are there tasks in which the student can provide supporting examples and counterexamples?

6.3.2 Context-specific suggestions

As described in Subsection 6.1.2, the Swedish textbooks are somewhat more oriented toward inquiry, conjecturing, and empirical arguments, and provide greater variation in natures of reasoning, while the Finnish ones are more oriented toward deductive arguments and general reasoning. This implies that some of the suggestions in the previous subsection are more directed at actors in the Swedish context, and others at actors in the Finnish one.

In the Swedish context, extra attention to the following issues is recommended:

- Is there sufficient support for formal aspects of proving, deductive arguments, and arguments involving general cases?
- Are proofs visible, and is the structure of the mathematics theory and of the addressed mathematics results made clear?

In the Finnish context, extra attention should be paid to the following:

- Are there opportunities for inquiry and conjecturing?
- Are there opportunities for investigating arguments in addition to opportunities for developing arguments?

6.3.3 Topic-specific suggestions

Some aspects of opportunities for proof-related reasoning seem to vary depending on topic. There is reason to believe that the mathematics itself sets boundaries for what is possible and desirable. Given the topic-specific findings and the discussion in Paper V, the list below gives some examples of topic-specific details that one can consider giving special attention. The list is only a selection.

Logarithms:

- Is there a discussion of existence and uniqueness of logarithms related to relevant properties of the exponential function?
- Are the similarities between formulations and proofs of logarithm rules for products, quotients, and powers, and for such laws for different bases, utilized for student activities involving conjecturing and proving?

Primitive functions:

- Is the close connection to differentiation rules utilized for activities involving finding elementary primitive functions and for conjecturing and proving general rules of integration?
- Is it made clear that being a primitive function is a global property?
- Is there a discussion of the existence of primitive functions?
- Is it made clear that the formula $F(x) + C$ is a representation of an infinite set of functions with a common property? Is it made clear that the statement of this formula involves both a necessary and a sufficient condition, and how this can be handled in a proof?

Definite integrals:

- Are there connections between rules of integration and intuitively clear properties of areas?
- Are connections to applications used for designing activities that include the formulation and testing of conjectures?

Combinatorics:

- Are there possibilities to use combinatorial reasoning to justify algebraic identities and vice versa?
- Are there possibilities to introduce proof by induction?
- Would it be advantageous to include the Dirichlet box principle to create opportunities for non-constructive existence proofs, and for reasoning that does not require high algebraic abilities?

6.4 Methodological discussion

The analyses conducted in the papers of this thesis only include the student textbooks. From the perspective of the teacher, teacher guides can offer important support for the teaching design, and hence for creating learning opportunities. Publishers may also provide digital resources such as web-based materials – for the student as well as the teacher. It is difficult to say whether the inclusion of such materials in the analysis would change the outcome; this would depend on whether the teacher guides have a different focus on proof-related reasoning than the textbooks do. However, as the use of teacher guides among Swedish and Finnish upper secondary teachers is not widespread, a reasonable stance is that they do not have a great effect on the enacted curriculum.

A limitation is that only two textbook series from each country have been investigated – and for Papers I, III, IV, and V, only one Finnish one. But as the analyzed textbooks dominate the markets in Sweden and Finland, the findings can be used to say something about what opportunities to learn proof-related reasoning most students in the two countries have been offered in their textbooks. This is how general statements about Swedish or Finnish textbooks should be interpreted. It is also important to note that the analyzed textbook series are intended for students preparing for higher studies in mathematics, science, and technology. Textbooks written for use in vocational programs may treat proofs differently; as such programs do not need to prepare students for the theoretical and formal exposition of university mathematics, one can assume that proof-related reasoning is focused on less in textbooks written for their mathematics courses.

The presented findings are based on an analysis of all justifications of main results addressed in expository sections, all worked examples, and all student tasks. However, only material in four selected topics was included. Further, this selection was not based on a belief that these topics are representative, and nor was it random. Consequently, the findings cannot be generalized to all topics in the mathematics curricula, and the data sample's representativeness cannot be judged by use of statistical tests. A certain "coherence" in style throughout a textbook is plausible, but findings concerning a certain textbook series should primarily be read as findings about that series' treatment of the four investigated topics. Studies of United States textbooks reveal that proof-related reasoning is more emphasized in geometry than in other subjects, and one can expect a similar situation in Swedish and Finnish textbooks. There are also separate sections on proof and proving in other parts of the Swedish and Finnish curricula. The Swedish national syllabus prescribes that elementary logic, implications and equivalences, and illustration of the concepts of definition, theorem, and proof be included in Course 1. In Finland, Part 11 is a course in number theory and logic, which, according to the steering documents, should include the formalization of statements, quantifiers, and proof by contrapositive and by contradiction. Hence, there is reason to believe that there are other parts of the analyzed textbook series in which various aspects of proof-related reasoning are emphasized more.

Regarding the analytical processes, it is a drawback that most of the textbook analyses have been conducted by one person only. Hence, reliability and validity rest heavily on a stringent use of the analytical frameworks. Here, reliability is strengthened by the fact that two researchers collaborated on the analysis for Paper I, compared coding, and determined coding principles that could be used in the other papers as well. When new topics were analyzed, discussions with colleagues aided in maintaining validity. Coding principles are described in the separate papers. Paper IV also includes detailed descriptions of how 18 different textbook tasks were coded so that the reader can judge the relevance and validity of the principles used. Still, there are issues worth extra mention.

First, the coding principles could differ from those used by other researchers using the same analytical frameworks. For instance, the higher rates of proof-related tasks in the Swedish and Finnish textbooks, compared to those reported for United States textbooks by Thompson et al. (2012), may be due to a more generous interpretation of what should count as proof-related. The choice to include all addressed results in the analysis of expository sections may have led to different rates of justified statements than if the analysis had focused on an *a priori* selected list of statements (as in (Thompson et al., 2012). On the other hand, these methodological choices should affect comparisons with other studies more than those between different parts of the analyzed textbook material.

Second, it has been difficult to formulate a principle for determining whether a task offers reasoning about a general case that works independent of topic. The requirement that proof-related tasks involving functions be counted as general only if they have some degree of freedom beyond an independent variable might be too restrictive for the comparison with other topics to be fair. The findings also indicate that there were fewer general tasks in integral calculus than in sections on logarithms and combinatorics. On the other hand, the tasks coded as general would have been even fewer with a requirement that they invoke infinite classes of functions. This shows that what “general” as a concept should mean needs to be specified. As is elaborated on in Paper II, there are “degrees of generality”.

6.5 Future research

Based on the methodological discussion, more solid evidence for this thesis’s findings requires investigations of more topics in more textbook series from more countries, and that the textbooks be analyzed independently by several researchers. Given that the findings indicate general patterns regarding textbooks from different educational contexts and topics, research can take other directions. Based on the findings, three specific areas for future textbook research are proposed below. Then, the perspective is widened beyond the textbook itself.

One of the findings in Papers I and IV was that there are topics in all analyzed textbook series that lack proof-related worked examples. As there were so few of these, patterns were difficult to discover. This raises questions as to the role of worked examples in textbooks, what they highlight as important, what learning opportunities they provide, etc. In relation to proof-related reasoning, this seems to be a field that the research so far has given little attention. Aside from worked examples, students can also be offered clues or complete solutions to exercises in answer sections. This means that one should consider including such textbook material if endeavoring to investigate what guidance a textbook offers in solving proof-related tasks.

Another finding in Papers I and IV was that the Swedish textbooks had more of a "conjecturing" approach to mathematics, and the Finnish ones a "deductive" one. There seems to be a tension between these two approaches. If one is emphasized, the other becomes less visible. This tension needs further attention. A thorough investigation of "good examples" of how to combine the two approaches in textbook expositions would likely be of value for textbook authors, as well as for teachers. This will require detailed analyses of expository sections, for instance as was done involving "embeddings" of justifications in Paper III.

The most obvious finding was that many important aspects of proof-related reasoning are virtually absent in the textbooks. On the one hand, one cannot expect upper secondary school to provide students with all the learning opportunities required for a mastery of all the items in Tables 8–10. However, if students are only acquainted with direct derivations of computational formulas, and are unfamiliar with other kinds of proof of other kinds of statements, the transition to university mathematics will be problematic. Already in introductory university courses in algebra and calculus, the student will face statements and arguments with all the various subproperties of generality, structure, logic, and methods listed in Table 10. A thorough comparison of upper secondary textbooks and university textbooks could provide valuable information regarding where student difficulties in relation to proof are most likely to occur. They could also offer advice on areas that are the most essential to focus more attention on in upper secondary school, or where university teaching needs to offer better learning support. The inclusion of a university textbook in Paper II did not provide sufficient data for drawing any conclusions in this direction.

To stay relevant, textbook research also needs to go beyond the textbook and into classroom practice. As the textbook analyses presented in this thesis analyzed textbook material as if it was lesson scripts, or documented lectures, the analytical approach should be possible to adapt for analyzing classroom teaching episodes. This means that the discussions in Section 6.2 have relevance for the design and analysis of classroom interventions and teaching.

Classroom interventions aimed at elevating reasoning and proving are relatively rare (Stylianides & Stylianides, 2017), and more research is needed on what guidance textbooks need to offer teachers in order to elevate their teaching of proof (Stylianides, 2014). A model for classroom implementation of a "rich perspective on proof" needs extensive and thorough testing. Such a model should include support for proof-related activities of various kinds: investigating patterns and formulating conjectures about general principles, as well as reading and investigating well-formulated general statements; outlining and constructing proofs, as well as reading and evaluating proofs; building mathematical structures, as well as studying the well-established structures that tradition has given mathematics. The model for classroom enactment should be reflected in the mathematics textbook and supported by other curriculum resources. The analytical frameworks used in this thesis conceptualize proof-related reasoning for the analysis of textbook material. For the development and analysis of this model, the frameworks need to be adapted for the analysis of classroom activities. Thus, what is proposed here is design research including the design of curriculum materials (textbooks), classroom enactment of curriculum (teaching), and analytical/theoretical tools (frameworks) for analyzing and understanding the relations between the written and enacted curriculum.

6.6 Concluding remarks

The aim of this thesis has been to contribute to an informed discussion about opportunities to learn proof-related reasoning offered in mathematics textbooks. This has been done by comparing and discussing empirical findings in Swedish and Finnish upper secondary textbooks, both from a Swedish-Finnish perspective and from a topic-specific perspective, with the literature on student difficulties as a background. Experiences from the analytical approaches used here have been used to discuss refinements to analytical frameworks and conceptualizations of proof-related reasoning.

Given that the various aspects of proof-related reasoning that are captured by the analytical frameworks used here are also desired learning outcomes in mathematics education, teachers and students need more support than textbooks typically offer. This thesis contributes by pointing to areas where such support can be assumed to be the most valuable, and offers specific suggestions in relation to logarithms, primitive functions, definite integrals, and combinatorics. It also makes a conceptual contribution by offering detailed lists of both general and specific analytical questions, kinds of reasoning activities, and important properties that characterize reasoning activities.

Hopefully, this thesis can offer teachers, textbook authors, and textbook analysists some aid in their great endeavor to expose the soul of mathematics (Schoenfeld, 2009).

References

Alcock, L., & Simpson, A. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 57*(1), 1–32. https://doi.org/10.1023/B:EDUC.0000047051.07646.92

Alcock, L., & Weber, K. (2010). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. In F. Hitt, D. Holton, & P. Thompson (Eds.), *Research in collegiate mathematics education. VII* (pp. 93–114). American Mathematical Society.

Almeida, D. (2001). Pupils' proof potential. *International Journal of Mathematical Education in Science and Technology, 32*(1), 53–60. https://doi.org/10.1080/00207390119535

Antonini, S., & Mariotti, M. A. (2008). Indirect proof: what is specific to this way of proving? *ZDM Mathematics Education, 40*(3), 401–412. https://doi.org/10.1007/s11858-008-0091-2

Balacheff, N. (1988a). Aspects of proof in pupils' practice of school mathematics. In D. Pimm (Ed.), *Mathematics, teachers and children* (pp. 216–235). Holdder & Stoughton.

Balacheff, N. (1988b). A study of students' proving processes at the junior high school level. In I. Wirszup & R. Streit (Eds.), *Proceedings of the second UCSMP international conference on mathematics education* (pp. 27–44). National Council of Teachers of Mathematics.

Bass, H. (2011). Proof in mathematics education: An endangered species? A review of teaching and learning proof across the grades: A K–16 perspective. *Journal for Research in Mathematics Education, 42*(1), 98–103. https://doi.org/10.5951/jresematheduc.42.1.0098

Begle, E. G. (1973). Some lessons learned by SMSG. *Mathematics Teacher, 66*(3), 207–214. https://doi.org/10.5951/MT.66.3.0207

Bell, A. W. (1976). A study of pupils' proof-explanations in mathematical situations. *Educational Studies in Mathematics, 7*(1–2), 23–40.

Bieda, K. N., Ji, X., Drwencke, J., & Picard, A. (2014). Reasoning-and-proving opportunities in elementary mathematics textbooks. *International Journal of Educational Research, 64*, 71–80. https://doi.org/10.1016/j.ijer.2013.06.005

Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2014). Developing mathematical competence: From the intended to the enacted curriculum. *The Journal of Mathematical Behavior, 33*, 72–87. https://doi.org/10.1016/j.jmathb.2013.10.001

Boyer, C. B., & Merzbach, U. C. (1991). *A history of mathematics*. John Wiley & Sons.

Cai, J., Ni, Y., & Lester, F. K. (2011). Curricular effect on the teaching and learning of mathematics: Findings from two longitudinal studies in China and the United States. *International Journal of Educational Research, 2*(50), 63–64. https://doi.org/10.1016/j.ijer.2011.06.001

Campbell, T. G., Boyle, J. D., & King, S. (2020). Proof and argumentation in K-12 mathematics: A review of conceptions, content, and support. *International Journal of Mathematical Education in Science and Technology, 51*(5), 754–774. https://doi.org/10.1080/0020739X.2019.1626503

Choppin, J. (2011). The impact of professional noticing on teachers' adaptations of challenging tasks. *Mathematical Thinking and Learning, 13*(3), 175–197. https://doi.org/10.1080/10986065.2010.495049

Coe, R., & Ruthven, K. (1994). Proof practices and constructs of advanced mathematics students. *British Educational Research Journal, 20*(1), 41–53. https://doi.org/10.1080/0141192940200105

Davis, E. A., Beyer, C., Forbes, C. T., & Stevens, S. (2011). Understanding pedagogical design capacity through teachers' narratives. *Teaching and Teacher Education, 27*(4), 797–810. https://doi.org/10.1016/j.tate.2011.01.005

de Villiers, M. D. (1990). The role and function of proof in mathematics. *Pythagoras, 24*, 17–24.

de Villiers, M. D. (1999). The role and function of proof with Sketchpad. In M. D. De Villiers (Ed.), *Rethinking Proof with Sketchpad* (pp. 3–10). Key Curriculum Press.

Dolev, S., & Even, R. (2015). Justifications and explanations in Israeli 7th grade math textbooks. *International Journal of Science and Mathematics Education, 13*(2), 309–327. https://doi.org/10.1007/s10763-013-9488-7

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. In E. Dubinsky, A. H. Schoenfeld, & J. Kaput (Eds.), *Research in collegiate mathematics education. IV* (pp. 239–289). American Mathematical Society.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics, 53*(1), 5–34. https://doi.org/10.1023/A:1024661004375

Durand-Guerrier, V., & Arsac, G. (2005). An epistemological and didactic study of a specific calculus reasoning rule. *Educational Studies in Mathematics, 60*(2), 149–172. https://doi.org/10.1007/s10649-005-5614-y

Finnish National Board of Education. (2004). *National core curriculum for upper secondary schools 2003: National core curriculum for general upper secondary education intended for young people*. Vammalan Kirjapaino Oy.

Fujita, T., & Jones, K. (2014). Reasoning-and-proving in geometry in school mathematics textbooks in Japan. *International Journal of Educational Research, 64*, 81–91. https://doi.org/10.1016/j.ijer.2013.09.014

Galbraith, P. L. (1981). Aspects of proving: A clinical investigation of process. *Educational Studies in Mathematics, 12*(1), 1–28. https://doi.org/10.1007/BF00386043

Garuti, R., Boero, P., & Lemut, E. (1998). Cognitive unity of theorems and difficulty of proof. In A. Olivier & K. Newstead (Eds.), *Proceedings of the 22th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345–352). The University of Stellenbosch, Faculty of Education and PME.

Gavin, M. K., Casa, T. M., Adelson, J. L., & Firmender, J. M. (2013). The impact of challenging geometry and measurement units on the achievement of grade 2 students. *Journal for Research in Mathematics Education, 44*(3), 478–509. https://doi.org/10.5951/jresematheduc.44.3.0478

Glasnovic Gracin, D. (2018). Requirements in mathematics textbooks: a five-dimensional analysis of textbook exercises and examples. *International Journal of Mathematical Education in Science and Technology, 49*(7), 1003–1024. https://doi.org/10.1080/0020739X.2018.1431849

Grouws, D. A., Smith, M. S., & Sztajn, P. (2000). The preparation and teaching practices of United States mathematics teachers: Grades 4 and 8. In P. Kloosterman & F. K. J. Lester (Eds.), *Results and interpretations of the 1990–2000 mathematics assessments of the National Assessment of Educational Progress* (pp. 221–267). National Council of Teachers of Mathematics.

Hanna, G. (1989). More than formal proof. *For the learning of mathematics, 9*(1), 20–23.

Hanna, G. (1990). Some pedagogical aspects of proof. *Interchange, 21*(1), 6–13. https://doi.org/10.1007/BF01809605

Hanna, G. (1995). Challenges to the importance of proof. *For the learning of mathematics, 15*(3), 42–49.

Hanna, G. (2000). A critical examination of three factors in the decline of proof. *Interchange, 31*(1), 21–33. https://doi.org/10.1023/A:1007630729662

Hanna, G., & de Bruyn, Y. (1999). Opportunity to learn proof in Ontario grade twelve mathematics texts. *Ontario Mathematics Gazette, 37*(4), 23–29.

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

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805–842). Information Age Publishing.

Healy, L., & Hoyles, C. (1998). *Justifying and proving in school mathematics: Technical report on the nationwide survey*. Institute of Education, University of London.

Healy, L., & Hoyles, C. (2000). A study of proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396–428. https://doi.org/10.2307/749651

Hemmi, K. (2006). *Approaching proof in a community of mathematical practice* [Doctoral dissertation, Stockholm University]. Stockholm, Sweden.

Hemmi, K. (2008). Students’ encounter with proof: The condition of transparency. *ZDM Mathematics Education, 40*(3), 413–426. https://doi.org/10.1007/s11858-008-0089-9

Hemmi, K., Koljonen, T., Hoelgaard, L., Ahl, L., & Ryve, A. (2013). Analyzing mathematics curriculum materials in Sweden and Finland: Developing an analytical tool. In B. Ubuz, C. Haser, & M. A. Mariotti (Eds.), *Proceedings of the eighth congress of the European Society for Research in Mathematics Education (CERME 8, February 6–10, 2013)* (pp. 1875–1884). Middle East Technical University and ERME.

Hemmi, K., Krzywacki, H., & Liljekvist, Y. (2019). Challenging traditional classroom practices: Swedish teachers’ interplay with Finnish curriculum materials. *Journal of Curriculum Studies, 51*(3), 342–361. https://doi.org/10.1080/00220272.2018.1479449

Hemmi, K., Lepik, M., & Viholainen, A. (2013, 2013/06/01). Analysing proof-related competences in Estonian, Finnish and Swedish mathematics curricula—towards a framework of developmental proof. *Journal of Curriculum Studies, 45*(3), 354–378. https://doi.org/10.1080/00220272.2012.754055

Hemmi, K., & Ryve, A. (2015a). The culture of the mathematics classroom during the first school years in Finland and Sweden. In B. Perry, A. MacDonald, & A. Gervasoni (Eds.), *Mathematics and transition to school: International perspectives* (pp. 185–198). Springer.

Hemmi, K., & Ryve, A. (2015b). Effective mathematics teaching in Finnish and Swedish teacher education discourses. *Journal of Mathematics Teacher Education, 18*(6), 501–521. https://doi.org/10.1007/s10857-014-9293-4

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics, 24*(4), 389–399. https://doi.org/10.1007/BF01273372

Hersh, R. (2009). What I would like my students to already know about proof. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 17–20). Routledge.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.

Hill, H. C., & Charalambous, C. Y. (2012). Teacher knowledge, curriculum materials, and quality of instruction: Lessons learned and open issues. *Journal of Curriculum Studies, 44*(4), 559–576. https://doi.org/10.1080/00220272.2012.716978

Hoyles, C., & Küchemann, D. (2002). Students' understandings of logical implication. *Educational Studies in Mathematics, 51*(3), 193–223. https://doi.org/10.1023/A:1023629608614

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education, 43*(4), 358–390. https://doi.org/10.5951/jresematheduc.43.4.0358

Jankvist, U. T., & Misfeldt, M. (2019). CAS assisted proofs in upper secondary school mathematics textbooks. *Redimat-Revista De Investigacion En Didactica De Las Matematicas, 8*(3), 232–266. https://doi.org/10.4471/redimat.2019.3315

Job, P., & Schneider, M. (2014). Empirical positivism, an epistemological obstacle in the learning of calculus. *ZDM Mathematics Education, 46*(4), 635–646. https://doi.org/10.1007/s11858-014-0604-0

Joutsenlahti, J., & Vainionpää, J. (2010). Oppimateriaali matematiikan opetuksessa ja osaamisessa [Learning materials in the teaching and learning of mathematics]. In E. K. Niemi & J. Metsämuuronen (Eds.), *Miten matematiikan taidot kehittyvät? Matematiikan oppimistulokset peruskoulun viidennen vuosiluokan jälkeen vuonna 2008 [How do pupils' mathematical skills develop? The learning outcomes at the end of fifth grade in compulsory school]* Opetushallitus [The Finnish National Board of Education].

Knuth, E. J., Choppin, J., Slaughter, M., & Sutherland, J. (2002). Mapping the conceptual terrain of middle school students' competencies in justifying and proving. In D. S. Mewborn, P. Sztajn, D. Y. White, H. G. Wiegel, R. L. Bryant, & K. Nooney (Eds.), *Proceedings of the 24th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (Athens, Georgia, October 26–29, 2002)* (pp. 1693–1670). ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Knutsson, M., Hemmi, K., Bergwall, A., & Ryve, A. (2013). School-based mathematics teacher education in Sweden and Finland: characterizing mentor–prospective teacher discourse. In B. Ubuz, C. Haser, & M. A. Mariotti (Eds.), *Proceedings of the eighth congress of the European Society for Research in Mathematics Education (CERME 8, February 6–10, 2013)* (pp. 1905–1914). Middle East Technical University and ERME.

Ko, Y.-Y., & Knuth, E. J. (2013). Validating proofs and counterexamples across content domains: Practices of importance for mathematics majors. *The Journal of Mathematical Behavior, 32*(1), 20–35. https://doi.org/10.1016/j.jmathb.2012.09.003

Lakatos, I. (1976). *Proofs and refutations*. Cambridge university press.

Li, Y. (2000). A comparison of problems that follow selected content presentations in American and Chinese mathematics textbooks. *Journal for Research in Mathematics Education, 31*(2), 234–241. https://doi.org/10.2307/749754

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41–51. https://doi.org/10.5951/jresematheduc.20.1.0041

Matematikcentrum. (2010). *Övningar i Analys i en variabel*. Studentlitteratur.

Matematikcentrum. (2011). *Övningar i endimensionell analys*. Studentlitteratur.

McCrory, R., & Stylianides, A. J. (2014). Reasoning-and-proving in mathematics textbooks for prospective elementary teachers. *International Journal of Educational Research, 64*, 119–131. https://doi.org/10.1016/j.ijer.2013.09.003

Moreno-Armella, L. (2014). An essential tension in mathematics education. *ZDM Mathematics Education, 46*(4), 621–633. https://doi.org/10.1007/s11858-014-0580-4

Morris, A. K. (2002). Mathematical reasoning: Adults' ability to make the inductive-deductive distinction. *Cognition and Instruction, 20*(1), 79–118. https://doi.org/10.1207/S1532690XCI2001_4

Movshovitz-Hadar, N. (1988). Stimulating presentation of theorems followed by responsive proofs. *For the learning of mathematics, 8*(2), 12–30.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International results in mathematics*. TIMSS & PIRLS International Study Center.

Mwadzaangati, L. (2019). Comparison of geometric proof development tasks as set up in the textbook and as implemented by teachers in the classroom. *Pythagoras, 40*(1), 1–14. https://doi.org/10.4102/pythagoras.v40i1.458

Månsson, J., & Nordbeck, P. (2011). *Endimensionell analys*. Studentlitteratur.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.

NCTM. (2000). *Principles and standards for school mathematics*. National Council of Teachers of Mathematics.

Ortiz, C. V., & Pastells, A. A. (2017). Proposiciones, procedimientos y argumentos sobre probabilidad en libros de texto chilenosde educación primaria [Mathematical objects in chilean textbooks: propositions, procedures and arguments about probability]. *Profesorado-Revista De Curriculum Y Formacion De Profesorado, 21*(1), 433–457. https://doi.org/10.30827/profesorado.v21i1.10373

Otten, S., Gilbertson, N. J., Males, L. M., & Clark, D. L. (2014). The mathematical nature of reasoning-and-proving opportunities in geometry textbooks. *Mathematical Thinking and Learning, 16*(1), 51–79. https://doi.org/10.1080/10986065.2014.857802

Otten, S., Males, L. M., & Gilbertson, N. J. (2014). The introduction of proof in secondary geometry textbooks. *International Journal of Educational Research, 64*, 107–118. https://doi.org/10.1016/j.ijer.2013.08.006

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics, 66*(1), 23–41. https://doi.org/10.1007/s10649-006-9057-x

Pepin, B., Gueudet, G., & Trouche, L. (2013). Investigating textbooks as crucial interfaces between culture, policy and teacher curricular practice: two contrasted case studies in France and Norway. *ZDM Mathematics Education, 45*(5), 685–698. https://doi.org/s11858-013-0526-2

Pepin, B., & Haggarty, L. (2001). Mathematics textbooks and their use in English, French and German classrooms: a way to understand teaching andlearning cultures. *Zentralblatt für Didaktik der Mathematik, 33*(5), 158–175. https://doi.org/10.1007/bf02656616

Persson, A., & Böiers, L.-C. (2010). *Analys i en variabel*. Studentlitteratur.

Pólya, G. (1954). *Mathematics and plausible reasoning: Induction and analogy in mathematics* (Vol. 1). Princeton University Press.

Porteous, K. (1990). What do children really believe? *Educational Studies in Mathematics, 21*(6), 589–598. https://doi.org/10.1007/BF00315946

Roh, K. H. (2008). Students' images and their understanding of definitions of the limit of a sequence. *Educational Studies in Mathematics, 69*(3), 217–233. https://doi.org/10.1007/s10649-008-9128-2

Ryve, A., Hemmi, K., & Börjesson, M. (2011). Discourses about school-based mathematics teacher education in Finland and Sweden. *Scandinavian Journal of Educational Research*, 1–16. https://doi.org/10.1080/00313831.2011.623178

Schoenfeld, A. H. (2009). Series editor's foreword: The soul of mathematics. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. xii–xvi). Routledge.

Selden, A., & Selden, J. (2003). Validations of proofs considered as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education, 34*(1), 4–36. https://doi.org/10.2307/30034698

Selden, A., & Selden, J. (2013). Proof and problem solving at university level. *The Mathematics Enthusiast, 10*(1), 303–334.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics, 29*(2), 123–151. https://doi.org/10.1007/BF01274210

Senk, S. L. (1985). How well do students write geometry proofs? *Mathematics Teacher, 78*(6), 448–456. https://doi.org/10.5951/MT.78.6.0448

Senk, S. L. (1989). Van Hiele levels and achievement in writing geometry proofs. *Journal for Research in Mathematics Education, 20*(3), 309–321. https://doi.org/10.5951/jresematheduc.20.3.0309

Sevimli, E. (2018). Undergraduates' propositional knowledge and proof schemes regarding differentiability and integrability concepts. *International Journal of Mathematical Education in Science and Technology, 49*(7), 1052–1068. https://doi.org/10.1080/0020739X.2018.1430384

Skolinspektionen. (2009). *Undervisningen i matematik—Utbildningens innehåll och ändamålsenlighet [The teaching of mathematics—The content and the effectiveness of the education]*. Skolinspektionen [The Swedish Schools Inspectorate].

Skolverket. (2011). *Läroplan, examensmål och gymnasiegemensamma ämnen för gymnasieskola 2011 [Curriculum, exam objectives, and upper secondary school subjects 2011]*. Skolverket [Swedish National Agency for Education].

Skolverket. (2012). *TIMSS 2011: Svenska grundskoleelevers kunskaper i matematik och naturvetenskap i ett internationellt perspektiv [TIMSS 2011: Swedish elementary school students' knowledge in mathematics and science in an international perspective]*. Fritzes.

Sowder, L., & Harel, G. (2003). Case studies of mathematics majors' proof understanding, production, and appreciation. *Canadian Journal of Math, Science & Technology Education, 3*(2), 251–267. https://doi.org/10.1080/14926150309556563

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

Stacey, K., & Vincent, J. (2009). Modes of reasoning in explanations in Australian eighth-grade mathematics textbooks. *Educational Studies in Mathematics, 72*(3), 271–288. https://doi.org/10.1007/s10649-009-9193-1

Staples, M., Newton, J., Kosko, K., Conner, A., Cirillo, M., & Bieda, K. (2016). Conceptions and consequences of what we call argumentation, justification, and proof. In M. B. Wood, E. E. Turner, M. Civil, & J. A. Eli (Eds.), *Proceedings of the 38th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 1704–1712). The University of Arizona.

Staples, M. E., Bartlo, J., & Thanheiser, E. (2012). Justification as a teaching and learning practice: Its (potential) multifacted role in middle grades mathematics classrooms. *The Journal of Mathematical Behavior, 31*(4), 447–462. https://doi.org/10.1016/j.jmathb.2012.07.001

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). Information Age Publishing.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for Research in Mathematics Education, 38*(3), 289–321. https://doi.org/10.2307/30034869

Stylianides, A. J., & Ball, D. L. (2008). Understanding and describing mathematical knowledge for teaching: Knowledge about proof for engaging students in the activity of proving. *Journal of Mathematics Teacher Education, 11*(4), 307–332. https://doi.org/10.1007/s10857-008-9077-9

Stylianides, A. J., Bieda, K. N., & Morselli, F. (2016). Proof and argumentation in mathematics education research. In A. Gutiérrez, G. C. Leder, & P. Boero (Eds.), *The second handbook of research on the psychology of mathematics education* (pp. 315–351). Sense Publishers.

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics, 55*(1), 133–162. https://doi.org/10.1023/B:EDUC.0000017671.47700.0b

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the learning of mathematics, 28*(1), 9–16.

Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*(4), 258–288. https://doi.org/10.1080/10986060903253954

Stylianides, G. J. (2014). Textbook analyses on reasoning-and-proving: Significance and methodological challenges. *International Journal of Educational Research, 64*, 63–70. https://doi.org/10.1016/j.ijer.2014.01.002

Stylianides, G. J., & Stylianides, A. J. (2017). Research-based interventions in the area of proof: the past, the present, and the future. *Educational Studies in Mathematics, 96*(2), 119–127. https://doi.org/10.1007/s10649-017-9782-3

Stylianides, G. J., Stylianides, A. J., & Philippou, G. N. (2007). Preservice teachers' knowledge of proof by mathematical induction. *Journal of Mathematics Teacher Education, 10*(3), 145–166. https://doi.org/10.1007/s10857-007-9034-z

Stylianides, G. J., Stylianides, A. J., & Weber, K. (2017). Research on the teaching and learning of proof: Taking stock and moving forward. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 237–236). The National Council of Teachers in Mathematics

Stylianou, D. A., Blanton, M. L., & Knuth, E. J. (Eds.). (2009). *Teaching and learning proof across the grades: A K-16 perspective*. Routledge.

Tarr, J. E., Reys, R. E., Reys, B. J., Chávez, Ó., Shih, J., & Osterlind, S. J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education, 39*(3), 247–280. https://doi.org/10.2307/30034970

Thompson, D. R. (1991, April). *Reasoning and proof in precalculus and discrete mathematics*. Paper presented at the meeting of the American Educational Research Association, Chicago.

Thompson, D. R., Senk, S. L., & Johnson, G. J. (2012). Opportunities to learn reasoning and proof in high school mathematics textbooks. *Journal for Research in Mathematics Education, 43*(3), 253–295. https://doi.org/10.5951/jresematheduc.43.3.0253

Tinto, P. P. (1990). *Students' views on learning proof in high school geometry: An analytic-inductive approach* [Doctoral dissertation, Syracuse University]. Syracuse, NY.

Toulmin, S. E. (2003). *The uses of argument*. Cambridge university press.

Törner, G., Potari, D., & Zachariades, T. (2014). Calculus in European classrooms: curriculum and teaching in different educational and cultural contexts. *ZDM Mathematics Education, 46*(4), 549–560. https://doi.org/10.1007/s11858-014-0612-0

Valverde, G. A., Bianchi, L. J., & Wolfe, R. G. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks*. Kluwer.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics, 48*(1), 101–119. https://doi.org/10.1023/A:1015535614355

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning, 12*(4), 306–336. https://doi.org/10.1080/10986065.2010.495468

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics, 56*(2), 209–234. https://doi.org/10.1023/B:EDUC.0000040410.57253.a1

Weber, K., & Alcock, L. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D. A. Stylianou, M. L. Blanton, & E. J. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 323–338). Routledge.

Weinberg, A., & Wiesner, E. (2011). Understanding mathematics textbooks through reader-oriented theory. *Educational Studies in Mathematics, 76*(1), 49–63. https://doi.org/10.1007/s10649-010-9264-3

Vinner, S. (1977). The concept of exponentiation at the undergraduate level and the definitional approach. *Educational Studies in Mathematics*, 17–26.

Wood, T. (1999). Creating a context for argument in mathematics class. *Journal for Research in Mathematics Education*, 171–191. https://doi.org/10.2307/749609

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Zandieh, M., Roh, K. H., & Knapp, J. (2014). Conceptual blending: Student reasoning when proving "conditional implies conditional" statements. *The Journal of Mathematical Behavior, 33*, 209–229. https://doi.org/10.1016/j.jmathb.2013.11.007

Zhang, D., & Qi, C. X. (2019). Reasoning and proof in eighth-grade mathematics textbooks in China. *International Journal of Educational Research, 98*, 77–90. https://doi.org/10.1016/j.ijer.2019.08.015

Appendix

The analysis of the textbook material requires (of course) more than an acquaintance with the mathematical content; and the discussion in Paper V even more so. A short summary of the most important details is included here, with notation matching that of the analyzed textbooks. For a more comprehensive account, any university textbook on algebra or calculus will do.

A.1 Logarithms

The *logarithm of a to the base b,* or the b-*logarithm of* a, is denoted $\log_b a$ and is defined implicitly as the solution x to the equation $b^x = a$; i.e., $b^{\log_b a} = a$. Important special cases are logarithms for base 10 (denoted $\lg a$) and the *natural logarithm* $\ln a$ where the base is the Euler number $e = 2{,}718 \ldots$. This is the only base for which the derivative of b^x is b^x.

For the definition to make sense, the equation $b^x = a$ must be uniquely solvable. This is the case if $b > 0$, $b \neq 1$ and $a > 0$. More precisely, the exponential function b^x is strictly increasing if $b > 1$ and strictly decreasing if $0 < b < 1$. Strict monotonicity implies uniqueness. Further, with these restrictions on b, the asymptotic behavior of b^x combined with the continuity of b^x guarantees that b^x attains every real value $a > 0$. This guarantees solvability.

The definition of logarithms makes them a tool for solving exponential equations and problems involving exponential growth. Logarithm laws (i.e. calculation rules) for products, quotients, and powers follow from corresponding power laws. For instance, the logarithm product law $\ln(a \cdot b) = \ln a + \ln b$ follows from the power law $e^x \cdot e^y = e^{x+y}$:

$$\ln(a \cdot b) = \ln\left(e^{\ln a} \cdot e^{\ln b}\right) = \ln e^{\ln a + \ln b} = \ln a + \ln b$$

Proofs of other laws are very similar (replace $\cdot$ and $+$ with / and – and one gets a proof of the logarithm quotient law), as are the corresponding laws for other bases.

A.2 Primitive functions

$F(x)$ is a *primitive function* to $f(x)$ in an open interval (a, b) if $F'(x) = f(x)$ for all x in (a, b). The derivative $F'(x)$ of F at x is the limit of

$\big(F(x+h)-F(x)\big)/h$ when $h \to 0$, if it exists and is finite. While the derivative is defined pointwise, being a primitive function is a property connected to an interval and expresses a relation between functions. Primitive functions are usually denoted with capital letters. They are also referred to as *anti-derivatives* or *indefinite integrals* and denoted $\int f(x)\,dx$. Primitive functions are important tools for computing definite integrals (see next subsection) and for solving differential equations.

Primitive functions (when defined on intervals) are uniquely determined up to an additive constant; that is, $G(x)$ is a primitive function to $F'(x)$ if and only if $G(x) = F(x) + C$ where C is a constant. The sufficiency is trivial, and the necessity follows from the fact that a function whose derivative vanishes in an interval must be constant there. This, in turn, can be proven using the mean value theorem. A classical existence result states that every function f that is continuous in an interval (a, b) has a primitive function in that interval. This is a consequence of the first fundamental theorem of calculus (see below).

As logarithms inherit properties from exponentials, so do primitive functions from derivatives. Two important classes of results are elementary primitive functions (e.g., $\int \cos(x)\,dx = \sin(x) + C$ since $(\sin x)' = \cos x$) and general rules of integration (e.g., $\int\big(f(x)+g(x)\big)dx = \int f(x)dx + \int g(x)dx$ since $(F+G)'(x) = F'(x) + G'(x)$).

A.3 Definite integrals

The definition of the *definite integral*

$$I = \int_a^b f(x)dx,$$

or more precisely the *Riemann integral*, is theoretically much more complicated than that of indefinite integrals. It is usually defined in one of the two following (and equivalent) ways:

- the limit I of Riemann sums $\sum_{k=1}^{n} f(x_k)\Delta x_k$ when the partition of the interval $[a, b]$ is infinitely refined (a process which in itself is quite difficult to formalize)
- the number I which is a lower bound to all upper sums of f on $[a, b]$ and an upper bound to all lower sums of f on $[a, b]$

In both cases, the number I must exist and be unique for the definition to make sense. Riemann was the first to prove (in the 1850s) that continuity of f is sufficient for this. The standard proof makes use of uniform continuity. A formal treatment also requires knowledge of limits and the completeness of the real field. Geometrically, the integral represents the area under the graph of f, and is hence a tool for solving problems that can be represented by such an area (e.g., distance is the area under the velocity graph, energy is the area under the power graph).

There are two fundamental results that connect definite integrals with derivatives. Newton and Leibniz discovered these (independently) in the late 17th century. *The first fundamental theorem of calculus* states that if f is continuous, then $\int_a^b f(x)dx$, if viewed as a function of the upper limit b, has the derivative $f(b)$. The standard proof is based on the mean value theorem of integral calculus. As a trivial corollary, continuous functions always have primitive functions. Another corollary is *the second fundamental theorem of calculus*: If F is a primitive function to a continuous function f then $\int_a^b f(x)dx = F(b) - F(a)$. The proof is rather simple. The representation formula for primitive functions guaranties that if F is a primitive function to f, then $\int_a^b f(x)dx = F(b) + C$ for some constant C. Putting $b = a$ in this identity yields $C = -F(a)$.

The second fundamental theorem of calculus implies that integral calculations boil down to finding primitive functions. Hence, even though the theoretical foundations of definite integrals are considered too theoretically advanced for upper secondary mathematics education, the mathematics results can be presented, understood, and applied to geometric and scientific problems.

A.4 Combinatorics

Combinatorics is not a specific concept but rather the branch of mathematics in which the possibilities to choose and arrange elements in a set are studied. At its base lies combinatorial representations of addition and multiplication. *The addition principle* states that when choosing an element from a set A with n elements or from a set B with m elements, there are $m + n$ possible choices. This principle is used to divide the set of possible choices into disjoint subsets that can be handled separately. *The multiplication principle* states that when two elements are chosen independently of each other, one from A and one from B, then the number of possible pairs of elements are $m \cdot n$. This principle is typically used for combinatorial problems involving sequences of choices.

From the multiplication principle, formulas for *permutations* and *combinations* can be derived. A permutation is an arrangement of the elements in a set. If the set contains *n* elements the first element of a permutation can be chosen in *n* different ways, the second in *n-1* different ways, and so on. Hence, the multiplication principle implies that the total number of permutations is $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 2 \cdot 1$. The number of *$k$-permutations* (i.e., permutations of k elements chosen from the set) is $P(n,k) = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = n!/(n-k)!$. A *$k$-combination* of a set with n elements is a subset with k elements. As the elements of such a subset can be arranged in $k!$ different ways, there will only be

$$C(n,k) = \frac{n!}{(n-k)!\,k!}$$

different k-combinations. This number is often denoted with the symbol $\binom{n}{k}$, which is read "n choose k".

Combinations also occur as *binomial coefficients*; according to the binomial theorem,

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Combinations satisfy the identity

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k},$$

which can be represented in an easily memorized triangular scheme called the *Pascal triangle*, in which $\binom{n-1}{k-1}$ is the kth entry of the nth row. There is also a symmetry property that represents the idea that picking k elements out of n is the same as choosing which $n-k$ elements *not* to pick: $\binom{n}{k} = \binom{n}{n-k}$.

Related to choosing and arranging subsets is *the Dirichlet box principle* (also called *the pigeonhole principle*), which states that if a set with $n+1$ elements is divided into n disjoint subsets, at least one subset will contain more than one element. More generally, if a set with more than nk elements is divided into k disjoint subsets, at least one will contain more than n elements.

Finally, combinatorics is an important tool in elementary probability theory. The classical definition of probability is the ratio of the number of favorable cases to the number of possible cases. Combinatorics provides a means to calculate these numbers of cases.

A.5 Logical inference

A mathematical statement P is either true or false. The negation of P is denoted $\neg P$, and is true when P is false and vice versa. Usually, mathematical statements involve variables, and the values of the variables affect the truth value of the statement. A *universal* statement is true for all values of the variable, whereas an *existential* statement is true for at least one value of each variable and a *uniqueness* statement is true for at most one value of each variable.

A mathematical *implication,* $P \Rightarrow Q$, is a conditional statement of the form "if P, then Q", where P and Q represent statements that can be true or false, but where Q always is true if P is true. The implication $P \Rightarrow Q$ is equivalent to its *contrapositive* $\neg Q \Rightarrow \neg P$, but not to its *inverse* $\neg P \Rightarrow \neg Q$ or its *converse* $Q \Rightarrow P$. In a *direct proof* of an implication $P \Rightarrow Q$, one assumes that P is true and argues for how this must lead to Q also being true. The argument must not build on any assumptions about the truth value of Q. A *contrapositive proof*, on the other hand, assumes nothing about P. Instead, the starting point is that Q is false, and then the proof outlines how this by necessity must lead to P also being false. A *proof by contradiction* assumes P to be true and Q to be false, and outlines how this leads to a contradiction.

If $P \Rightarrow Q$and $Q \Rightarrow P$, then P and Q are *equivalent*, which is written $P \Leftrightarrow Q$. An equivalence is often proven by proving the two implications $P \Rightarrow Q$ and $Q \Rightarrow P$ one at a time.

When an implication $P \Rightarrow Q$ is known, and P is known to be true, one can conclude that Q is true. This is the most common mode of deductive reasoning, referred to as *modus ponens*. If $P \Rightarrow Q$ is true and P is false, no conclusion can be drawn regarding Q's truth value. But if $P \Rightarrow Q$ is true and Q is false, one can conclude that P must be false. This mode of reasoning is called *modus tollens*.