

## The development of algebraic proficiency

**Citation for published version (APA):**

Stiphout, van, I. M. (2011). *The development of algebraic proficiency*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Eindhoven School of Education]. Technische Universiteit Eindhoven.  
<https://doi.org/10.6100/IR719774>

**DOI:**

[10.6100/IR719774](https://doi.org/10.6100/IR719774)

**Document status and date:**

Published: 01/01/2011

**Document Version:**

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

**Please check the document version of this publication:**

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

**Take down policy**

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

# The development of algebraic proficiency

Irene van Stiphout

To Anton van Stiphout and Peet van Nimwegen

A catalogue record is available from the Eindhoven University of Technology  
Library.

ISBN: 978-90-386-2979-7

NUR: 846

Printed by Printservice TU/e.

Cover by Oranje Vormgevers.

This thesis was typeset with L<sup>A</sup>T<sub>E</sub>X.

© 2011 Irene van Stiphout.

# The development of algebraic proficiency

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de  
Technische Universiteit Eindhoven, op gezag van de  
rector magnificus, prof.dr.ir. C.J. van Duijn, voor een  
commissie aangewezen door het College voor  
Promoties in het openbaar te verdedigen  
op woensdag 14 december 2011 om 16.00 uur

door

Irene Martine van Stiphout

geboren te Breda

Dit proefschrift is goedgekeurd door de promotoren:

prof.dr. K.P.E. Gravemeijer  
en  
prof.dr. W.M.G. Jochems

Copromotor:  
dr. P.H.M. Drijvers



This doctoral thesis was financially supported by Fontys University of Applied Sciences and facilitated by the Eindhoven School of Education.  
The research was carried out in the context of the Dutch Interuniversity Centre for Educational Research.

# Contents

<b>Dankwoord</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Dutch discussion of mathematics education . . . . .	1
1.2 The Dutch educational system . . . . .	2
1.3 Research aims . . . . .	4
1.4 Research strategy . . . . .	4
1.5 Outline of the thesis . . . . .	5
<b>2 The development of students' algebraic proficiency in Dutch pre-university education</b>	<b>7</b>
2.1 Introduction . . . . .	7
2.2 Different aspects of algebraic proficiency . . . . .	8
2.2.1 The relation between procedural fluency and conceptual understanding . . . . .	8
2.2.2 Transition from arithmetic to algebra . . . . .	10
2.2.3 Research questions . . . . .	10
2.3 Methods . . . . .	11
2.3.1 Test design . . . . .	11
2.3.2 Data collection . . . . .	12
2.3.3 Creating Rasch scales for algebraic proficiency . . . . .	13
2.3.4 Fit . . . . .	14
2.4 Cross-sectional and longitudinal results . . . . .	15
2.4.1 Cross-sectional results . . . . .	15
2.4.2 Longitudinal results . . . . .	17
2.5 Results on individual items . . . . .	19
2.5.1 Algebra . . . . .	19
2.5.2 Arithmetic . . . . .	22
2.5.3 Comparing students of the social and science streams . . . . .	24
2.6 Conclusions and discussion . . . . .	25
2.6.1 Main findings . . . . .	25
2.6.2 Discussion . . . . .	26

<b>3</b>	<b>Structure sense as an aspect of algebraic proficiency</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	Structure sense . . . . .	30
3.3	Student performance in structure sense . . . . .	31
	3.3.1 Algebraic tasks . . . . .	34
	3.3.2 Numerical tasks . . . . .	39
3.4	Conclusions and discussion . . . . .	41
<b>4</b>	<b>Operational–structural transitions as indicators of difficulty in secondary school algebra tasks</b>	<b>43</b>
4.1	Introduction . . . . .	43
4.2	Research context . . . . .	44
4.3	Sfard’s Theory of Reification . . . . .	45
4.4	Operational and structural conceptions . . . . .	47
4.5	Indicators of complexity . . . . .	48
	4.5.1 Transition related to the process–product duality . . . . .	49
	4.5.2 Transition related to the fixed-value–variable duality . . . . .	50
	4.5.3 Operational–structural transitions as indicators of difficulty . . . . .	51
4.6	Testing the hypothesis . . . . .	51
	4.6.1 Test construction and data collection . . . . .	52
	4.6.2 Results . . . . .	53
4.7	Conclusion and discussion . . . . .	57
<b>5</b>	<b>Instructional efficiency as a measure for evaluating learning outcomes: some limitations</b>	<b>59</b>
5.1	Introduction . . . . .	59
5.2	Main tenets of CLT . . . . .	60
5.3	Instructional efficiency . . . . .	62
5.4	Research context . . . . .	64
5.5	First limitation . . . . .	65
5.6	Second limitation . . . . .	65
5.7	Third limitation . . . . .	68
5.8	Conclusion . . . . .	70
<b>6</b>	<b>Contexts and models in Dutch textbook series: the case of linear relations and linear equations</b>	<b>73</b>
6.1	Introduction . . . . .	73
6.2	Frame of reference . . . . .	75
	6.2.1 Emergent modeling . . . . .	75
	6.2.2 Linear relations and linear equations from an emergent modeling perspective . . . . .	76
6.3	Linear relations . . . . .	78
	6.3.1 Method . . . . .	78
	6.3.2 Analysis . . . . .	80

<i>CONTENTS</i>	vii
6.3.3 Findings . . . . .	85
6.4 Linear equations . . . . .	86
6.4.1 Method . . . . .	86
6.4.2 Analysis . . . . .	87
6.4.3 Findings . . . . .	91
6.5 Conclusion and discussion . . . . .	94
<b>7 Conclusions</b>	<b>97</b>
7.1 Main findings . . . . .	97
7.1.1 The development of algebraic proficiency . . . . .	97
7.1.2 Analysis of Dutch textbook series . . . . .	103
7.1.3 Overall findings . . . . .	106
7.2 Discussion . . . . .	107
7.2.1 Limitations . . . . .	107
7.2.2 Reflection . . . . .	108
<b>Summary</b>	<b>131</b>
<b>List of publications</b>	<b>137</b>
<b>Curriculum Vitae</b>	<b>139</b>
<b>ESoE dissertation series</b>	<b>141</b>





# Dankwoord

Er zijn veel mensen die een belangrijke rol hebben gespeeld in en rond het onderzoek dat in dit proefschrift wordt beschreven. Daarvoor ben ik ze enorm dankbaar. De volgende personen en organisaties wil ik in het bijzonder noemen.

Allereerst Fontys bedankt voor de gelegenheid om aan dit onderzoek te werken. Ik heb het fantastisch gevonden om de kans te krijgen me zo te verdiepen in een onderwerp dat me na aan het hart ligt.

Johan van der Sanden wil ik bedanken omdat hij me heeft gestimuleerd een promotieonderzoek te gaan doen, en een eerste voorstel daarvoor heeft geschreven. Na zijn overlijden kwam ik bij Wim Jochems terecht, die ervoor heeft gezorgd dat ik bij Fontys in dienst kon. Bedankt Wim dat je me deze kans hebt gegeven.

Koeno, bedankt voor je begeleiding en het vertrouwen. Ik heb veel van je geleerd en denk met veel plezier terug aan de discussies en gesprekken die we hebben gehad.

Paul, bedankt dat je halverwege 2009 de dagelijkse begeleiding op je hebt willen nemen. Tijdens heel wat gezellige bijeenkomsten in Nijmegen en Utrecht ben je een sterke gesprekspartner geweest en heb je veel en goed commentaar geleverd op mijn werk.

Docenten van het Merewade College, Lorentz Casimir Lyceum, Nassau scholengemeenschap, RSG Slingerbos, Sondervick College, Christiaan Huygens College en Stedelijk College Eindhoven die aan het onderzoek hebben meegewerkt bedankt voor de tijd en moeite die jullie er in hebben gestoken. Speciaal de contactpersonen van de scholen bedankt die hebben geholpen bij de dataverzameling: Henri de Keyzer, Bas Friesen, Peter van Alem, Annemarie Glas, Frans van Scherpenzeel, Marcel Laarhoven, Annemieke Vennix, Marjolein Seegers, Marianne Lambriex en Marieke Thijssen. De leerlingen bedank ik voor de moeite die ze hebben genomen om de vragen in de toetsen te beantwoorden. De schriftelijke uitwerkingen die zij hebben geleverd vormen de basis van het onderzoek. Suzanne en Lotte bedankt voor het maken van de eerste versies van de toetsen. De namen op de omslag van dit proefschrift symboliseren de leerlingen die hebben meegedaan en over wiens prestaties het onderzoek gaat.

De collega-promovendi van FLOT Femke, Ellen, Mariska, Maud en Alexan-

der bedankt voor het delen van de zoektocht van het promoveren. De collega's van FLOT en ESoE, Martin, Gijs, Zeger-Jan en Geeke, bedankt voor de plezierige samenwerking en het delen van ervaringen. De collega's van de ESoE bedankt voor de samenwerking en de belangstelling. Speciaal dank aan Zeger-Jan, Evelien, Jannet en Marieke voor de vrolijkheid en gezelligheid in TR 3.32.

Geeke bedankt voor alles wat we de laatste jaren samen hebben gedaan. Ik ervaar het als heel bijzonder dat we over veel van het wiskundeonderwijs dezelfde opvattingen hebben. Ik hoop van harte dat we in de toekomst samen aan de verbetering van het wiskundeonderwijs kunnen werken. We hebben in ieder geval al plannen genoeg. Fijn dat je mijn paranimf wilt zijn.

De capaciteitsgroep wiskunde van de TU/e bedankt voor de open contacten, de belangstelling en de mogelijkheid de resultaten van het onderzoek te delen.

Audrey Debije Popson bedankt voor het corrigeren van het Engels onder grote tijdsdruk.

Antoon en Gertrudis, Jos en Saskia, Toin en Margot: bedankt voor de (onverwachte) logeerpartijen, het medeleven, en het begrip voor mijn afwezigheid op familie-uitjes.

Désirée, Theo, Suzanne en Inge bedankt voor jullie hulp. De manier waarop jullie hebben meegedacht en meegezocht heeft veel voor me betekend. Ik ben blij, Désirée, dat jij mijn paranimf wilt zijn.

Mijn moeder wil ik ontzettend bedanken voor het geloof en vertrouwen dat ze in me heeft gehad. Mama, je bent op veel verschillende manieren nauw bij mijn onderzoek betrokken geweest. Niet alleen heb je praktische dingen geregeld, ook hebben we veel gediscussieerd over het reken- en wiskundeonderwijs.

Dit proefschrift is opgedragen aan mijn vader en mijn buurman. Zij hebben met veel liefde, interesse en vertrouwen mijn leven voor een deel gevolgd. Als voormalig wiskundeleraar zou mijn vader mijn onderzoek zeker bijzonder interessant gevonden hebben. Jammer dat hij het niet heeft mogen meemaken. Peet van Nimwegen was een bijzondere buurman die op hoge leeftijd met enorme betrokkenheid mijn onderzoek heeft gevolgd. Het is een voorrecht hem gekend te hebben.

Mijn kinderen Isabelle, Christiaan en Martine: jullie zijn de afgelopen jaren een voortdurende bron van inspiratie en eindeloos proefkonijn voor me geweest. Speciaal bedankt voor jullie lieve aanmoedigingen in de laatste maanden van dit onderzoek en voor jullie begrip als ik ergens niet bij kon zijn. Jullie veerkracht heeft me de moed gegeven om door te gaan, ook toen het moeilijk werd.

Roel, je hebt me ongelooflijk veel geholpen met mijn proefschrift. Jouw geloof in mij en de onvoorwaardelijke steun die je me hebt gegeven zijn van cruciaal belang geweest. Met jou kan ik verzuchten: samen zijn we er gekomen!

Irene van Stiphout

# Chapter 1

## Introduction

For years, the student level of algebraic proficiency has had the interest of teachers, educational researchers, and politicians. The discussion of this issue in the Netherlands led to a debate. This debate served as the background for the research described in this thesis. In the current chapter, we give an overview of the debate and discuss the aims of the research described in this thesis.

### 1.1 The Dutch discussion of mathematics education

In the Netherlands, the discussion of arithmetic and algebra proficiency took place against the background of curriculum changes in lower and higher secondary education. In 1992, an educational change took place in lower secondary education, called *Basic Education* (in Dutch: Basisvorming). This change involved the introduction of a broad curriculum with 15 subjects for students of ages 12–14.

Drijvers (2006b, p. 57) points to a number of features that stand out in this change. We may summarize these as follows. Much attention is paid to meaningful contexts and to modeling. Students have to translate problems into algebra, or make calculations with the model given. The results of the algebraic manipulation have to be translated back into the context. In this way, students have to switch between the world of concrete problems and the world of algebra. Central in the curriculum of lower secondary education are relations and functions that require students to switch between representations such as graphs, formulas, and tables. Furthermore, students are stimulated to develop their own situational, informal and preformal strategies. Little attention is being paid to developing procedural routines. Furthermore, students are used to working independently a

great deal in both the classroom and at home. In this tradition, textbook series are tailored to avoid cognitive conflicts. By means of many subquestions in the exercises, students are guided to the intended procedures and insights.

In higher secondary education, ages 15–18, the curriculum changed in 1998 with the introduction of the so-called *Second Phase* (in Dutch: Tweede Fase) in higher secondary education. This change aimed at a better balance between knowledge, insight, and skills; and a broader curriculum for all students consisting of four profiles that do justice to students' different capabilities and interests (Tweede Fase Adviespunt, 2005b). These four profiles, Culture & Society, Economy & Society, Nature & Health, Nature & Science, all have different mathematical programs, tailored to the specific needs of the profile and continuing higher education. Beside Algebra, mathematics programs also included mathematical analysis, geometry, and statistics. Soon after the introduction of these programs, secondary school teachers were dissatisfied because of the disappointing level of student algebraic proficiency and because of the students' disappointing performance on the final exam (Boon et al., 2002; Zwaneveld, 2004). Educators in higher education complained about the algebraic skill level of first-year students (Van Gastel et al., 2007).

University teachers indicated that students in higher science, technical, and economics courses were not able to do simple mathematical calculations (e.g. Tweede Fase Adviespunt, 2005a). The heated discussion that followed concerned the whole curriculum from primary education to higher education. The polarization of the debate in newspapers and specialist journals resulted in two opposing camps: those who advocate reform mathematics, in the Netherlands influenced by "Realistic Mathematics Education" (RME), and those who advocate a more traditional approach to teaching mathematics. The empirical data used in the debate came from large scale assessments, both nationally (Janssen et al., 2005) and internationally, such as PISA (OECD, 2007) and TIMSS (Olson et al., 2008). One camps used these studies to claim that Dutch students perform relatively well, and the other, to claim that the student level of proficiency decreased. This discussion was one of the reasons for the Dutch government to adapt the *Second Phase* in 2007. This adoption involved, among other things, a more prominent place for algebraic skills in the curriculum. The research project of thesis started during this change and in the middle of the discussion of the student level of algebraic proficiency.

## 1.2 The Dutch educational system

Education in the Netherlands is compulsory for students from age 5 onwards. Depending on their basic qualification, students may quit school at ages varying from 16 to 18. However, most students start to go to school just after their fourth birthday and begin grade 1 of primary education. After eight years, students leave primary education. Secondary education is divided into three dif-

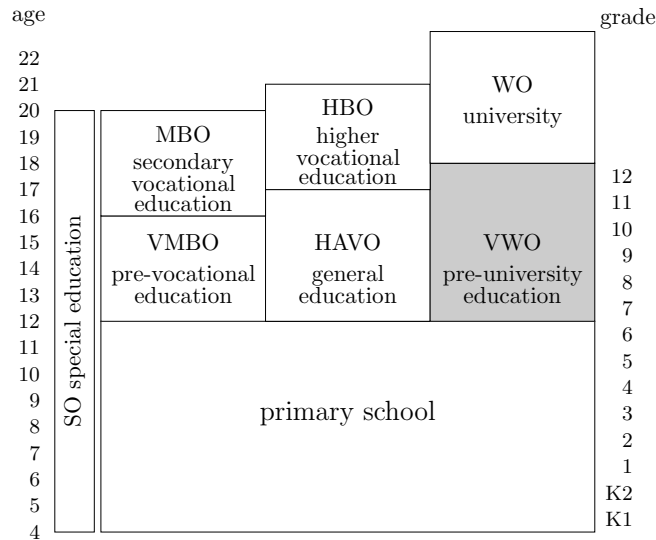


Figure 1.1: The Dutch educational system.

ferent streams: pre-vocational education, general education, and pre-university education, see Figure 1.1. About 55%–60% of the students that leave primary education go to pre-vocational education (VMBO) (Ministerie van Onderwijs, Cultuur en Wetenschap, 2010), which prepares them for vocational education. About 25% of the students that leave primary education go to general education (HAVO), which prepares them for higher vocational education. About 15%–20% of the students go to pre-university education (VWO) which prepares for university. In this thesis, we focus on the development of algebraic proficiency of students in pre-university education, the gray rectangle in Figure 1.1. Alongside primary and secondary education, there is special education (SO) for students with special needs because of learning disabilities or behavioral problems.

Upper secondary education (from grade 10) in HAVO and VWO is called Second Phase. This Second Phase follows on the partly common curriculum in grades 7, 8 and 9 in lower secondary education. As we mentioned already, students in the Second Phase are divided into four streams: Culture & Society, Economy & Society, Nature & Health, and Nature & Science. These so-called profiles have different mathematics programs. For example, unlike the students in the Nature & Science stream, students in the Culture & Science stream do not learn calculus. In 2007, the Dutch government decided to renew the Second Phase. One of the adaptations concerned the mathematics programs. In the so-called Renewed Second Phase, more attention is paid to algebra.

### 1.3 Research aims

Dutch universities played a central role in the discussion of algebraic proficiency in the Netherlands. They saw themselves confronted with students whose level of algebraic proficiency was considered as insufficient. In reaction, they organized remedial courses to repair the deficiencies. Although part of the problems in the universities could be ascribed to the fact that the universities were not prepared to adapt to the new mathematics programs of the Second Phase, there also seemed to be problems with the student level of algebraic proficiency (Sterk & Perrenet, 2005; Tempelaar, 2007; Vos, 2007).

In the discussion in the Netherlands, different causes were put forward for students' disappointing level of algebraic proficiency. For example, that the students had to take responsibility for their own learning process. Another cause that was mentioned is that the algebraic skills taught in lower secondary education were not maintained in higher secondary education. In addition, the introduction of the graphing calculator in upper secondary education was mentioned as a cause for the students' disappointing level. Using the graphing calculator was supposed to reduce students' ready knowledge.

This discussion served as input for this study. The point of departure is the students' debated level of algebraic proficiency in the transition from secondary to higher education. This study aims to obtain insight in the actual level of student algebraic proficiency in pre-university education. And, if this level is disappointing, to obtain insight into the deeper causes of the problems students experience in the preparation for pre-university education to higher education.

### 1.4 Research strategy

In order to obtain more insight into the actual level of student algebraic proficiency, we constructed four tests. The theoretical foundation of these tests was the relation between procedural fluency and conceptual aspects of algebraic proficiency, which also played a role in the discussion in the Netherlands about algebraic proficiency. Another aspect that played a role in this discussion was the relation between arithmetic and algebra. As we mentioned before, the discussion was triggered by problems in the transition from secondary to higher education, but eventually concerned the whole curriculum from primary education up to higher education. Therefore, this relation served as input into the construction of the tests. In our view, it is hard to draw a line between arithmetic and algebra. For example, calculating  $7 \times 17$  by means of  $7 \times 10 + 7 \times 7$  reinforces the distributive law, and  $7 \times 237 + 3 \times 237$  gives meaning to algebraic simplifications such as  $7a + 3a = 10a$  (French, 2002). In addition, in Dutch lower secondary education, students learn to simplify expressions with square roots, such as  $2\sqrt{5} + 4\sqrt{5} = 6\sqrt{5}$ . In a certain sense, this type of calculation is arithmetic as well as algebra: it is algebra because this kind of topic is not taught in

arithmetic in primary education; and it is arithmetic because this expression does not have a letter in it. In our view, the ability to deal with square roots, negative numbers, and fractions is part of algebraic proficiency. Therefore, we decided to add these topics into the tests. In Chapter 2 we further elaborate on the relation between arithmetic and algebra and on the relation between procedural fluency and conceptual aspects.

Based on these two aspects, we constructed four sequential paper and pencil tests in a partly cross-sectional and partly longitudinal design. A thousand-odd students in grades 8–12 (ages joined at least one of the sessions of the assessments, which were used to investigate how student algebraic proficiency developed. We used the Rasch model to analyze the data (Rasch, 1980; Bond & Fox, 2007; Linacre, 2009). This one parameter item response model has proven to provide good scales for mathematics proficiency in large assessments such as TIMSS (Olson et al., 2008) and PISA (OECD, 2003). With this model we were able to compare the results of the students of the four assessments even though the tasks of these four assessments were different.

The results of these assessments gave rise to further research on higher-order thinking skills. Based on the findings of this study, we decided to perform an analysis of the two most-used textbook series in the Netherlands.

## 1.5 Outline of the thesis

In Chapter 2, we describe the construction of the assessments and the Rasch analysis. We investigate the development of students algebraic proficiency, both cross-sectionally and longitudinally. In addition, we discuss results on individual tasks. The results of the tests created the need to get a better handle on the difficulties students experienced. Therefore, in Chapters 3 and 4, we focus the analysis to conceptual aspects of algebraic proficiency. In Chapter 3, we focus the analysis of the assessments on the ability to deal with the mathematical structure of algebraic expressions. In Chapter 4, we find an explanation for the difficulties students experience in the nature of mathematics.

In Chapter 5, we address methodological issues concerning a measure from the cognitive load theory that we used in the assessments. This measure was supposed to provide a more detailed view on student level of algebraic proficiency than just performance scores. However, using the measure revealed some methodological issues.

In Chapter 6, we describe the analysis of two Dutch textbook series. In this analysis, we focus on the role of contexts and models, and how these contribute to the learning process. We conclude this dissertation with a summary of our main findings, reflections on these findings, and practical recommendations in Chapter 7.

We aimed at writing chapters in such a way that they can be read independently. As a consequence, there are some overlaps.





## Chapter 2

# The development of students' algebraic proficiency in Dutch pre-university education

In this partly cross-sectional, partly longitudinal study, we investigate the development of algebraic proficiency in Dutch secondary education. As a theoretical background we use the relation between procedural and conceptual knowledge and the transition from arithmetic to algebra. Rasch analysis shows that students make progress, both cross-sectionally and longitudinally, but this progress is small. Furthermore, the development is not in the area of the conceptual aspects of algebraic proficiency. Finally, there is no significant difference in performance between students in the social stream and students in the science stream.

### 2.1 Introduction

Discussions of students' level of algebraic proficiency are taking place worldwide. International comparative studies such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA) induced studies on how to improve students' algebraic proficiency (e.g., National Mathematics Advisory Panel, 2008). In the Netherlands, the discussion focused on the level of basic algebraic skills in the transition from secondary education to higher education (Tweede Fase Adviespunt, 2005a; Van

Gastel et al., 2010). Complaints were heard that students were not proficient in basic algebraic algorithms and that they could not apply them correctly when entering higher education. As a result, educators and politicians called for a stronger emphasis on procedural skills (Van Gastel et al., 2007). However, from the mathematics education perspective, there is an urge to teach a deeper understanding that transcends the algebraic procedures. This deeper understanding appears to be of great importance in higher education. The following examples of McCallum (2010) may illustrate that the demands in higher education exceed the level of superficial procedural fluency:

- recognizing that  $P \cdot \left(1 + \frac{r}{12}\right)^{12n}$  is linear in  $P$  (finance);
- identifying  $\frac{n(n+1)(2n+1)}{6}$  as being a cubic polynomial with leading coefficient  $\frac{1}{3}$  (calculus);
- observing that  $L_0 \sqrt{1 - \left(\frac{v}{c}\right)^2}$  vanishes when  $v = c$  (physics);
- understanding that  $\frac{\sigma}{\sqrt{n}}$  halves when  $n$  is multiplied by 4 (statistics).

Our starting point is that the reasons for an unsatisfactory level of skills have to be sought in the preceding learning process. We focus on the algebraic skills taught in the lower secondary school, because the complaints by students as well as educators concern algebraic skills taught at the lower secondary level (Van Gastel et al., 2007, 2010). In order to better understand why students experience difficulties with these skills in higher secondary education, we chart students' algebraic proficiency in grades 8 through 12. The underlying rationale is that weaknesses or flaws will become manifest in the way the students' proficiency develops.

## 2.2 Different aspects of algebraic proficiency

This study aims at investigating students' development in algebraic proficiency. In order to investigate this development, we constructed tests based on two aspects of algebraic proficiency. The first aspect concerns the relation between basic skills and the deeper understanding that also plays a role in the aforementioned discussion. The second aspect concerns the transition from arithmetic to algebra. This transition is a central theme in lower secondary school algebra and is considered crucial for the development of algebraic proficiency. Below we elaborate on these two aspects.

### 2.2.1 The relation between procedural fluency and conceptual understanding

The distinction between fluency and understanding is central in discussions of algebraic proficiency and has been widely discussed in the past forty years (e.g.



expertise as one single dimension ranging from basic skills to symbol sense. Basic skills involve procedural work for which a local focus and algebraic calculations suffice. Symbol sense involves strategic work which requires a global focus and algebraic reasoning, see Figure 2.1. In this chapter we follow Drijvers (2006a) in viewing algebraic proficiency as a sliding scale ranging from basic skills to symbol sense.

### 2.2.2 Transition from arithmetic to algebra

This study focuses on students' development in algebraic topics taught in lower secondary pre-university education. We decided to include topics such as calculating with negative numbers, fractions, and square roots that are related to the transition from arithmetic to algebra. Knowledge about this transition is important for teachers who want to prepare students for algebra (Herscovics & Linchevski, 1994). This transition has been studied extensively (e.g. Filloy & Romano, 1989; Linchevski & Herscovics, 1996). The literature has identified several difficulties in the transition from arithmetic to algebra such as the lack of closure obstacle (Tall & Thomas, 1991), and the process-product dilemma (Sfard & Linchevski, 1994b). These difficulties also play a role in calculating with negative numbers and square roots. The lack of closure obstacle refers to the difficulty students experience when they have to handle an algebraic expression which represents a process that cannot be carried out. For example, the expansion of the brackets of  $2(3x - 5)$  leads to the expression  $6x - 10$  which has to be accepted as an answer instead of as a subtraction that has still to be carried out. An example of the process-product dilemma is handling square roots in expressions such as  $\sqrt{20} + \sqrt{5} = 3\sqrt{5}$  which represents both a process of extracting the root and the product of that process.

### 2.2.3 Research questions

The two aspects of algebraic proficiency discussed above play an important role in the development of the algebraic skills taught in lower secondary education. They form the basis for the construction of tests that aim at answering the following research questions.

1. How does students' algebraic proficiency develop from a cross-sectional perspective?
2. How does students' algebraic proficiency develop from a longitudinal perspective?
3. How does students' algebraic proficiency develop in terms of basic skills and symbol sense?

## 2.3 Methods

Our project aims at investigating the development of the algebraic proficiency of students in pre-university education. To monitor this development, the students are assessed four times, spread out over a calendar year in a partly cross-sectional and partly longitudinal design. In the Netherlands, students in upper secondary education follow one of four different programs. For convenience of comparison, we divided these four programs into two streams: a society stream and a science stream. Students in these streams follow partly the same and partly different mathematical programs. To complicate matters, a curriculum reform took place during the data collection. The new curriculum pays more attention to algebraic skills. So, in the analysis, this curriculum change has to be taken into account. Because of the different programs in upper secondary education, we focused on the common topics of these programs. A textbook analysis combined with an analysis of formal documents of the government showed that the following topics are taught in all programs: algebraic tasks concerning expanding brackets, simplifying expressions, and solving equations; and arithmetical tasks concerning negative numbers, fractions, and square roots.

### 2.3.1 Test design

The tests are based on the attainment targets formulated by the Dutch Ministry of Education, Culture and Science combined with the theoretical considerations discussed in Section 2.2. Because the attainment targets of the science stream include those of the social stream, we have based the tests on the attainment targets of the social stream (CEVO, 2006).

Numerical items are included that are related to the transition of arithmetic to algebra. These items are related to minus signs (for example, calculate  $-7 - (4 - 3) \cdot (-8) - 2$ ), calculating with fractions (for example, simplify  $\frac{8}{21} - \frac{2}{7} + \frac{1}{14}$ ), and square roots (for example, simplify  $2\sqrt{5} + 4\sqrt{5}$ ).

Algebraic items are included ranging from basic skills to symbol sense. These items are taken from the textbook series (e.g. Reichard et al., 2005) and teaching material of SLO (2008), as well as from the literature, for instance Arcavi (1994); Harper (1987); Wenger (1987); Matz (1982). The complete list of tasks can be found in the Appendix A.

For ease of following the individual students, we have used similar items in different assessments, put differently, we used an anchor design. We have chosen our tests to consist of open questions, to be worked out with paper and pencil. In this way, we avoid students' guessing answers. During the tests, students were not allowed to use calculators or notes. The tests consisted of 12 to 16 items and were designed to be completed in half an hour, in order not to overburden the students and teachers.

Grade	March 2008	May 2008	October 2008	February 2009	Part. 4×	Part. ≥ 1
8/9	164	227	173	171	94	266
9/10	163	160	129	114	56	217
10/11 Total	243*	185	163	144	90	268*
Social	95	60	49	51	27	103
Science	144	125	114	93	63	161
11/12 Total	244*	204	188	72	37	269*
Social	117	103	90	26	17	132
Science	125	101	98	46	20	134
Total	814	776	653	501	277	1020

\* The total is larger than the sum of the social and science stream because some students did not report their stream.

Table 2.1: Number of students in the cross-sectional and longitudinal data collection.

### 2.3.2 Data collection

We assessed students in March 2008, May 2008, October 2008, and February 2009. Students of grades 8, 9, 10 and 11 (ages 13–16) participated in the first and second assessments. After the summer vacation, these students were in grades 9 up to 12 in October 2008 and February 2009. Table 2.1 provides an overview of the numbers of students that participated. Four schools participated, two schools using textbooks of *Getal & Ruimte*, and two schools using textbooks of *Moderne Wiskunde*, see for example (Reichard et al., 2006) and (De Bruijn et al., 2007). These textbooks together have an estimated market share of 90% (cTWO, 2009), and so can be seen as representative. Three schools are situated in the south of the Netherlands and one school is situated in the middle of the Netherlands. All four schools are situated in urbanized parts of the Netherlands mainly with children of Dutch origin. From each school, two classes of each grade participated. In total, 1020 students of four schools participated at least once. Figure 2.1 shows the numbers of students per grade that participated. In grades 8 and 9 from each school, two classes of each grade participated. In grades 10, 11 and 12, students of the social stream and of the science stream participated. The written answers were coded 1 for correct and 0 for incorrect. Doubtful cases were discussed with colleagues.

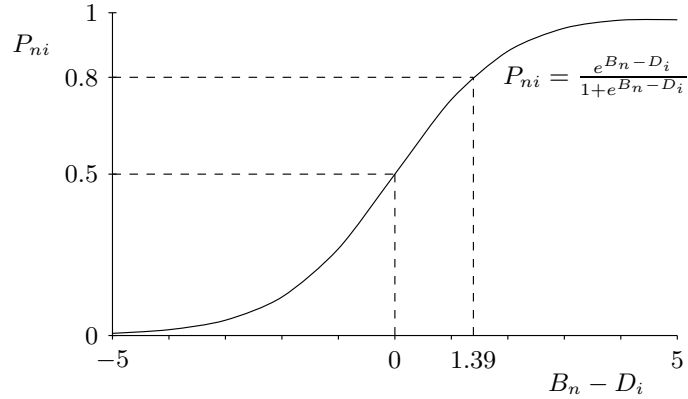


Figure 2.2: Probability of success as a function of ability relative to item difficulty.

### 2.3.3 Creating Rasch scales for algebraic proficiency

For the analysis of our data, we used the Rasch model, a one parameter item response model (Rasch, 1980; Bond & Fox, 2007; Linacre, 2009). With a Rasch analysis, one linear scale is created on which both persons are situated according to their ‘ability’ and items according to their ‘difficulty’. On this scale, not only the order but also the distances between the items and the students have meaning. Rasch theory supposes that the probability of a person’s giving a correct answer on an item is a logistic function of the difference between person’s ability and the difficulty of the item, see Figure 2.2. To put it more precisely, the probability  $P_{ni}$  of person  $n$  with ability  $B_n$  to correctly answer item  $i$  with difficulty  $D_i$  is given by

$$P_{ni} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}}.$$

Both the ability of the persons and the difficulty of the items are measured in so-called units of log odds ratios, or logits. The local origin of the Rasch scale is usually situated in the center of the range of item difficulties. As a consequence, if the ability equals the item difficulty, that is, if  $B_n = D_i$ , then

$$P_{ni} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}} = \frac{e^0}{1 + e^0} = \frac{1}{2}.$$

We consider this model appropriate for three reasons. First, we expect the Rasch model to provide a more detailed view of students’ algebraic proficiency level than  $p$ -values, because the Rasch model takes the difficulties of the items into account. The second reason is that we conjecture that algebraic proficiency



satisfies the assumption of the Rasch model of possessing one unidimensional construct as discussed in Section 2.2. Furthermore, Rasch models are also used in international mathematics education surveys such as TIMSS (Olson et al., 2008) and PISA (OECD, 2003). The third reason to use the Rasch model is that it provides the opportunity to anchor the scales of the four assessments. This makes it possible to compare the results of the students of our four assessments even though the items of these four assessments are different.

For each assessment, we created its corresponding Rasch scale. However, the purpose of the project is also to follow individual students in time. Therefore, we connected the four Rasch scales by using anchor items. Although all four assessments consisted of different items and none of the items occurred in more than one assessment, some of the items in different assessments are quite similar. Consider for instance the four items in which students are asked to expand the brackets, namely  $-4(3a+b)$ ,  $-5(2p+q)$ ,  $-3(4p+q)$ , and  $-4(5p+q)$ . Whether or not these items are suitable to use as anchor items is determined by a Differential Test Functioning (DIF) (Linacre, 2009). DIF indicates that after adjusting for the overall scores of the respondents, one group of respondents scores better on a specific item than another group of respondents.

The results of the DIF-analysis suggest that five items of each assessment are suitable for serving as anchor items to connect the Rasch scales. Based on the item measures of these five items, we connected the four Rasch scales of the four assessments. As a consequence, items of all four assessments are placed on one scale. Also, students of different assessments have a Rasch measure on one scale.

In our analysis, we want to determine which items students master. Therefore it is necessary to determine what we view as mastery. We can arguably consider a probability of 80% of answering an item correctly as an expression of mastering that item. From the Rasch model it follows that a probability of 0.8 of person  $n$  answering item  $i$  correctly corresponds to an ability  $B_n$  which is 1.39 logit higher than the difficulty  $D_i$  of item  $i$  because if  $B_n - D_i = 1.39$ , then

$$P_{ni1} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}} = \frac{e^{1.39}}{1 + e^{1.39}} \approx 0.8.$$

As a consequence, we consider that students with a measure at least 1.39 higher than the measure of the item master that particular item.

### 2.3.4 Fit

From the assumptions of the Rasch model, it follows that students and items with extreme scores are not directly estimable. In our case, none of the items in our tests had an extreme score. That is, no item was answered correctly by all students and no item was answered incorrectly by all students. There were, however, individual students with extreme scores: in the first assessment, 23 out of 814 students answered all items correctly; one students answered all items

incorrectly. In the second assessment 40 out of 776 students had extreme scores; in the third assessment, 5 out of 653 students and in the fourth assessment, 6 out of 501 students had extreme scores. In the analysis, students with extreme scores are left out. These students are imputed a Rasch measure based on those of the other students. As a consequence, Rasch measures at the tail end have to be treated with care.

To evaluate the fit of the Rasch model on our data, we checked item polarity, infit and outfit, reliability, and multidimensionality. At first, we verified the point-measure correlations related to item polarity. These correlations reflect the extent to which items are aligned in the same direction on the latent variable. If these correlations are positive, students with a high ability perform better on these items than students with low ability. In our case, all correlations are positive.

From the 1020 participating students, 163 students show an outfit greater than 2 in at least one of the four assessments. Mostly, these high outfit scores are due to incorrect solutions to items with low measure, for example, “expand  $-5(2p + q)$ .” We couldn’t find a specific pattern for the group of students with a high outfit that failed some low measure item, therefore we decided not to exclude these students from the analysis. Furthermore, we conclude that the high outfit is due to the small number of items in our assessments. Consequently, we must draw conclusions carefully.

With respect to the reliability, we found values of .70, .70, .66 and .68 for assessments 1, 2, 3 and 4, respectively. These reliability scores can be compared to Cronbach’s alpha.

With respect to multidimensionality, for each assessment the Rasch analysis yielded that the raw unexplained variance of the model almost equals the raw unexplained variance of the empirical data. Furthermore, the contrasts of the principal component analysis yielded no other significant factors.

Based on the considerations above, we conclude that the items in our assessments reflect one and the same underlying latent variable.

## 2.4 Cross-sectional and longitudinal results

The analysis is performed from different perspectives. First, we investigate the development of algebraic proficiency cross-sectionally and longitudinally. Next, we analyze student ability from a more theoretical point of view, related to the aspects of algebraic proficiency we described in Section 2.2.

### 2.4.1 Cross-sectional results

First, we perform a cross-sectional comparison of grades 8 through 12. Figure 2.3 shows the percentiles of Rasch measures in logits of four groups of students who were assessed on March 2008, May 2008, October 2008 and February 2009

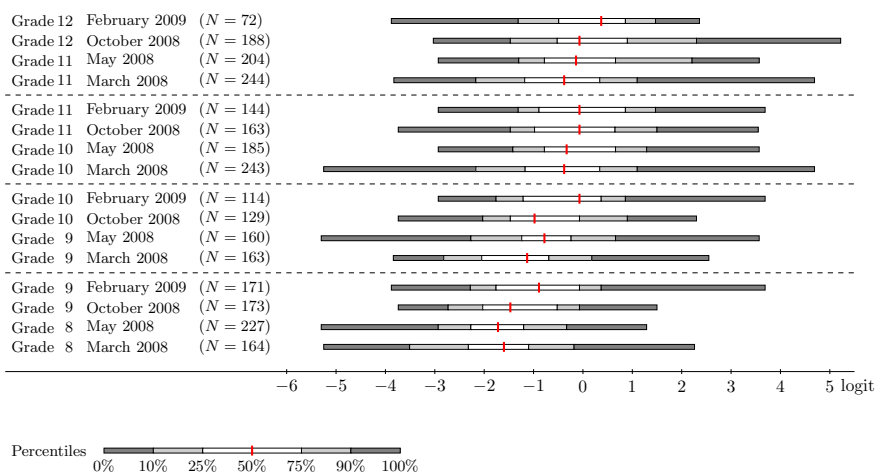


Figure 2.3: Cross-sectional percentiles of all grades in all assessments.

respectively. As we explained above, the number of participating students varied over the assessments. As a consequence, the bars between the dashed lines in Figure 2.3 partly represent the same students. From the percentiles of the Rasch measures, we conclude that generally, the averages of the different assessments increase with the grades. If we focus on the little lines that represent 50% of the students, we see that there is a difference of approximately two logits between the lowest average (grade 8, May 2008) and the highest average (grade 12, February 2009). A more detailed view of Figure 2.3 reveals that grade 9/10 performs better than grade 8/9 and grade 10/11 performs better than grade 9/10. The difference in ability between grade 10/11 and grade 11/12 is less obvious. Here we note that the curriculum of students in grade 11/12 differs from the curriculum of students of the other grades. Recall that the difference between these curricula concerns algebraic skills. As a consequence, it is hard to draw a conclusion from the lack of difference between grades 10/11 and grades 11/12. The only tentative conclusion we propose is that the new curriculum may lead to a higher proficiency, if the growth in ability from grade 8/9 to grade 9/10 to grade 10/11 continues.

When we look at the dispersion of the percentile scores displayed by the bars in Figure 2.3, we see that in each bar, 50% of the students (the white area) is within two logits of each other. Thus, the dispersion within the twelve or sixteen different assignments is rather small with regard to the difference between the worst and the best scoring student (the length of the whole bar). If we focus on 80% of the students (i.e., leaving out the best and worst 10%), we see that they are within a range of at most 3.8 logits of each other.

With respect to the different streams the students followed starting from

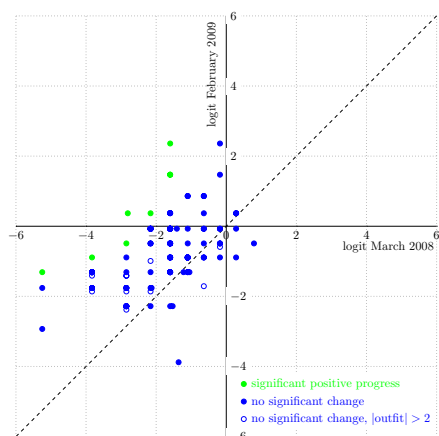


Figure 2.4: Rasch measures of grade 8/9 students in the first and fourth assessments ( $N = 104$ ).

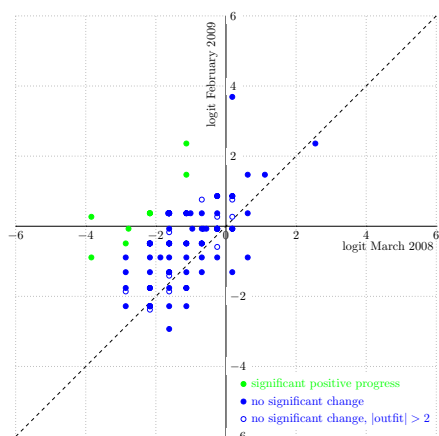


Figure 2.5: Rasch measures of grade 9/10 students in the first and fourth assessments ( $N = 90$ ).

grade 10, students of the science stream outperform those of the society stream. This difference is not significant. The Rasch model provides standard errors which provide a 95% confidence interval around the measures. A rule of thumb considers differences greater than 1.5 times the sum of the standard errors as significant (Linacre, 2009). In this case, the averages differ by less than 1.5 logit with standard errors around 0.75 logit, which is not significant. Note, however, that because of the relatively small number of items, differences would have to be quite large before they become significant.

To sum up, the cross-sectional analysis showed that there is progress and there is only little dispersion among the middle 50% of the students. There seems to be a growth in ability between the generation of grade 10/11 and that of grade 11/12, which might be a positive effect of the curriculum changes. In grades 10, 11 and 12, students of the science stream outperformed students of the social stream. However, this difference is not significant.

### 2.4.2 Longitudinal results

The cross-sectional analysis in subsection 2.4.1 showed progress in subsequent grades. For a more detailed view of the development of algebraic proficiency, we analyze the development of individual students over the four assessments.

Figure 2.4 shows the Rasch measurements in logits of individual students in the first assessment in March 2008 (horizontally) compared to the Rasch measurements of these students in the fourth assessment in February 2009 (vertically). Each dot represents one student.

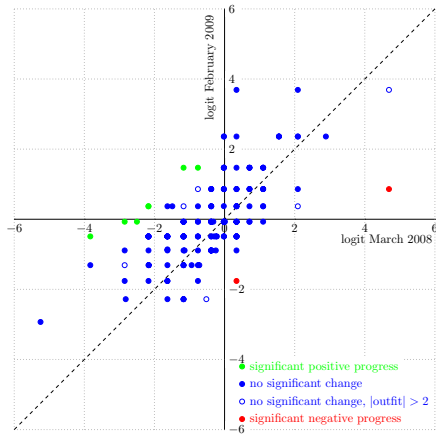


Figure 2.6: Rasch measures of grade 10/11 students in the first and fourth assessments ( $N = 131$ ).

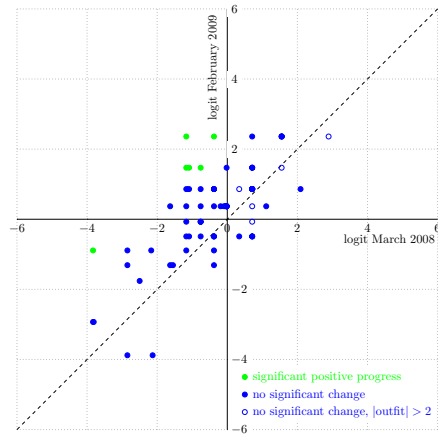


Figure 2.7: Rasch measures of grade 11/12 students in the first and fourth assessments ( $N = 65$ ).

Students above the dashed line score better in the fourth assessment than in the first assessment. Hence these students have shown progress. Students below the dashed line show lower ability in the fourth assessment than in the first assessment. Hence these students retrogressed. The standard error of the Rasch measures provides a 95% confidence interval to determine whether or not the changes are significant. Figures 2.4, 2.5, 2.6 and 2.7 show the ability of students of subsequent grades in the first assessment compared to the ability in the fourth assessment. The majority of the students did not make significant progress: 96 students out of 104 students in grade 8/9; 81 of 90 students of grade 9/10; 122 out of 131 students in grade 10/11; and 59 out of 65 students in grade 11/12 did not make significant progress. In subsequent grades, 8 out of 104; 9 out of 90; 7 out of 131; and 6 out of 65 students made positive progress. In grade 10/11, 2 students out of 131 retrogressed. With respect to the social and science streams of students in grade 10, 11 and 12, we found no difference between students in different streams. Both streams possess some students who progress and some who retrogress. The two retrogressing students in grade 10/11 are both in the science stream.

Summarizing, the longitudinal analysis showed that students in grade 8/9, 9/10, 10/11 and 11/12 make progress during a calendar year. Only few students made significant progress. In the following section, we relate the students' abilities to the difficulties of the items.

## 2.5 Results on individual items

As discussed in Section 2.3, Rasch analysis yielded one scale on which the students as well as the tasks of all four assessments have a measure. This construction provides the opportunity to relate students' abilities to all tasks used in the tests. We did not try to separate items that test basic skills from those that test symbol sense for several reasons. The main reason is that from the evaluation of the written answers of the students, it became clear that the degree to which an item appealed to symbol sense is a property of that item as well as of the strategy used by the students to solve the problem. Another reason is that an item that appeals to symbol sense for a student in grade 8 might be basic for a student in grade 12.

Tasks are divided into two main categories: algebra and numbers. These two categories are both divided into three subdivisions: the category algebra is divided into the subdivisions expanding brackets, simplifying, and solving equations; the category numbers is divided into negative numbers, fractions, and square roots. Figure 2.8 shows the percentiles of Rasch measurements of the students of the fourth assessment in relation to the Rasch measures of the tasks. The Rasch scale with unit logits is in the center of the figure. Above the axis, the percentiles of students in the fourth assessment are presented. Below the axis, the measures of all tasks are presented, added with 1.39. In this way, students with the same measure as a task have a probability of 0.80 of answering that task correctly (see also Section 2.3). The vertical dashed line in the center of the figure at  $-0.49$  logit indicates the ability of 75% of the participating students in grade 12. Tasks on the left-hand side of the dashed line are mastered by more than 75% of the students in grade 12; tasks on the right-hand side of the dashed line are mastered by less than 75% of the students in grade 12. Below, we discuss two categories of tasks: algebraic tasks and arithmetical tasks.

### 2.5.1 Algebra

In Table 2.2, we listed the difficulties of a selection of tasks. These difficulties correspond to a probability of 0.80 of answering that task correctly. The full list of tasks can be found in the Appendix.

#### Expanding brackets

In the easiest task, students are asked to expand the brackets in  $-4(3a + b)$ . Almost all students master this task in grade 12. The other two tasks are mastered by less than 75% of the students. A common error in expanding the brackets in  $-2(4x - y) + 3(-2y - 4)$  is the oversimplification to  $2x + y + 3$  by dividing  $-8x - 4y - 12$  by  $-4$ . This step would be appropriate in case the expression was followed by  $= 0$ . Another type of error is caused by minus signs. This kind of error might occur because of an increase of cognitive load due to a clutter of

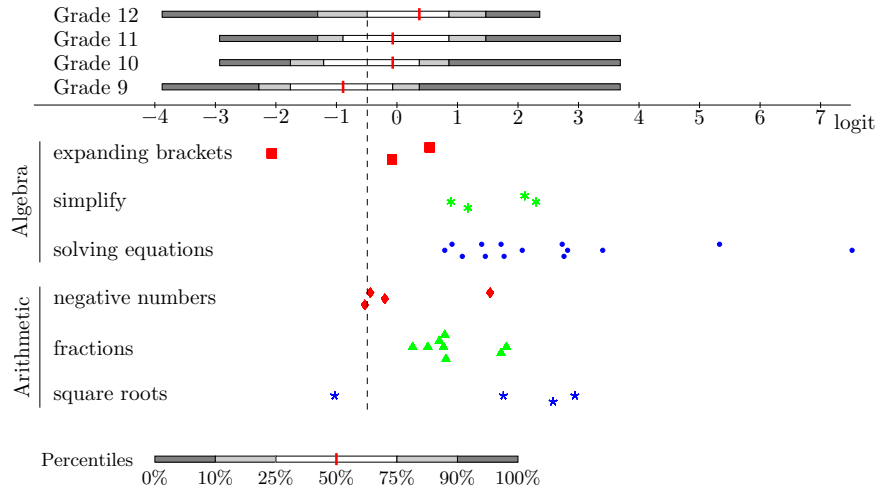


Figure 2.8: Percentiles of student ability in the fourth assessment in relation to the difficulty of mastering the items, on the Rasch-scale.

minus signs, rather than because of misunderstanding (Ayres, 2001). We conclude that students are able to expand brackets, but only up to a certain degree of difficulty.

### Simplifying

Simplifying the expression  $\frac{5x^2+10-2(2x^2+4)}{x^2+2}$  leads to the expression  $\frac{x^2+2}{x^2+2}$ . Recognizing that this expression can be reduced to  $\frac{x^2+2}{x^2+2} = 1$  requires students to identify  $x^2 + 2$  as one entity and realize that  $x^2 + 2 \neq 0$  for all  $x \in \mathbb{R}$ .

The substitution of  $a = -2$  and  $b = -1$  in  $-(a^2b)^3 - 2(ab^2)^2$  is mastered by less than 10% of the students. Substituting  $a = -2$  and  $b = -1$  in the expression yields a clutter of minus signs which is apparently difficult for students to handle. A possible strategy is to first calculate  $a^2 = 4$  and  $b^2 = 1$  and substitute these expressions separately to reduce the amount of minus signs. Almost all students in grade 8 to grade 12 lack the flexibility to manage this large amount of minus signs. This flexibility, however, is part of symbol sense, so we conclude that students lack this part of symbol sense.

### Solving equations

All tasks related to solving equations are mastered by less than 50% of the students of grade 12. The equations  $\frac{21}{6+\frac{5}{1+x}} = 3$  and  $(x-5)(x+2)(x-3) = 0$

Task	Measure
Expand the brackets: $-4(3a + b) =$	-2.07
Simplify: $-2(4x - y) + 3(-2y - 4) =$	-0.08
Simplify: $(3a^2 + 2a + 7)(a + 8)$ . Show your working.	0.54
Simplify: $\frac{5x^2+10-2(2x^2+4)}{x^2+2} =$	1.17
Substitute $a = -2$ and $b = -1$ in $-(a^2b)^3 - 2(ab^2)^2$ .	2.30
Solve: $\frac{21}{6+\frac{5}{1+x}} = 3$ .	1.40
Solve: $(x - 5)(x + 2)(x - 3) = 0$ .	2.07
Is there any $x$ for which $\frac{2x+3}{4x+6} = 2$ ? If so, calculate $x$ ; if not, please explain why such an $x$ doesn't exist.	2.76
Solve: $2(3x + 2) = 3(2x - 1) + 7$ .	3.40
Solve: $a\sqrt{2} = 1 + 2a\sqrt{3}$ .	5.33
If $a\sqrt{b} = 1 + 2a\sqrt{1+b}$ , then $a =$	7.52

Table 2.2: Algebraic tasks with corresponding difficulty (probability of success 0.80).

both require that students identify a part of the expression as a whole.

In the first equation, whether students use the cover-up method and replace the equation for example with  $\frac{21}{\square} = 3$ , or multiply both sides of the equation by  $1+x$  or  $6+\frac{5}{1+x}$ , these strategies all have in common that a part of the expression has to be treated as a whole, as an object. Similarly, in the second equation,  $(x-5)$ ,  $(x+2)$  and  $(x-3)$  have to be treated as wholes. The work of the students showed that students tend to expand the brackets, after which they can neither find the factorization, nor suddenly realize the equation is simple, see Figure 2.9.

Almost no students were able to conclude that the equation  $\frac{2x+3}{4x+6} = 2$  has no solutions because  $\frac{2x+3}{4x+6} = \frac{1}{2}$  for all  $x \neq -1\frac{1}{2}$ . Arcavi (1994) argues that the ability to withstand the invitation to solve this equation directly and instead perform an a priori inspection and conclude that the quotient equals  $\frac{1}{2}$  except for  $x = -1\frac{1}{2}$  is an expression of symbol sense. Solving this equation directly as well as performing an a priori inspection is beyond the ability of almost all students.

Expanding the brackets in the equation  $2(3x + 2) = 3(2x - 1) + 7$  yields  $6x + 4 = 6x + 4$ . A few students were able to conclude that every  $x \in \mathbb{R}$  is a solution of this equation. The equation  $a\sqrt{2} = 1 + 2a\sqrt{3}$  is an adapted version of Wenger's equation that is mastered by only a few students. The most difficult task is Wenger's equation: exactly 2 out of the 650 participating secondary school students in the assignment of March 2008 were able to solve this equation.

The equations discussed above call for taking a broad view of solving equa-



Los op:

$$(x-1)(x+3)(x-4) = 0.$$

$$x^2 + 3x + x^2 - 4x - x - 3 + x + 4 + x^2 - 4x + 3x - 12 = 0$$

$$3x^2 - 5x - 11 = 0$$

$$x - 1 = 0 \quad x = 1$$

$$x + 3 = 0 \quad x = -3$$

$$x - 4 = 0 \quad x = 4$$

Figure 2.9: Typical student work on the task: solve  $(x - 1)(x + 3)(x - 4) = 0$ .

tions. This broad view, which is seen as a part of symbol sense, includes the acceptance of more than one possible solution and having a Gestalt view. From the students' performance on solving equations, we conclude that the majority of the students did not have such a broad view of equations.

## 2.5.2 Arithmetic

In Table 2.3, we listed the difficulties of a selection of tasks, the full list of tasks can be found in the Appendix. We discuss students' performance on these tasks below.

### Negative numbers

Tasks related to negative numbers are grouped around  $-0.50$  logit, except for one that is at  $1.54$  logit. The group of tasks around  $-0.50$  logit is mastered by almost 75% of the students. The task around  $1.50$  logit involves calculating the expression  $-7 - (4 - 3) \cdot (-8) - 2$ . This task is perceived as much more difficult than the other tasks in the subcategory of negative numbers. In our view, the difficulty of this task can be related to the ability to see parts of an expression as entities. For example, if students recognize the structure of the expression and know the priority rules for arithmetic,  $(4 - 3)$  as 1, the expression immediately transforms into  $-7 - 1 \cdot (-8) - 2$ , which is more manageable.

Summarizing, we conclude that the students mastered simple expressions with negative numbers, but more complicated expressions rapidly grew beyond their capabilities.

Task	Measure
Calculate: $\frac{18-5}{7-20} + 4 =$	-0.53
Calculate: $3 \cdot (-2) \cdot 5 - 2 \cdot 5 =$	-0.44
Calculate: $-7 - (4 - 3) \cdot (-8) - 2 =$	1.54
Simplify: $\frac{7}{15} - \frac{2}{5} + \frac{1}{10} =$	0.70
Jantine claims that $2\frac{1}{2} \times 3\frac{1}{3} = 6\frac{1}{6}$ . Explain why this is incorrect.	0.79
Simplify: $\frac{8}{21} - \frac{2}{7} + \frac{1}{14} =$	1.81
Simplify: $2\sqrt{5} + 4\sqrt{5} =$	-1.03
Simplify: $\frac{\sqrt{12}}{\sqrt{3}} =$	2.58
Martijn claims that $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ . Explain why you do or do not agree with Martijn.	2.94

Table 2.3: Arithmetical tasks with corresponding difficulty (probability of success 0.80).

## Fractions

In Figure 2.8, tasks related to fractions are split into two groups. One group of tasks involves the multiplication of mixed numbers; the second group of tasks refers to the addition of unlike items. In the first type of task, students were asked to explain why  $2\frac{1}{2} \times 3\frac{1}{3}$  does not equal  $6\frac{1}{6}$ . Students did not necessarily have to calculate  $2\frac{1}{2} \times 3\frac{1}{3}$ , because a reasoning such as that  $2 \times 3\frac{1}{3}$  already exceeds  $6\frac{1}{6}$  is a correct answer. Less than half of the students of grade 12 mastered this task.

In the second group of items, students were asked to add unlike fractions. Similar tasks, with the same mathematical structure, were included in all four assessments. In the first and the third assessment, the denominators were multiples of 7; in the second and the fourth assessment, the denominators were multiples of 5. Although the mathematical structure of these four tasks is the same, the results show a large difference in difficulty. The tasks with denominator multiples of 7 are more than one logit more difficult than the tasks with denominator 5. This suggests that the underlying mathematical structure is less important to students than the familiarity with the numbers 5, 10 and 15.

From these results we conclude that calculating with fractions, although taught in primary education, remained difficult for the students in secondary education.

### Square roots

The addition  $2\sqrt{5} + 4\sqrt{5}$  is the easiest task of the number tasks and is mastered by the majority of the students in grade 12. Tasks in which a square root has to be manipulated by extracting the square are much more difficult. In our view, rewriting  $\sqrt{12}$  to  $2\sqrt{3}$  does not only require applying the distributive rule  $\sqrt{ab} = \sqrt{a}\sqrt{b}$ . Instead, several steps have to be taken which all together form a sequence of steps which is difficult for students. These steps involve the recognition that 12 is a number that contains a square ( $12 = 4 \cdot 3$ ), the application of the distributive rule ( $\sqrt{4 \cdot 3} = \sqrt{4} \cdot \sqrt{3}$ ), and taking the square root of 4 ( $\sqrt{4} \cdot \sqrt{3} = 2\sqrt{3}$ ). We conclude that students master simple addition of square roots. Manipulating square roots, however, is beyond the ability of the majority of the students.

Summarizing, the results on numbers show that calculating with negative numbers, fractions and square roots was difficult for the majority of the students. The ability to flexibly deal with minus signs, fractions and square roots is seen as a prerequisite for developing algebraic proficiency. Further research focused on arithmetic is necessary to place these findings in the ongoing curriculum from primary education to secondary education.

### 2.5.3 Comparing students of the social and science streams

As we mentioned in subsection 2.4.1, students in the science stream perform better than students in the social stream. This difference turned out not to be significant. The difference between the performance of students in the social and science streams is presented in Figure 2.10. In this figure, percentiles of Rasch measures of grade 11 and grade 12 students in February 2009 are shown from the total group of students, and of students split up into social and science streams. In total, 144 grade 11 students participated, of which 51 students were in the social stream and 93 were in the science stream. In grade 12, 72 students participated, of which 26 students were in the social stream and 46 were in the science stream. These numbers of students are really too small to calculate percentiles. Nevertheless, from Figure 2.10 it becomes clear that the difference between the performance of social and science students is small.

Furthermore, in Figure 2.8 we related the measures of the students to the measures of the items by drawing the 75% line. Students on the right side of this line master items on the left side of this line. The 75% line of the total group is at  $-0.49$  logit, whereas the 75% line of the science stream is at  $0.37$  logit. As a consequence, the science students only master a few items more than the students of the total group. So, for the science students as well, the majority of the items are not mastered.

Summarizing, science students outperform the social students, but the difference in the number of items that is mastered is small. In our view, this is

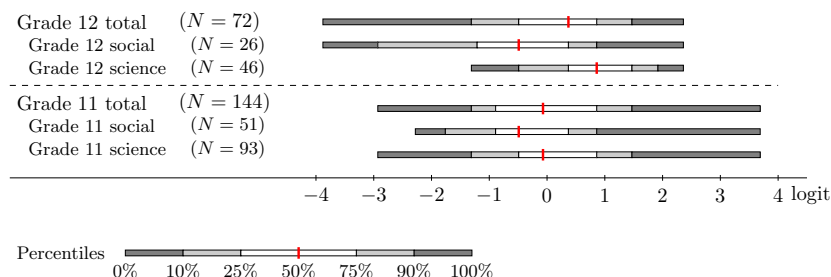


Figure 2.10: Percentiles of social and science students on the Rasch scale in February 2009.

an indication that the Dutch educational system concentrates on students with average ability. As such, the educational system misses serving the more talented students.

## 2.6 Conclusions and discussion

### 2.6.1 Main findings

This study reports on the development of algebraic proficiency in Dutch pre-university education and aims at answering the following research questions.

1. How does students' algebraic proficiency develop from a cross-sectional perspective?
2. How does students' algebraic proficiency develop from a longitudinal perspective?
3. How does students' algebraic proficiency develop in terms of basic skills and symbol sense?

The answer to the first question is that students made progress. The difference between grade 10/11 and grade 11/12 is small. This might be due to a curriculum change. Students in grade 8/9, grade 9/10 and grade 10/11 followed a curriculum program that includes more algebra than the program of students of grade 11/12. This curriculum change might have had a positive effect on students' algebraic proficiency if the growth in ability from grade 8/9 to grade 9/10 to grade 10/11 continues. Furthermore, there is little dispersion among the middle 50% of the students.

The answer to the second question is that the longitudinal analysis yields that the majority of the students made progress between the first and the fourth

assessments that was administered over a period of one year. However, for the majority of the students, this progress was not significant. Furthermore, students of the science stream performed better than students of the social stream. These differences are not significant either.

The answer to the third question is that there are only few tasks that were mastered by the majority of the students. In general we conclude that students mastered simple tasks, but tasks become too complicated rather quickly, for example due to a clutter of minus signs or brackets. The addition and multiplication of fractions and the multiplication of square roots was too difficult for the majority of the students. Furthermore, the majority of the students did not show important aspects of symbol sense. For example, students lacked the flexibility to manage minus signs in the substitution of  $a = -2$  and  $b = -1$  in the expression  $-(a^2b)^3 - 2(ab^2)^2$ . Another example is the ability to see a part of an expressions as an object in its own right, e.g., in  $\frac{21}{6 + \frac{5}{1+x}} = 3$  and  $(x - 5)(x + 2)(x - 3) = 0$ . In the latter, students tend to expand the brackets and then try to find the factorization, instead of recognizing the mathematical structure  $A \cdot B \cdot C = 0$  that implies  $A = 0$  or  $B = 0$  or  $C = 0$ . These flexible skills are important aspects of symbol sense. However, the range in which the development of the majority of the students takes place does not affect the range of tasks related to these aspects.

With respect to social and science students we found that science students outperformed social students. This difference however is not significant. For the science students as well it holds that the development is not in the area where most of the items are. The small difference between social and science students gives rise to the worry that education is missing opportunities for serving the more talented students.

## 2.6.2 Discussion

The results of this study show that although students showed progress both cross-sectionally and longitudinally, this progress did not lie in the range of the difficulties of the majority of the items. In other words, the majority of the items were too difficult for students of grade 8, and were still too difficult for students of grade 12.

The results must however be seen in the context of the limitations of this study. Firstly, because we did not want to place too heavy a load on the teachers and the students, we chose to keep the number of items relatively low. As a consequence, differences would have had to have been quite large to be significant. Secondly, a curriculum change took place during the data collection. The new curriculum pays more attention to algebra. We took this change into account by concluding that the curriculum change might have a positive effect if the growth continues. Thirdly, the number of students in the social and science stream is low, which affect the generalization of the findings to all social and all science students.

Finally, due to the explorative character of this study, the assessments contained many different tasks covering different underlying concepts. As a consequence, it is not completely clear which concepts students did not master.

We consider the results disappointing for two reasons. First, students' development does not lie in the range of the majority of the tasks. This is disappointing because these tasks fit in the regular curriculum. Secondary education teachers considered these specific tasks as being appropriate and as lying within the abilities of the students (Schutte, 2010). We are aware that only a small part of the students who participated will choose technical studies in the future. Nevertheless, in our view, students who will choose a study in social science or economics also need algebraic proficiency. Secondly, the small difference between the performance of students in the social and the science streams makes it likely that the present Dutch curriculum does not provide enough opportunities to students in the science stream.

These findings are in line with international research (e.g. Stacey et al., 2004) and with recent research in the field of fractions among Dutch students about the same age (Bruin-Muurling, 2010). In our view, the disappointing performance of the students is not only due to a lack of symbol sense. More importantly, we see the lack of symbol sense as a symptom of a lack of mastering a formal level of reasoning at which the mathematical structure and ambiguous nature of the mathematics is central.

Different researchers have argued that reaching this higher level is inherently difficult and involves a shift of thinking. For example, Freudenthal (1971) stated that "The student applies certain new rules unconsciously until at a certain moment he becomes conscious about them." To Sfard (1991), reaching this higher level is so difficult, that it remains practically out of reach for certain students. However, as the examples of McCallum (2010) in the Introduction show, reaching a more formal level of mastering algebraic expressions is especially important to students in higher education.

How to reach this higher level is a core concern of the mathematics education research community. In our view, the call from educators and politicians for more attention to routine and procedural skills will not solve students' problems. This study has shown that these problems come not so much from a lack of procedural skills, but more from a lack of conceptual understanding.

Teaching tasks that require symbol sense carries the risk of a focus on specific advanced strategies, where the procedure is not grounded in understanding. Furthermore, a focus on symbol sense carries the risk that number and situation dependent strategies are seen as an expression of symbol sense, whereas symbol sense concerns a feeling for the mathematical structure and the ability to flexibly deal with ambiguity of mathematical concepts (e.g. Byers, 2007).



## Chapter 3

# Structure sense as an aspect of algebraic proficiency

In this chapter, we investigate the development of algebraic proficiency in Dutch secondary education using structure sense as a theoretical lens. The results indicate that students' development on the Rasch scale does not coincide with the range of tasks concerning structure sense. In addition we find that the relation between structure sense and students' level of proficiency is not always clear.

### 3.1 Introduction

In Chapter 2, we investigated the development of students' algebraic proficiency in Dutch pre-university education. The results indicated that the students made progress, both cross-sectionally and longitudinally, but this progress is small. Further, students' development was not in the area of the conceptual aspects of algebraic proficiency.

In this chapter, we focus on one underlying conceptual aspect, namely structure sense. In this way, we aim at obtaining a more detailed view of students' level of algebraic proficiency. Structure sense can be described as the ability to deal with the mathematical structure of algebraic expressions. Many researchers have argued that flexible manipulation skills are an important aspect of algebraic proficiency (Tall & Thomas, 1991; Sfard & Linchevski, 1994b; Kilpatrick et al., 2001; Stacey et al., 2004). These flexible skills include the ability to see the algebraic structure of an expression. In the analysis, we focus on tasks that appeal to structure sense. These tasks include algebraic tasks as well as numerical tasks.

The outline of this chapter is as follows. First, we discuss structure sense and how this theoretical lens suits the theoretical background underlying the



assessments of Chapter 2. Next, we analyze in depth student performance in relation to structure sense.

## 3.2 Structure sense

One of the conclusions of Chapter 2 was that the students' development did not lie in the area of the conceptual aspects of algebraic proficiency. In this chapter, we focus on one underlying conceptual aspect, namely structure sense. In this way, we aim at obtaining a more detailed view on the students' level of algebraic proficiency.

The term structure sense is introduced by Linchevski & Livneh (1999) to describe the ability "to use equivalent structures of an expression flexibly and creatively." Their research concerned students in the transition from arithmetic to algebra, and the structure they examined concerned the order of arithmetical operations such as  $50 - 10 + 10 + 10$ , which is viewed as  $50 - (10 + 10 + 10)$  by half of the grade 6 and grade 7 students. In high school algebra, structure sense is seen as a collection of abilities, such as: recognize a structure, see a part of an expression as a unit; divide an expression into meaningful sub-expressions; recognize which manipulation is possible and useful to perform; and choose appropriate manipulations that make the best use of the structure (Hoch & Dreyfus, 2004, 2006; Novotná & Hoch, 2008). More precisely, Novotná & Hoch (2008) define that students display structure sense if they can 1. recognize a familiar structure in its simplest form; 2. deal with a compound term as a single entity and through an appropriate substitution recognize a familiar structure in a more complex form; 3. choose appropriate manipulations to make best use of a structure.

In contrast with Novotná & Hoch (2008) who see structure sense as an extension of symbol sense, we see structure sense as a part of symbol sense. Symbol sense is a complex feel for symbols that include a positive attitude towards symbols and a global view of expressions (Arcavi, 1994, 2005). Parts of this global view are the ability to read through symbols and flexible manipulation skills that can be related to seeing the algebraic structure in an expression. As an example of the ability to read through symbols, Arcavi (1994) discusses the equation  $(2x + 3)/(4x + 6) = 2$ . Reading through the symbols reveals that the left-hand side of the equation equals  $\frac{1}{2}$  for all  $x \neq -1\frac{1}{2}$ , because the numerator equals half the denominator. Inspecting the equation before starting to solve it with the purpose of gaining a feeling for the meaning of the problem is seen as an instance of symbol sense. Flexible skills include the ability to sense symbols not only as a concatenation of letters, but as arranged in a special pattern. As an example, Arcavi (1994) discusses Wenger's Equation,  $v\sqrt{u} = 1 + 2v\sqrt{1 + u}$ , which is required to be solved for  $v$  (Wenger, 1987). The difficulty of this equation is to recognize this equation as linear in  $v$  and overcome the visual salience of the square roots in the equation (Kirshner & Awtry, 2004). This requires identifying parts of the expression as units, an ability that is referred to as the Generalized Substitution

Principle (Wenger, 1987). Recognizing the special pattern in expressions is part of structure sense.

The flexible approach is also mentioned by Sfard & Linchevski (1994b), who argue that flexible manipulation skills can be seen as a function of the versatility of available interpretations, and the adaptability of the perspective. In their view, these abilities are part of a structural mode of thinking. In Chapter 4 we will further elaborate this issue.

The notion of structure sense fits into the framework used to construct the assessments in Chapter 2. This framework used two perspectives on algebraic proficiency: the relation between basic skills and symbol sense, and the transition from arithmetic to algebra. As we have argued above, structure sense plays a role in both perspectives. Therefore, we use structure sense as a theoretical lens to investigate algebraic proficiency.

### 3.3 Student performance in structure sense

The background of this chapter is a study of the development of algebraic proficiency in Dutch pre-university education. To monitor this development, we constructed tests based on two perspectives of algebraic proficiency: the relation between basic skills and symbol sense, and the transition between arithmetic and algebra, see Chapter 2. The analysis showed that students make progress, both cross sectionally and longitudinally, but this progress is small. The Rasch analysis indicated that the growth of student development in algebraic proficiency did not lie in the area of the majority of the tasks. Part of these tasks have as a common difficulty that a keen eye for the mathematical structure of the expression is required. Therefore, we decided to analyze these tasks in greater depth.

We used the data of the research discussed in Chapter 2, in which we assessed students four times: the first assessment took place in March 2008; the second in May 2008; the third in October 2008; and the fourth in February 2009. The assessments consisted of paper and pencil tests. During the tests, students were not allowed to use calculators or notes. In total, 1020 students participated in at least one assessment. In order to monitor the development of individual students, we used an anchor design. To evaluate the students' results, a Rasch analysis was performed (Rasch, 1980; Bond & Fox, 2007). The Rasch model is a one parameter item response model. The Rasch model assumes that the probability  $P_{ni}$  of person  $n$ 's answering item  $i$  correctly is a logistic function of the difference between that persons' ability  $B_n$  and the difficulty of the item  $D_i$ :  $P_{ni} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}}$ . Rasch provides a common scale for items and persons. However, each assessment created one Rasch scale for the items and persons of that particular assessment. Now the resulting four Rasch scales are connected by anchor items, so that as a result, students and items are located on one and the

same Rasch scale. The unit of Rasch scale is the logit. We consider a probability of 0.80 of answering an item correctly as an expression of mastering that item (see Chapter 2). This corresponds to a difference between  $B_n$  and  $D_i$  of 1.39 logit, as 1.39 is the solution of  $B_n - D_i$  in the expression  $P_{ni} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}} = 0.80$ . The focus of this chapter is to investigate student development in relation to the notion of structure sense. Therefore, we analyze in greater detail student ability based on the tasks of the four assessments in relation to the Rasch measures of the tasks that appeal to structure sense. Drawing on the above discussion of structure sense, we selected 16 tasks from the four assessments that appeal to recognizing the algebraic structure of an expression or seeing a part of an expression as a unit. Ten of these tasks concern algebraic tasks and six concern numerical tasks (for an overview, see Appendix B).

An example of an algebraic task in which the recognition of the mathematical structure plays a role is the calculation of  $-6 - (5 - 4) \cdot (-8) - 3$ . In this task, the algebraic structure  $-6 - [(5 - 4) \cdot (-8)] - 3$  is covered by the brackets and the minus signs. This hinders the recognition of the structure. A non-example is the task of rewriting  $-4(3a + b)$ . Although an answer on this task such as  $-4(3a + b) = -12a + b$  can be seen a lack of structure sense, we argue that this task primary appeals to applying just the distributive rule, whereas the former task requires recognizing the structure and then applying the rules.

As we discussed above, the four assessments have served as input for the construction of one Rasch scale on which students and tasks have a measure. Figure 3.1 provides an overview of student measures of the fourth assessment in February 2009, and the task difficulties of selected tasks corresponding to a probability of 0.80 to answer the task correctly. For clarity's sake, in case of similar tasks, we included only one task in the Figure. Seven tasks were left out in this way, resulting in nine tasks in the Figure.

Central in this Figure is the horizontal axis with logits as units. The gray scaled bars above the axis represent student ability in percentiles in the fourth assessment in February 2009. The bars below the horizontal axis represent the difficulty of the tasks on the Rasch scale. The left-hand side of a bar corresponds to a probability of 0.50 of answering the corresponding task correctly, this is usually referred to as the Rasch measure of the task; the right-hand side of the bar corresponds to a probability of 0.80 of answering that task correctly. In Chapter 2 we argued that in our view, a probability of 0.80 of answering a task correctly indicates that this task has been mastered.

Tasks for which the corresponding bars lie on the left-hand side of the dashed line are mastered by at least 50% of the grade 12 students. From Figure 3.1 it follows that the task “calculate  $3 \cdot (-2) \cdot 5 - 2 \cdot 5$ ” is the only task that is mastered by more than 50% of the grade 12 students. The other eight tasks in the Figure are mastered by less than 25% of the grade 12 students.

Below we discuss student performance on the tasks presented in Figure 3.1.

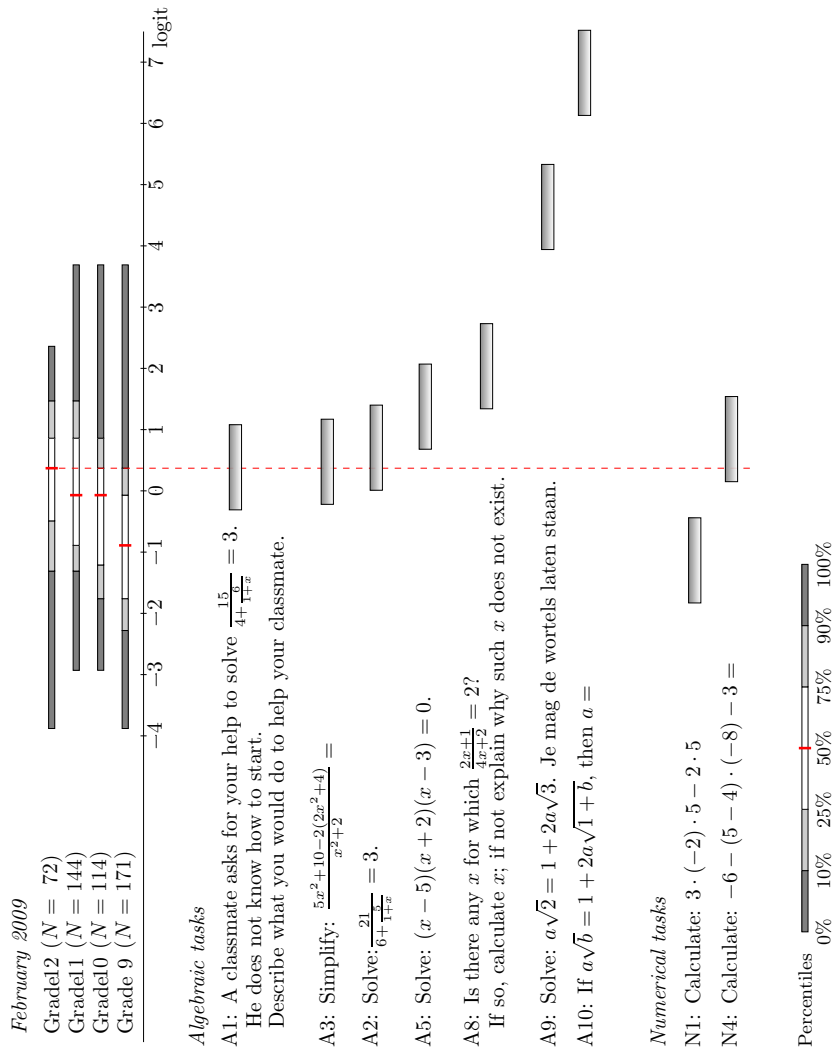


Figure 3.1: Student percentiles of February 2009 and structure sense tasks on the Rasch scale.

### 3.3.1 Algebraic tasks

One of the algebraic tasks presented in Figure 3.1 involves simplification of an expression, the other six algebraic tasks concern solving equations. Below, we discuss the student performance on each task.

#### Fraction in equation

In the first assessment (March 2008), students were asked to report the first clue they would provide a classmate to help solve the equation  $\frac{15}{4 + \frac{6}{1+x}} = 3$ . The Rasch measure of this task, A1, is  $-0.31$  logit (probability of success 0.50). The way the task was formulated left room for answers such as “I would call the teacher for help.” This kind of answer revealed problems in the dichotomous coding of student answers. To avoid this kind of answer, a formal version, A2, was included in the third assessment. The Rasch measure of the formal version is 0.01 logit, so the tasks did not differ much.

To solve these equations, different strategies can be used. For example, students could use the cover-up method and cover the denominator. In this way, the equation becomes  $\frac{15}{\square} = 3$ , which is easily solved. Following this line of thought, the next step would be  $4 + \Delta = 5$  which implies  $\frac{6}{1+x} = 1$ . This equation in turn could be solved by covering the denominator, thus yielding  $\frac{6}{\diamond} = 1$ , which implies  $1 + x = 6$ . An example of a more formal strategy is to multiply the numerator and the denominator both with  $1 + x$ , or multiply both sides of the equation with the denominator  $4 + \frac{6}{1+x}$ . These strategies have in common that students have to identify a part of the equation (the denominator, or a part of the denominator) as an object, which is seen as an expression of structure sense.

#### Simplification

In the second assessment (May 2008), we asked students to simplify the expression

$$\frac{5x^2 + 10 - 2(2x^2 + 4)}{x^2 + 2}.$$

The Rasch measure of this task, A3, is  $-0.22$  logit (probability of success 0.50). A similar task, A4, with Rasch measure  $-0.50$ , was included in the fourth assessment (February 2009). This formula has the structure  $\frac{5\square - 4\square}{\square}$  which leads to  $\frac{x^2+2}{x^2+2}$ . But simply expanding the brackets and taking similar terms together yields the same. In both cases the next step is to recognize that this fraction consists of two similar expressions that are divided, so the fraction yields one.

An example of the difficulties students experienced is presented in Figure 3.2. This student rewrote the numerator  $5x^2 + 10 - 2(2x^2 + 4)$  as  $5x^2 + 8(2x^2 + 4)$  by erroneously taking 10 and  $-2$  together. To this student, the algebraic structure of the numerator was not clear. Another type of error is presented in Figure 3.3.

Vereenvoudig zo ver mogelijk:

$$\frac{5x^2 + 10 - 2(2x^2 + 4)}{x^2 + 2} =$$

$$\downarrow$$

$$\frac{5x^2 + 8(2x^2 + 4)}{x^2 + 2}$$

$$\downarrow$$

$$\frac{5x^2 + 16x^2 + 32}{x^2 + 2}$$

$$\downarrow$$

$$\frac{21x^2 + 32}{x^2 + 2}$$

Translation:

Vereenvoudig zo ver mogelijk: Simplify

Vereenvoudig zo ver mogelijk:

$$\frac{5x^2 + 10 - 2(2x^2 + 4)}{x^2 + 2} =$$

$$\frac{5x^2 + 10 - 4x^2 - 8}{x^2 + 2} =$$

$$\frac{x^2 + 2}{x^2 + 2} = 0$$

Figure 3.2: Work of a grade 8 student. Figure 3.3: Work of a grade 10 student.

This grade 10 student correctly simplified the expression to  $\frac{x^2+2}{x^2+2}$ . Then he concluded that  $\frac{x^2+2}{x^2+2}$  equals zero instead of one. This error might stem from an inability to see the algebraic structure of the expression, but other explanations such as a deep misunderstanding of fractions or the student being used to the form “expression = 0,” seem just as reasonable.

### Factorization

In the second assessment, students were asked to solve the equation

$$(x - 5)(x + 2)(x - 3) = 0.$$

This requires students to identify the underlying algebraic structure, which involves three factors on the left side of the equation, and the product of these three factors equals zero. The structure is  $A \cdot B \cdot C = 0$ , which implies  $A = 0$  or  $B = 0$  or  $C = 0$ . The Rasch measure of this task, A5, is 0.68 logit (probability of success 0.50). A similar task, A6, was included in the fourth assessment. This task was perceived as less difficult, with a Rasch measure of  $-0.60$  logit, probably due to a test-retest bias. The difficulty of these tasks can be explained by the students’ tendency to expand the brackets, after which they could not find the factorization. Some of the students came to see the error and recovered by giving the correct answer. Figure 3.4 shows such a work of a grade 12 student who started to solve the equation by expanding the brackets. In the second line of this work, he correctly wrote that

$$x^3 - 4x^2 + 3x^2 - 12x - x^2 + 4x - 3x + 12 = 0.$$

8. Los op:

$$(x-1)(x+3)(x-4) = 0.$$

$$(x^2 + 3x - x - 3)(x-4) = 0$$

$$x^3 - 4x^2 + 3x^2 - 12x - x^2 + 4x - 3x + 12 = 0$$

$$x^3 - 2x^2 + x + 12 = 0$$
~~$$x^3 - 2x^2 + x + 12 = 0$$~~

$$x = 1 \quad \vee \quad x = -3 \quad \vee \quad x = 4$$

Translation:

Los op: Solve

Figure 3.4: Work of grade 12 student.

Going to the next line, he forgot the factor  $-12x$ , so erroneously concluded that  $x^3 - 2x^2 + x + 12 = 0$ . In the next line, this is scratched, he started to factor out  $x$ . Apparently, then he realized that the solutions can be found more easily and wrote the correct solutions down.

The students' tendency to expand the brackets indicates the visual salience of the brackets. The underlying structure of  $A \cdot B \cdot C = 0$  is overlooked.

### Arcavi's Equation

As we have argued in Section 3.2, in our view, structure sense can be seen as a part of symbol sense. The ability to read through symbols described by Arcavi (1994) as an aspect of symbol sense, lies close to the ability to recognize the algebraic structure of a formula. Arcavi (1994) argues that resisting the impulse to immediately solve the equation  $(2x+3)/(4x+6) = 2$ , and instead to try to read meaning into the symbols, is an expression of symbol sense. In this example, the ability to read through the symbols requires that students recognize the expression  $(2x+3)/(4x+6)$  as an expression of the form  $\frac{A}{2A}$ . Therefore, it is necessary for students to see  $2x+3$  as a unit. However, realizing that  $(2x+3)/(4x+6)$  equals  $\frac{1}{2}$  is not necessary for the solving process. Just solving the equation and finding the solution  $x = -1\frac{1}{2}$  and realizing that the denominator equals zero, is a correct procedure for solving this equation. So, reading through the equation and seeing the structure is practical, but not necessary.

The ability to see  $\frac{A}{2A}$  is a manifestation of more symbol sense than finding the solution for which the denominator equals zero and then concluding that the

equation does not have a solution. This raises questions about the relation between symbol sense, operationalized by the strategy used, and a student's level of algebraic proficiency, operationalized by the measures on the Rasch scale. Since more symbol sense is supposed to be an expression of a higher level of algebraic proficiency, we expect students with more symbol sense to have a higher Rasch measure, and thus we expect students who see  $\frac{A}{2A}$  to have a higher Rasch measure than students who found the solution for which the denominator equals zero. To investigate the relation between these strategies and students' proficiency, we performed a qualitative analysis of the written answers of the students. This analysis is performed on the written answers to the task presented in the fourth assessment, because in contrast to the similar task in the first assessment, this task was presented to students of all grades.

Because we were interested in the strategies used to solve this equation, we restricted the analysis to the correct answers. The analysis yielded three categories of students. First, students that argued that the quotient equals  $\frac{1}{2}$ . These students recognized the structure  $\frac{A}{2A}$  in the equation. The second category contains students who found the solution for which the denominator equals zero. These students correctly argued that this solution is no solution. The third category contains students who gave another argument. For example, these students only wrote "no." Another example is a student who argued that  $(2x + 1)/(4x + 2)$  equals a fraction between 0 and 1. From the 501 participating students in the fourth assessment, 105 gave a correct solution. Half of these students used the strategy of the first category. The other students were nearly evenly distributed between the other categories. The relation between strategy and ability is presented in Figure 3.5. Again, in the center of the Figure is the Rasch scale of proficiency. The gray scaled bars below the axis represents percentiles of students' Rasch measures in the fourth assessment in February 2009. The dots above the axis represent students in a particular grade with a particular strategy.

From the figure we see that, although more students solve the task correctly using the strategy "quotient equals  $\frac{1}{2}$ " than using the strategy "denominator zero," the Rasch measure of the students using the former strategy is not higher than those using the latter. In other words, the strategy that is supposed to be a manifestation of more symbol sense does not imply a higher Rasch measure. In our view, this means that either the relation between strategy and symbol sense is not as strict as the literature suggests, or the relation between symbol sense and the underlying latent variable of the Rasch scale is troublesome. In the former case, the strategy students use might depend on the strategy they think they are expected to use. From this point of view, the more broadly applicable strategy "denominator zero" might be viewed as more valuable, because this strategy could also be used in case the equation had been  $(2x + 1)/(3x + 2) = \frac{1}{2}$ . However, since the number of students in this group and the differences between the categories are small, this single case cannot justify radical conclusions, but further research on the relation between strategies, symbol sense and algebraic



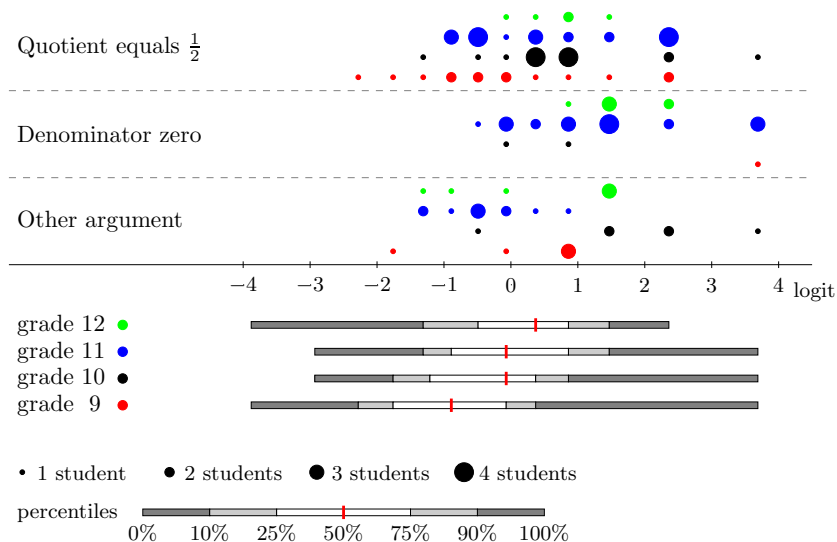


Figure 3.5: Students' ability combined with the strategy used.

proficiency seems appropriate.

### Wenger's Equation

The two tasks with the highest difficulty, A9 and A10, are adapted versions of Wenger's Equation. These two tasks were far above the ability of all students, see Figure 3.1. In the first assessment, we included Wenger's Equation in which we replaced the letters  $u$  and  $v$  for readability with the letters  $a$  and  $b$ . We did not ask students to solve  $a\sqrt{b} = 1 + 2a\sqrt{1+b}$  for  $a$ , because Dutch students are not familiar with this kind of question. Instead, we asked students to rewrite  $a\sqrt{b} = 1 + 2a\sqrt{1+b}$  as an expression of the form  $a = \dots$ . In the analysis we found that this way of asking is ambiguous to students. For example, students divided both sides of the equation by  $\sqrt{b}$ , which yields to  $a = \frac{1+2a\sqrt{1+b}}{\sqrt{b}}$ . This is not the kind of answer we intended to see, but somehow meets the purport of the question.

As we argued in Section 3.2, recognizing the linear structure of this equation is crucial for solving it. The ability to see the linear structure requires students to sense the symbols as arranged in a special pattern, which is an expression of structure sense. Exactly two students out of the 650 of grades 9, 10 and 11 were able to solve this equation. These two students recognized the linear form of the equation and gave the correct answer,  $a = \frac{1}{\sqrt{b}-2\sqrt{1+b}}$ .

In the third assessment, an adapted version of the equation was included. In

this version, we substituted  $b = 2$ , so the equation that had to be solved became  $a\sqrt{2} = 1 + 2a\sqrt{3}$ . Exactly 19 out of the 653 participating students in the third assessment were able to solve this task.

The difficulty of Wenger's Equations is explained by students' inability to recognize the linear form of the equation. Wenger (1987) found that students are able to perform manipulations correctly, but these manipulations do not lead to a solution. Rather, students go round in circles, create more complex expressions and then reduce these terms. The square roots that serve as coefficients in this equation invite students to square both sides of the equation. The unfamiliarity of students with square roots as coefficients can enhance the visual salience of square roots (Kirshner & Awtry, 2004).

We agree that recognizing the linear form is crucial for solving the equation. However, in our view, there are other hurdles. First, the unconventional sequence of symbols in  $a\sqrt{2} = 1 + 2a\sqrt{3}$ . In this equation, the variable is followed by the numerical coefficient in the left-hand side of the equation. In the right-hand side, the variable is in the middle of the numerical coefficient. This is an unusual sequence of symbols for Dutch students, because usually, the numerical coefficient precedes the variable. In this way, the equation would have been  $\sqrt{2}a = 1 + 2\sqrt{3}a$ . We realize that this is only a slight difference for an expert. But we believe that to students, this change of order perhaps makes the difference between being able or not being able to solve the equation.

The second hurdle concerns the different roles of the variables  $a$  and  $b$  in the equation. The  $a$  in the equation serves as unknown, whereas the  $b$  serves as a variable. The versatility of the use of variables is a well known difficulty in mathematics and has been studied by many researchers (e.g. Matz, 1982; Janvier, 1996; Rosnick, 1981; Wagner, 1983; Trigueros & Ursini, 1999; Drijvers, 2003; Ursini & Trigueros, 2001; Schoenfeld & Arcavi, 1988). The skill of flexibly dealing with the different roles of variables can be seen as part of a broad view on equations. The results for these tasks suggest that students do not have such a broad view. We elaborate on this in Chapter 6.

Summarizing, we argue that Wenger's Equation presents students with several difficulties. These difficulties all concern the ability to recognize the linear structure of the equation which can be seen as a part of structure sense. The students' performance on the two linear equations does not allow us to conclude which of these difficulties is paramount, but in our view, these equations all require structure sense.

### 3.3.2 Numerical tasks

In the analysis we have included two types of numerical tasks (see also Figure 3.1). Below we elaborate on both types.

9. Bereken:

$$-6 - (5 - 4) \cdot (-8) - 3 =$$

$$5 - 4 = 1$$

$$-6 - 1 \cdot -8 - 3 =$$

$$-7 \cdot -11 =$$

$$77$$

Translation:

Bereken: Calculate

Figure 3.6: Work of a grade 10 student.

3. Bereken:

$$-6 - (5 - 4) \cdot (-8) - 3 =$$
~~$$-6 - (5 - 4) \cdot (-8) - 3 =$$~~

$$= -6 - 1 \cdot (-8) - 3$$

$$= -7 \cdot (-8) - 3$$

$$= 56 - 3$$

$$= 53$$

Figure 3.7: Work of a grade 10 student.

### Product minus product

In the first and fourth assessment, we included the tasks N1:  $3 \cdot (-2) \cdot 5 - 2 \cdot 5$  and N2:  $2 \cdot (-3) \cdot 5 - 2 \cdot 5$ , with Rasch measures  $-1.83$  and  $-1.59$  respectively (probability of success 0.50). These two tasks differ only slightly and are mastered by almost 75% of the grade 12 students, see Figure 3.1. A conceivable error in this task would be to calculate first  $5 - 2$ , leading to the expression  $3 \cdot (-2) \cdot 3 \cdot 5 = -90$ . The majority of students avoided this error and thus recognized the structure of the formula as “a product minus another product.”

### Brackets and minus

The other four tasks have the structure  $-6 - (5 - 4) \cdot (-8) - 3$ , with different numbers, see Appendix B for the full list. These four tasks have functioned as anchor items in the Rasch analysis. As a consequence, these tasks all have equal Rasch measures of 0.15 logit. From Figure 3.1 it follows that this task is mastered by 10% of the grade 12 students. So these tasks are perceived as more difficult than the two tasks above.

The difficulty with these tasks can be explained by the student’s inability to see the algebraic structure of the expression. Seeing the structure is hindered by the dot in the middle. Students tend to see the structure  $A \cdot B$ , as is illustrated in the work of a grade 10 student in Figure 3.6. This student calculated  $-6 - (5 - 4)$  and  $(-8) - 3$  separately, yielding  $-7$  and  $-11$ , and then multiplied both numbers. An alternative explanation could be that this student did not know the right-of-way rule for operations. Perhaps she thought that subtracting comes before

multiplying, but we deem this explanation less likely in view of the results of students for the other numerical task above.

A variant of this error is made by the student in Figure 3.7. She also simplified the left-hand side of the dot to  $-7$ . Then she multiplied by  $-8$  and subtracts 3, yielding 53. Both students were distracted by the visual salience of the dot.

### 3.4 Conclusions and discussion

In the previous sections, we reported the results of students in Dutch secondary education. In Chapter 2 we concluded that the development of the majority of the students is not in the area of the conceptual aspects of algebraic proficiency. One of these aspects is structure sense (Linchevski & Livneh, 1999). Structure sense in high school algebra consists of several abilities including the ability to see the algebraic structure of a formula and the ability to see a part of an expression as a unit (Hoch & Dreyfus, 2004, 2006; Novotnà & Hoch, 2008). We argued that structure sense can be seen as a part of symbol sense (Arcavi, 1994). The results indicate that the range which the majority of the students mastered did not include tasks related to structure sense. In general we may conclude that the majority of the students were not able to deal flexibly with the mathematical structure of expressions. This is the case with both the numerical tasks and the algebraic tasks.

Having structure sense is seen as a *sine qua non* for having algebraic proficiency. Novotnà & Hoch (2008) argue that attention to algebraic structure in secondary education is important because teaching mathematics as a set of procedures and algorithms can lead to a lack of understanding of the conceptual structure, thus resulting in an inability to solve real mathematical problems. Arcavi (1994) argues that symbol sense is at the heart of what it means to be competent in algebra. In his view, instruction that supports the development of symbol sense should support and encourage taking the time to talk about issues regarding symbol sense instead of focusing on finding the answer.

Structure sense also includes the ability to choose appropriate manipulations that make the best use of the structure (Novotnà & Hoch, 2008). This ability is also mentioned by Arcavi (1994) as an expression of symbol sense. Our analysis did not yield a direct relation between success on items that express more structure sense and the latent variable of the Rasch scale. This indicates a more severe problem: the amount of structure sense can not be judged on one strategy used for just one task alone. Structure sense is described as a set of abilities and must be judged as such. Therefore structure sense is not easy to teach: it is not just a matter of exhibiting one particular strategy, but more so a matter of exhibiting a whole range of abilities. This range of abilities includes a broad and flexible view on equations, in which square roots can serve as coefficients, in which a fraction can be inside another fraction, in which equations may have multiple solutions, or none at all. The results of this chapter have shown that

most students did not have such a broad view on equations. In Chapter 4 and Chapter 6 we elaborate on this point.

## Chapter 4

# Operational–structural transitions as indicators of difficulty in secondary school algebra tasks

In this chapter we use Sfard’s notion of operational and structural conceptions as a theoretical lens, in order to explain the results of a study on the development of algebraic proficiency. This leads to the hypothesis that transitions between operational and structural conceptions can be seen as a complexity factor in solving algebraic problems. This hypothesis was operationalized in a test that was presented to 92 students. The results corroborated the hypothesis.

### 4.1 Introduction

Discussions of the level of algebraic proficiency of secondary school students nowadays take place worldwide. In the Netherlands, this discussion induced a study on the development of algebraic proficiency in secondary education (see Chapter 2). The results of this study revealed that student development did not cover the conceptual aspects of algebraic proficiency such as recognizing mathematical structures in formulas or in seeing a part of an expression as an object; such tasks were beyond the capabilities of the majority of the students.

This chapter seeks to explain these results by concretizing and operationalizing Sfard’s theory of reification. In this theory, the ambiguous nature of mathematical concepts plays a central role (Sfard, 1991; Sfard & Linchevski, 1994b;

Sfard, 1995). Their nature is seen as ambiguous because for example one and the same mathematical expression refers to different entities (Tall & Thomas, 1991; Byers, 2007). For example, on the one hand the expression  $2x + 3$  can be seen as an addition in which  $2x$  and  $3$  have to be added. On the other hand, the expression  $2x + 3$  can be seen as the result of the addition. In this case,  $2x + 3$  is an entity, an object. Whether  $2x + 3$  has to be perceived as a process or as an object depends on the specific situation in the process of problem solving. It is important that students are able to switch between different modes of thinking.

Sfard & Linchevski (1994b) identified two transitions in which this ambiguity plays a role: the process–product duality, and the algebra of a fixed value or as variable. Operationalization of these transitions leads to the hypothesis that transitions between operational and structural conceptions might serve as a complexity factor in solving algebraic problems. This hypothesis is tested among 92 students of grade 11.

The outline of this chapter is as follows. Below, we first elaborate on the research context that lead to the results we want to explain. Next, we elaborate on Sfards’ theory of reification and discuss operational and structural conceptions related to the transitions between these two. Finally, we discuss the results of the role of these transitions as a factor of complexity by testing this hypothesis in a small setting.

## 4.2 Research context

The tasks we mentioned above are part of a study on the development of algebraic proficiency in Dutch pre-university education. To monitor this development, we constructed tests, based on two perspectives of algebraic proficiency: the relation of procedural fluency and conceptual understanding and the relation of arithmetic and algebra, see Chapter 2.

We assessed students four times, spread out over one calendar year. The assessments consisted of paper and pencil tests. During the tests, students were not allowed to use calculators or bring on notes. In total, 1020 students participated at least in one assessment. In order to monitor the development of individual students, we used an anchor design. To evaluate the results, a Rasch analysis was performed (Rasch, 1980; Bond & Fox, 2007). The Rasch model supposes that the probability  $P_n$  of a person  $n$ ’s answering an item  $i$  correctly is a logistic function of the difference between the student’s ability  $B_n$  and the difficulty of the item  $D_i$ :

$$P_{ni} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}}.$$

Rasch provides a common scale for items and persons. Each assessment created one Rasch scale. The four Rasch scales are connected with anchor items. As a result, the students and the items are located on one and the same Rasch scale with the unit being the logit, see Figure 4.1.

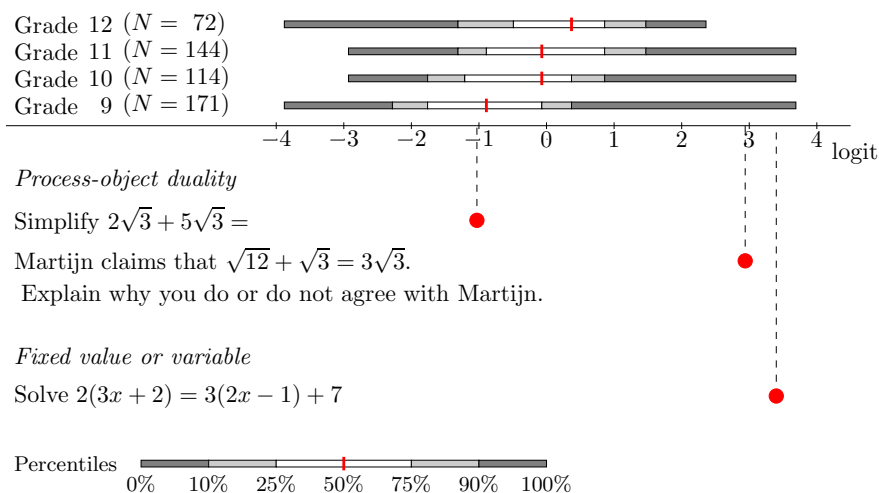


Figure 4.1: Percentiles of students and tasks on Rasch scale.

We consider a probability of 0.80 of answering an item correctly as an expression of mastering that item, see also Chapter 2. This corresponds to a difference between  $B_n$  and  $D_i$  of 1.39 logit because 1.39 is the value of  $B_n - D_i$  which satisfies the equation  $P_{ni} = \frac{e^{B_n - D_i}}{1 + e^{B_n - D_i}} = 0.80$ . In Figure 4.1, students with the same measure as a task have a probability of 0.80 to answer that task correctly. The analysis showed that the students made progress, both cross sectionally and longitudinally, but this progress is small. Further, only a few items were mastered by the majority of the students. More specifically, the analysis revealed a remarkable difference in performance on tasks concerning calculating with square roots. The task “simplify  $2\sqrt{5} + 4\sqrt{5}$ ” is mastered by almost 90% of the students, whereas the task of answering the question “do you or do you not agree that  $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ ” is mastered by less than 10% of the students of grade 12. Further, the task in which students had to solve the linear equation  $2(3x + 2) = 3(2x - 1) + 7$  was mastered correctly by less than 10% of the grade 12 students, see Figure 4.1. In our view, these tasks require the transitions identified by Sfard & Linchevski (1994b) concerning the process-product duality and the value-variable duality.

### 4.3 Sfard's Theory of Reification

Central in the theory of reification is that in mathematics, abstract concepts can be conceived in two fundamentally different ways: structurally as objects and operationally as processes. Sfard's theory of reification provides a com-



bined ontological-psychological theoretical framework for advanced mathematical thinking.

The evidence for the ontological dimension of this theory lies in the historical development of mathematics. As an example, Sfard (1991) illustrates the development of the concept of number. The origins of this concept lie in the process of counting. For years, the concept of number was closely related to this process. An important step in the development of the concept of number was the discovery of the incommensurability of square roots such as  $\sqrt{2}$ . Disconnecting numbers from measurability provides mathematicians with the opportunity to broaden the concept of numbers to even complex numbers such as  $\sqrt{-1}$ . The development of the concept of number illustrates the formation of mathematical knowledge. According to Sfard (1995), the historical development of algebra can be characterized as a long sequence of transitions from operational to structural conceptions. Again and again, processes performed on certain abstract objects at one level turn into new objects that in turn serve as the objects of new higher level processes.

The psychological foundation of the theory of reification can be found in the framework of genetic epistemology of Piaget (1970). The role of processes and objects in mathematical thinking can also be found in theoretical extensions of Piaget's reflective abstraction (e.g., Dubinsky, 1991).

Based on these foundations, Sfard argues that in mathematics, one can distinguish two kinds of entities: abstract objects and computational processes. For example a mathematical function can be seen as a method of computing values by performing some computational process. In this view, a function is a kind of recipe to get the output of a given input like “multiply the input by 3 and then add 2.” On the other hand, a function is a set of ordered pairs with certain characteristics such that, given the first element, the second element of the pair is unique. In the example above, a more formal approach could be  $f$  from  $\mathbb{R} \rightarrow \mathbb{R}$  with  $f : x \mapsto 3x + 2$  as an object with characteristics such as continuity and differentiability. Although abstract objects and computational processes look very different, in fact, they are “different sides of the same coin” (Sfard, 1991, p. 1). This means that the abstract objects are an alternative way of referring to computational processes. For instance any number may be conceived operationally, as a process, or structurally, as an object.

In the process of concept formation, three stages are distinguished: interiorization, condensation, and reification. These stages correspond to three degrees of structuralization based on a theoretical analysis of the relation between processes and objects. In the first stage, interiorization, the learner gets acquainted with a process, for instance a function as a recipe. A process is interiorized if this process does not necessarily have to be carried out in order to be analyzed. In the second phase, condensation, a person becomes more capable of thinking about a process as a whole. There is a growing easiness in which mathematical entities are seen as different representations of the same concept. The third

phase, reification, is defined by Sfard as an ontological shift. Whereas interiorization and condensation are gradual processes, the phase of reification is not gradual, but “a sudden ability to see something familiar in a totally new light” Sfard (1991, p.19). Reification is the point where the object turns into input for interiorization on a higher level.

By definition, these three phases are hierarchical. One can not reach the next stage without having taken the former steps. This is built on the conjecture that the operational conceptions precede the structural conceptions. This conjecture is based on theoretical as well as empirical considerations on the historical development of mathematics and algebra in particular.

According to Sfard, there are two reasons why it is difficult to reach the stage of reification. The first reason is the so called vicious circle of reification. For example,  $3 - 5$  must be treated as an object before one is able to conclude that  $3 - 5$  is a legitimate mathematical object which obeys rules similar to the rules of natural numbers. The second reason is that people have to free themselves from certain deeply rooted convictions before the new abstract object can be accepted (Sfard, 1995). An example is the conviction that all numbers signify amounts.

## 4.4 Operational and structural conceptions

The most important observation in the theory of reification is that the same mathematical concept may sometimes be interpreted operationally, as a process, and sometimes structurally, as an object. The duality of mathematical entities and the crucial role of this duality in the process of concept formation is also addressed by Gray & Tall (1992). They introduced the term *procept* to refer to “a combined mental object consisting of a process, a concept produced by that process, and a symbol which may be used to denote either or both.”

In Sfard’s theory of reification, abstract objects such as different kinds of numbers or functions mediate between the operational and the structural conceptions. These abstract objects emerge as lower level processes become subject to manipulations on a higher level. In the stage of reification, abstract objects serve as links between the operational and structural conceptions. With respect to a given concept, processes on the lower level are referred to as primary processes, whereas processes on the higher level are referred to as secondary processes.

For example, in the concept formation of numbers, the process of the division of an integer by another integer is a primary process. Calculating with fractions is a secondary process linked to the lower level process of dividing an integer by an integer. In the chain of transitions from operational to structural conceptions, secondary processes on one level might act as primary processes on the next level. Structural conceptions are a result of the reification of the underlying primary process.

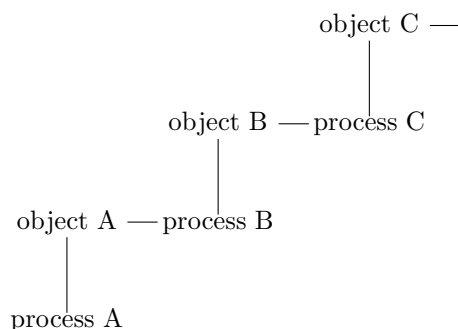


Figure 4.2: Development of mathematical concepts as transitions from operational to structural conceptions (Sfard and Linchevski, 1995).

In order to perform secondary processes, students have to be able to link the secondary process to the primary processes because secondary processes are grounded in primary processes. Sfard & Linchevski (1994b) argue that without the link to the primary processes, secondary processes seem arbitrary and become easily forgotten. The core in the fluency of algebra is seen as the capacity to make transitions from one level to another, the capacity to quickly alternate between operational conceptions and structural conceptions. To Sfard & Linchevski (1994a), an important component of algebraic proficiency is the flexible approach in which a student is able to switch between an operational approach in the focus is on the processes, and a structural approach in which the focus is on the objects. These switches play a role in the process–product duality and the value–variable duality. In the next section, we elaborate on these transitions.

## 4.5 Operational–structural transitions as indicators of complexity

According to Sfard & Linchevski (1994a), the power of algebra lies in the possibility of performing processes on the secondary level, without constantly making the link to the underlying primary process. The ability to perform processes in an unthinking mode is what algebra makes a powerful tool, as Whitehead (1924, p. 61) put it:

“It is a profoundly erroneous truism, repeated by all copy-books and by eminent people when they are making speeches, that we should cultivate the habit of thinking of what we are doing. The precise opposite is the case. Civilization advances by extending the number of

important operations which we can perform without thinking about them. Operations of thought are like cavalry charges in a battle—they are strictly limited in number, they require fresh horses, and must only be made at decisive moments.”

However, performing processes on the secondary level in a meaningful way requires the constant ability to return to the underlying primary process. The ability to make transitions from the primary to the secondary level and the reverse is seen as the core of algebraic fluency. For example, Arcavi (2005) argued that algebraic proficiency requires students to flexibly switch between meaningless actions (automatic application of rules and procedures) to sense making in the way of “symbol sense.” In other words, students have to be able to postpone meaning in order to apply procedures quickly. And also, to “unclog an automatism” as Freudenthal (1983, p. 469) put it, to reconsider, to reflect, or to question, if necessary.

As we have mentioned before, Sfard & Linchevski (1994b) identified two crucial transitions: the transition from a pure operational to a dual process–product approach, and the transition from the algebra of a fixed value to the algebra of a variable. Related to these two transitions, we argue that for students who have not reached the stage of reification, each transition from the primary processes in an operational mode of thinking to the secondary processes in a structural mode of thinking and the reverse is a difficult step in the solution process of a specific problem. For this purpose, we discuss student performance on tasks related to the first and second kinds of transition. This leads to the hypothesis that each transition can be interpreted as a contribution to the complexity of a task.

#### 4.5.1 Transition related to the process–product duality

In section 4.1, we discussed two tasks concerning square roots that showed a huge difference in student performance. The first task was to simplify the expression  $2\sqrt{3} + 5\sqrt{3}$ . The second task was to decide whether or not they agree with “ $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ .” The first task was mastered by the majority of the students; the second task was mastered by less than 10% of the students, see Figure 4.1. Without going into detail, we can state that there is a huge difference in performance on those two tasks. Searching for an explanation of this difference, we analyzed both tasks in greater depth in the light of the theory of reification.

In simplifying  $2\sqrt{3} + 5\sqrt{3}$ , it is important that students treat  $\sqrt{3}$  as a single entity or object. Students have to realize that estimating  $\sqrt{3}$  is not appropriate. Rather,  $\sqrt{3}$  has to be seen as an object itself of which there are two ( $2\sqrt{3}$ ) and then again 5 ( $5\sqrt{3}$ ) which makes 7 of these objects, in total,  $7\sqrt{3}$ . The main aspect of solving this problem is that students recognize and treat  $\sqrt{3}$  as an object. Once students do so, the further calculation is simple.

In the second task, the situation is different because here, it is not appropriate to see  $\sqrt{12}$  as a static object. In this case, students have to recognize that  $\sqrt{12}$  has to be manipulated. Next, students have to realize that

$$\sqrt{12} = \sqrt{4 \cdot 3} = \sqrt{4} \cdot \sqrt{3} = 2\sqrt{3}.$$

After this step, it is important that students perceive  $\sqrt{3}$  as an object of which there are two ( $2\sqrt{3}$ ) plus one ( $\sqrt{3}$ ), so that makes, in total,  $3\sqrt{3}$ .

In our view, in the example above, the addition of  $2\sqrt{3}$  and  $\sqrt{3}$  is an addition of objects, whereas the rewriting of  $\sqrt{12}$  as  $2\sqrt{3}$  requires the process of extracting a root. It is a transition from an object to a process. Hence, the transition from the expression  $\sqrt{12} + \sqrt{3}$  in which  $\sqrt{12}$  is an object to the calculating and treating of  $\sqrt{12}$  as a process is a transition in the process–product duality. This transition has to be made in the reverse order as students have to see  $\sqrt{3}$  as an object in calculating  $2\sqrt{3} + \sqrt{3}$ .

#### 4.5.2 Transition related to the fixed-value–variable duality

The second transition is related to the role of letters in algebra. Multiple interpretations of variables can be found by Ursini & Trigueros (2001) who made a decomposition of ‘variable’ related to its main uses in elementary algebra: specific unknown, general number, and relationship between variables. Drijvers (2003) distinguishes five different roles related to different approaches in algebra: the variable as placeholder for numerical values; as changing quantity in functional algebra; as generalized number in the algebra of pattern and structure; as an unknown in problem-solving; and as a symbol in algebraic language. Because of the different meanings letters can have, one of the main difficulties of the concept of variable is to deal with this versatility of meanings in a flexible way. This flexibility is important because the meaning of variables can differ while solving the problem (e.g. Arcavi, 1994). In the theory of reification, Sfard & Linchevski (1994b) make a distinction between the algebra of fixed values and functional algebra where letters are seen as variables.

Related to this distinction, we discuss a task in which students have to solve the equation  $2(3x + 2) = 3(2x - 1) + 7$ . Expanding the brackets in this equation leads to the tautological expression  $6x + 4 = 6x + 4$  or, equivalently,  $0 = 0$ . This tautology is true for every value of  $x$ , so the correct answer is that the equation holds for every  $x \in \mathbb{R}$ . Less than 10% of the students mastered this task, see Figure 4.1.

The rewriting of the expression  $2(3x + 2) = 3(2x - 1) + 7$  as the expression  $6x + 4 = 6x + 4$  is a secondary process. In the view of the theory of reification, algebraic manipulations at the secondary level are inherited from the underlying numerical calculation. The legitimacy of the manipulation is grounded in the fact that whatever number is substituted for  $x$ ,

$$2(3x + 2) = 3(2x - 1) + 7 \iff 6x + 4 = 6x + 4.$$

So, for the manipulation of the equation, treating  $x$  as an unknown but fixed value, suffices. Mostly, the manipulations in the standard procedure of solving equations on the secondary level lead to an expression of the form “ $x = \text{number}$ ” for the students. The exact meaning of this final formula is not an issue, what is important is that the formula “ $x = \text{number}$ ” indicates the end of the solving procedure. However, reaching the tautology  $6x + 4 = 6x + 4$  or  $0 = 0$  does not fit into the algebra of fixed values. At this point, a transition is needed to the algebra of variables in which  $x$  can have different values in order to be able to conclude that  $2(3x + 2) = 3(2x - 1) + 7$  holds for every  $x \in \mathbb{R}$ .

### 4.5.3 Operational–structural transitions as indicators of difficulty

Above, we used the theory of reification to explain the difficulty of manipulating square roots and linear equations. This difficulty is explained by the necessity for students to make a transition in the process–product duality and the value–variable duality. The flexibility involved in being able to deal with this duality is an important aspect of algebraic proficiency. However, this stage is inherently difficult and out of reach for many students. In our view, to students who have not reached the stage of reification, each transition between the operational and structural approach is difficult. Each time the solving of a specific problem requires a switch from process to object or the reverse, or from the algebra of fixed values to the algebra of variables or the reverse, a difficulty is created. Based on these considerations and the results discussed in the previous section, we formulate the following hypothesis.

*Hypothesis:*

The more transitions are needed for solving an algebraic problem, the more difficult this problem is.

With this hypothesis we are able to explain the results of students on the square root tasks and the linear equation. However, we want to use this hypothesis as a predictor for the difficulties students experience while solving algebraic problems. Therefore, we decided to test the hypothesis in a small setting.

## 4.6 Testing the hypothesis

In order to test the hypothesis formulated above, we constructed a test. These tasks lie close to the tasks we discussed in the previous section and are related to the process–product duality and the value–variable duality.

### 4.6.1 Test construction and data collection

Regarding square roots, we constructed two tasks that narrow the gap between the two tasks we discussed in the previous section (task R1 and task R4 in the list below). In the second task, the square in the root has been made visible. This might bring students to recognize  $\sqrt{9 \cdot 2}$  as the square root that needs to be manipulated. In the third task, the roots are split up already, so students only have to extract  $\sqrt{4}$ .

- R1 Harry claims that  $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ .  
Explain why you do or do not agree with Harry.
- R2 Sophie claims that  $\sqrt{9 \cdot 2} + \sqrt{2} = 4\sqrt{2}$ .  
Explain why you do or do not agree with Sophie.
- R3 Thomas claims that  $\sqrt{4} \cdot \sqrt{7} + \sqrt{7} = 3\sqrt{7}$ .  
Explain why you do or do not agree with Thomas.
- R4 Simplify:  $2\sqrt{5} + 4\sqrt{5} =$

Students that are able to answer the first task correctly might be in the reification phase. Therefore, we expected that students would perform better on higher numbered tasks than on lower numbered tasks. Further, we hypothesized that a correct answer to a lower numbered task implies a correct answer to a higher numbered one.

As regards the value–variable duality, we asked students to solve three linear equations. The first equation in the list below is exactly the same as the one we discussed in Section 4.5 and holds for all  $x \in \mathbb{R}$ . The second equation transforms to  $6x + 4 = 6x + 2$ , which has no solutions; and the third equation has exactly one solution:  $x = 1$ .

- E1 Solve:  $2(3x + 2) = 3(2x - 1) + 7$ . Show your work.
- E2 Solve:  $2(3x + 2) = 3(2x - 1) + 5$ . Show your work.
- E3 Solve:  $2(3x + 2) = 2(2x - 1) + 8$ . Show your work.

In our view, a correct answer to the first task indicates the ability to deal with the value–variable duality. Therefore we hypothesized that students would perform better on higher numbered equations than on the lower numbered ones. Further, we hypothesized that a correct answer to a lower numbered equation implies a correct answer to a higher numbered equation. Finally we hypothesized that incorrect answers to the first and second equation are not due to faulty simplifications of the expressions (secondary processes). Rather, incorrect answers are due to the inability to draw the correct conclusion from expressions such as

Stream	$N$	R1	R2	R3	R4	E1	E2	E3
Social	28	0.11	0.14	0.21	0.50	0.00	0.14	0.68
Science	64	0.59	0.70	0.78	0.92	0.16	0.38	0.81
Total	92	0.45	0.53	0.61	0.79	0.11	0.30	0.77

Table 4.1:  $p$ -values of social and science students on square root (R1–R4) and equation (E1–E3) tasks.

$6x + 4 = 6x + 4$  (first equation) and  $4 = 6$  (second equation), the inability to relate the secondary process to the underlying primary process.

We chose to present the seven tasks discussed above as open questions that had to be worked out with paper and pencil to grade 11 students. We wanted to present the tasks to experienced students because we are interested whether the hypothesis holds for students who have almost finished secondary school. To avoid overloading the students in the highest grade, we chose grade 11. In order to keep the sample representative for secondary education, we decided to include students of both the social stream and the science stream (see Chapter 2). In total, 92 students of two schools situated in the south of the Netherlands participated. Students were asked to motivate their answers. Further, students were not allowed to use calculators or bring on notes.

#### 4.6.2 Results

First, we performed a coarse-grained analysis on the written answers. Correct answers were coded with 1, incorrect answers were coded with 0. The average performance of the students is presented in Table 4.1. In line with the hypothesis, the  $p$ -values of both series of tasks (square roots and linear equations) are increasing. Below, we elaborate on the performance scores of both series of tasks. In addition, we performed a qualitative analysis on both series in order to find out whether the incorrect answers can be connected with the dualities discussed above.

##### Results on the process–product transition

We hypothesized that the square root tasks are of decreasing difficulty. The quantitative analysis on the raw performance scores corroborated this hypothesis. The next step in the analysis involved checking the hypothesis that the performance patterns of students are increasing, i.e., students that answered task R1 correctly also gave a correct answer to tasks R2, R3 and R4; students who gave a correct answer to task R2 also gave a correct answer to tasks R3 and R4, and so on. In the square root tasks, performance scores of 81 of 92 students



Type of answer	R1	R2	R3	R4
Correct	41	49	56	73
Incorrect, lacks connection	42	27	27	12
Incorrect, other	9	17	9	7
Total	92	92	92	92

Table 4.2: Number of students per category of answers.

showed an increasing pattern. We consider this result as a corroboration of our hypothesis. Next, a qualitative analysis was performed on the written answers to investigate whether incorrect answers are due to the inability to deal with the process–product duality. Therefore, written answers on the square root tasks were categorized into three types: correct; incorrect with a clue that the connection between the process and the product of square root is absent; incorrect, other.

A typical example of a second category answer to task R1 is “ $\sqrt{12} = 3$  so  $3 + \sqrt{3}$  does not equal  $3\sqrt{3}$  so no.” In this answer, the process of extracting the root of 12 is not performed correctly, this gives rise to the question whether the process of extracting roots is clear to this student.

Students in the third category did not give any explanation at all, or gave an answer that does not directly indicate that the connection to the underlying process is absent. For example, a student answered  $\sqrt{12} = \sqrt{4 \times 3} = 2 + \sqrt{3}$ . In this case, the problem does not seem to be the connection with the process of extracting the root, because the first step in this calculation is correct. Rather, this student mixed up multiplication and addition.

Table 4.2 shows the number of students for each category. Incorrect answers are mostly due to the lack of a connection to the underlying process. Therefore, we conclude that the transition between a square root as a process and as a product or vice versa is a factor of complexity.

The focus of the hypothesis is on students’ difficulties with the process–object duality. However, these difficulties might not be the only kind which they experience. The following dialogue between student Arvin and the interviewer illustrates that difficulties might go beyond the process–object duality. The interviewer asked Arvin to reflect on his answer to task R2 in which students were asked whether or not they agree with Sophie who claims that  $\sqrt{9 \cdot 2} + \sqrt{2} = 4\sqrt{2}$ . The answer of Arvin to this task is “no,  $\sqrt{9 \cdot 2} = \sqrt{18} = 18^{\frac{1}{2}} = 9$  so it has to be  $9\sqrt{2}$ .”

A: Hmm... So I did 9 times 2, equals 18 and 18 to the power  $\frac{1}{2}$  yields 9.

I: And how does 18 to the power  $\frac{1}{2}$  yield 9?

A: I just divided 18 by 2, equals 9.

I: OK.

A: Yes, next times plus 2.

I: Plus  $\sqrt{2}$ .

A: So that actually should have been 9 plus  $\sqrt{2}$ .

I: That should be 9 plus  $\sqrt{2}$ ?

A: Yes, yes, finished?

This student correctly makes the connection between a fraction power and a square root. However, in the evaluation of  $18^{\frac{1}{2}}$  Arvin mix up raising to the power  $\frac{1}{2}$  and multiplying by  $\frac{1}{2}$ . In addition, he confuses plus  $\sqrt{2}$  with times  $\sqrt{2}$ . This student does not know the meaning of raising to a power. Even when he elucidates his reasoning during the interview, he does not take a different view. To Arvin, the process of extracting a root is not clear. Treating  $\sqrt{5}$  as an object however in the task  $2\sqrt{5} + 4\sqrt{5}$  was no problem for Arvin. So to Arvin, manipulating with  $\sqrt{2}$  as an object is not the difficulty, rather it is the connection to the underlying process of extracting roots or raising to the power one-half. In this case,  $\sqrt{2}$  as an object is not the result of a process that is reified.

Arvin's problem with fractional powers does not directly relate to the process-object duality. This indicates that the difficulties students experience in manipulating square roots might be more than the process-object duality only.

### Results on the value-variable transition

The quantitative analysis of the raw performance scores corroborated the hypothesis that the first equation is the most difficult and the third equation is the easiest. We also hypothesized that students that are able to answer the first question correctly are also able to answer the second and third questions correctly. Analogously, we expected that students that are able to answer the second question correctly are able to answer the third question correctly. Of the 92 participating students, 88 showed an increasing pattern, which we consider as a corroboration of the hypothesis.

Next, we performed a qualitative analysis on the written answers to investigate whether the incorrect answers are due to difficulties in manipulating the expressions or due to an inability to draw the correct conclusion. Therefore, the answers are categorized into three groups: correct; correct manipulation and incorrect conclusion; incorrect manipulation. Table 4.3 shows the percentages of students in each category. The analysis shows that the majority of students are able to manipulate the equations correctly and that the difficulty lies in drawing the correct conclusion out of expressions such as  $0 = 0$  and  $0 = 2$ .

Type of answer	E1	E2	E3
Correct	10	27	71
Correct manipulation; incorrect conclusion	74	53	0
Incorrect no explanation	8	12	21
Total	92	92	92

Table 4.3: Number of students per category of answers.

The following dialogue between student Catherine and the interviewer illustrates this problem. Catherine solved Equation E1:  $2(3x + 2) = 3(2x - 1) + 7$  by expanding the brackets, which leads to the equation  $6x + 4 = 6x - 3 + 7$ . Next, she erroneously concluded that  $x = 0$ . The interviewer asked Catherine to reflect on the thoughts she had while solving this equation.

C: Well, first expand the brackets, thus you do this so to say with the half parrot's bill<sup>1</sup> as I learned that

I: Yes

C: Which yields [points to  $6x + 4 = 6x - 3 + 7$  on her work]

I: Yes

C: And  $6x$  cancels out

I: Yes

C: Next  $-3$  plus  $4$  because that becomes  $+$  so you get  $+7$

I: Yes

C: But this one is  $7$  move to the other side becomes  $-7$  so you get  $0$

I: So you get  $0$ ?

C: So  $x = 0$

I: And where does this  $x$  comes from?

C: [Silence, then with astonishment] But you actually want to know how much  $x$  is?

I: Yes

C: But I canceled that out

---

<sup>1</sup>A half parrot's bill refers to the arcs in  $2(3x + 2) = 6x + 4$ .

I: Yes

C: O yes, I conjure that up or something, actually I don't know. I should have divided! No, I don't know, I did not think about it, I just thought I have to find  $x$ , so in the end there is  $x =$  and my solution was 0. I did not think about deleting  $x$ .

To Catherine, the expectation to find an expression, in the sense of a row of symbols of the form  $x =$  number, was very strong, as is beautifully expressed by Catherine's exclamation "but you actually want to know how much  $x$  is." To Catherine, the string of symbols  $x =$  number just seemed to indicate the end of the procedure of solving the problem. At this point, she did not seem to be aware of the meaning of the row of symbols (namely the solution). The string of symbols only indicated the end of the procedure.

Sfard & Linchevski (1994b) already noticed that the inability to draw the correct conclusion out of  $0 = 0$  might be due to the tacit assumption that  $x$  equals only one value. As an alternative they suggest that manipulations might be done in a routine way without any thought on the meaning of the letters. This routine way always ends with the line  $x =$  number. Solving the equation is identified with reduction of the equation to the expression  $x =$  number.

## 4.7 Conclusion and discussion

In this Chapter, we used Sfard's theory of reification to explain the results of Chapter 2. These results fit into two transitions described by Sfard & Linchevski (1994b), in which ambiguity plays a role: the process-object duality and the duality of the algebra of a fixed number and the algebra of variables. We hypothesized that these transitions can be seen as factors of the complexity of a task. Testing this hypothesis among 92 grade 11 students corroborated this hypothesis. In the case of square roots, the majority of the incorrect answers seemed to be due to a lack of a realization of the connection between the process of extracting the square root and the root as an object. As for the linear equations, the majority of students were able to perform the secondary process of expanding the brackets correctly. Here, the difficulty is in returning to the primary process to draw the correct conclusion.

In the view of the theory of reification, the stage of reification is crucial in the development of mathematical concepts. Reification means a switch from a detailed and diffuse operational mode of thinking to a general and concise structural mode of thinking. This switch is inherently difficult and the majority of students do not reach the stage of reification during secondary education, see also Chapter 2.

Gray & Tall (1994) argued that a possible way out could be to start with processes rather than with ready-made algebraic objects. Curricula often assume

a structural approach even though the process–object duality is not grasped by the student (Sfard & Linchevski, 1994b). For example, a letter is introduced as a variable and not just as an unknown, even though from a historical perspective, the latter precedes the former, which is more advanced. In this way, the curriculum reverses the order of concept formation.

The results of this study are in line with other studies concerning the process–concept duality (Dubinsky, 1991; Gray & Tall, 1994) and the process–object duality in algebra (Sfard & Linchevski, 1994b). Central in these studies is that flexibility with the duality of mathematical entities is at the core of algebraic proficiency and that this flexibility is hard to achieve for many students. The way we brought into practice these transitions between operational and structural conceptions might be helpful for educators in identifying and classifying students' difficulties with algebra.

## Chapter 5

# Instructional efficiency as a measure for evaluating learning outcomes: some limitations

Cognitive load theory (CLT) plays an important role in the educational research literature. Meanwhile, CLT has been the subject of both conceptual and methodological criticism. This article addresses methodological limitations related to the measure “instructional efficiency” introduced by Paas and Van Merriënboer in 1993. Although originally introduced to compare different instructional designs, Van Gog and Paas (2008) suggest that this measure also might be suitable to evaluate learners’ levels of expertise. However, using instructional efficiency in a study of algebraic proficiency at secondary level revealed three methodological limitations related to the construction of instructional efficiency, the interpretation of mental effort, and the validity of mental effort.

### 5.1 Introduction

This chapter reports on the use of a measurement of the cognitive load theory (CLT) in a study of mathematics education. Cognitive load theory (Sweller, 1988; Sweller et al., 1998) is an influential theory in the field of cognition and instruction. In the past 20 years, the theory has developed substantially. A main concern of cognitive load theory (CLT) is that well designed instruction should

take into account the constraints on the learner's cognitive system. Instructional techniques should avoid overloading the learner's working memory, which is seen as the central bottleneck for learning.

The context of this study is a research project on the development of algebraic proficiency in Dutch pre-university education. In algebraic proficiency, both fluency and mastery are seen as important aspects. These aspects are reflected in CLT, where expertise is seen as a combination of performance and mental effort. Especially the latter triggered our interest, as the inclusion of mental effort seemed to offer a means of getting a handle on fluency.

CLT has been proven to provide valuable insights in mathematics education (e.g., Ayres, 1993, 2001; Cooper & Sweller, 1987; Owen & Sweller, 1989; Sweller, 1989; Sweller & Low, 1992). From a CLT perspective, it is not only important whether a solution is correct or incorrect: the invested mental effort should also be taken into account in order to obtain a more detailed view of the learner's level of expertise (e.g., Van Merriënboer & Sweller, 2005). The combination of performance and mental effort scores is expected to provide a more fine-grained view of learners' levels of expertise than performance scores alone.

Although the combination of performance and mental effort originally was used to compare different instructional designs, Van Gog & Paas (2008) suggest that this combination can also be useful to evaluate learners' relative levels of expertise. This chapter explores this suggestion, and aims at investigating the potential of the combination of performance and mental effort as a method to evaluate learners' relative levels of expertise by using data from a recent study on the development of algebraic proficiency in secondary education.

## 5.2 Main tenets of CLT

The central tenet of cognitive load theory (Sweller, 1988; Sweller et al., 1998; Van Merriënboer & Sweller, 2005) is that instruction should be designed in line with learners' cognitive architecture. Human cognitive architecture is assumed to consist of a limited working memory and a practically unlimited long-term memory. Miller (1956) revealed that working memory is only able to hold about seven items of information at a time. While processing information, working memory might even be able to deal with only four, plus or minus one (Van Gog & Paas, 2008). New information has to pass through working memory before it can be stored in long-term memory.

In this view, the limited capacity of one's working memory is a bottleneck for learning. Learning, in cognitive load theory, is seen as the construction, elaboration and automation of schemas (Sweller, 1988; Sweller et al., 1998). Schemas are constructed by combining single elements into increasingly complicated structures. These schemas can be elaborated by adding new information to already existing schemas. Once a schema has been constructed, interacting elements incorporated in the schema can be treated as single elements in working memory.

Also, such a schema can be incorporated in higher order schemas. According to CLT, learning material differs in the extent to which it imposes working memory load. The working memory load in its turn depends on the number of elements that must be processed in working memory simultaneously. This number depends on the element interactivity, where an element is “anything that has to be learned, most frequently a schema” (Sweller et al., 1998, p. 259).

Paas et al. (2003b) define cognitive load as a multidimensional construct that represents the load that performing a particular task imposes on the learner’s cognitive system. Measuring the multidimensional construct of cognitive load has proven to be difficult (Paas et al., 2003b). However, from the model of Paas & Van Merriënboer (1994), it follows that cognitive load can be assessed by measuring mental load, mental effort and performance. Because mental load is supposed to be equal for all participants, the measure of performance and mental effort suffices to measure students’ levels of expertise.

CLT distinguishes three types of cognitive load (Sweller et al., 1998). *Intrinsic cognitive load* is determined by the number of interacting elements of the learning material. High element interactivity imposes high intrinsic load because all elements have to be processed simultaneously in working memory. *Extraneous cognitive load* is the ineffective load due to the design of the instructional materials. It is unnecessary, does not contribute to learning and may even hamper learning. Intrinsic load and extraneous load are additive (Paas et al., 2003a; Sweller et al., 1998): so if the intrinsic load is low, the level of extraneous load might not exceed working memory. Lowering the extraneous load is especially important if the intrinsic load is high (Paas et al., 2003a). Finally, *germane cognitive load*, introduced by Sweller et al. (1998), refers to the additional cognitive load imposed by the design of the material that does contribute to learning.

According to CLT, appropriate instruction should avoid both overload and underload because under these conditions learning deteriorates. As a consequence, appropriate instruction should reduce the extraneous load imposed by processes that hamper learning and increase the germane load that fosters learning (Sweller et al., 1998; Van Gog & Paas, 2008). CLT has distinguished several effects that reduce extraneous load (Van Merriënboer & Sweller, 2005). For example, the goal-free effect is imposed by the replacement of conventional problems with goal-free problems, i.e., problems with a non-specific goal. This effect reduces the extraneous load because the learner does not have to relate the current problem state with the goal state while working on the task (Ayres, 1993). Other effects that reduce the extraneous load are the worked example effect, the completion problem effect, the split attention effect, the modality effect, and the redundancy effect (Van Merriënboer & Sweller, 2005).

In the past decades, CLT has developed into an influential theory with many valuable contributions to the field of cognition and instruction. Recently, CLT has been the subject of criticism, both conceptual (Schnotz & Kürschner, 2007) and methodological (Hoffman & Schraw, 2010; De Jong, 2010; Gerjets et al.,



2009). The conceptual criticism relates to the distinction between different kinds of cognitive load. The methodological criticism refers to its measurement. The focus of this chapter is to investigate the idea of using instructional efficiency as a measure for evaluating the learner's level of expertise.

### 5.3 Instructional efficiency

In 1993, Paas & Van Merriënboer (1993) introduce a measure for efficiency as a combination of performance and mental effort. This measure is introduced as a means to compare the effects of different instructional approaches and has become known as “instructional efficiency.” It is based on performance scores and mental effort scores. In Paas & Van Merriënboer (1993), performance scores are determined by the percentage of correct answers. Mental effort scores are the measures for the perceived amount of mental effort, reported by the students on a nine-point symmetrical category scale. The values and categories of this scale range from (1) very, very low mental effort to (9) very, very high mental effort.

The performance and mental effort scores are standardized. These standardized performance  $z$ -scores and mental effort  $z$ -scores are presented in a coordinate plane—performance vertically, mental effort horizontally—in which each unit of (standardized) invested mental effort equals one unit of performance. For a particular instructional approach with certain standardized scores, the efficiency is defined as the distance of this point to a baseline condition  $E = 0$ , combined with a sign indicating at which side of the baseline the condition is located. Figure 5.1 shows that efficiency increases as performance increases and mental effort decreases. From the standardization, it follows that efficiency is a relative measure. The performance and mental effort scores of a subgroup of persons and/or items are related to the performance and mental effort of the whole group of persons and/or items. Subgroups are for example persons with particular instructional design or students of a particular grade. In the case of using efficiency to evaluate learners' levels of expertise, a subgroup consists of one particular student.

Since the introduction of this measurement, many researchers have used instructional efficiency either in its original form or in an adapted form (Kester et al., 2006; Paas et al., 2007; Paas & Van Merriënboer, 1994; Tuovinen & Sweller, 1999; Van Gog et al., 2008). Although originally developed to compare different instructional designs, Van Gog & Paas (2008) emphasize that instructional efficiency also can be viewed as an expression of a learner's level of expertise. In their view, instructional efficiency can be used to evaluate a learner's level of expertise. This chapter explores this suggestion. Therefore, we now go more deeply into instructional efficiency. We use the notations and abbreviations as introduced by Paas & Van Merriënboer (1993):

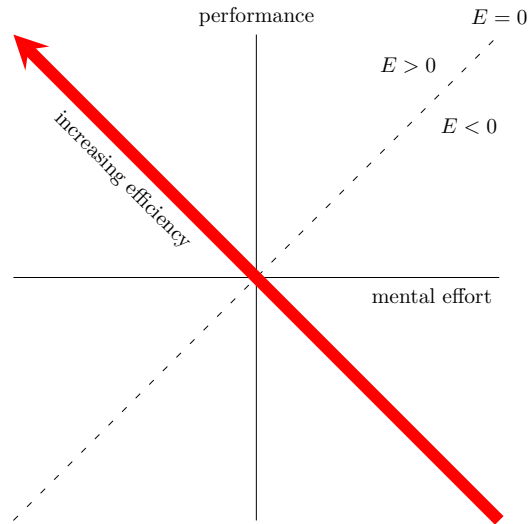


Figure 5.1: Standardized performance and mental effort represented graphically.

$P$ : average performance score of a particular subgroup of one or more persons and/or items,

$R$ : average mental effort score of a particular subgroup of one or more persons and/or items,

$P_M$ : the grand mean of performance scores of the whole group of persons and/or items,

$P_S$ : the standard deviation of performance scores of the whole group of persons and/or items,

$R_M$ : the grand mean of mental effort scores of the whole group of persons and/or items,

$R_S$ : the standard deviation of mental effort scores of the whole group of persons and/or items,

$P_z$ : the standardized performance score,  $P_z = \frac{P - P_M}{P_S}$ ,

$R_z$ : the standardized mental effort score,  $R_z = \frac{R - R_M}{R_S}$ .

Here we note that a subgroup of persons can refer to just one student as well as to all students of a class or school. Analogously, a subgroup of items can refer to one single item as well as to a cluster of items. Let us further analyze the relationship between those variables in order to obtain a deeper understanding of the behavior of the measure instructional efficiency. By definition  $P_z = \frac{P - P_M}{P_S}$

and  $R_z = \frac{R - R_M}{R_S}$ . With these two variables, instructional efficiency of  $P$  and  $R$  scores of a particular subgroup of persons and/or items can be calculated.

$$\begin{aligned}
 E(P, R) &= \frac{P_z - R_z}{\sqrt{2}} \\
 &= \frac{1}{\sqrt{2}} \left( \frac{P - P_M}{P_S} - \frac{R - R_M}{R_S} \right) \\
 &= \frac{1}{\sqrt{2}} \left( \frac{1}{P_S} \cdot P - \frac{1}{R_S} \cdot R + \frac{R_M}{R_S} - \frac{P_M}{P_S} \right) \\
 &= \frac{1}{\sqrt{2}P_S} \cdot P + \frac{-1}{\sqrt{2}R_S} \cdot R + \frac{1}{\sqrt{2}} \left( \frac{R_M}{R_S} - \frac{P_M}{P_S} \right).
 \end{aligned}$$

Note that  $P_M$ ,  $P_S$ ,  $R_M$  and  $R_S$  are constants, independent of which particular subgroup of persons and/or items is being considered. Therefore, the measure instructional efficiency is a linear transformation of the variables  $P$  and  $R$ .

Criticism of the construction of instructional efficiency addresses its relativity (De Jong, 2010; Hoffman & Schraw, 2010) and reliability (Hoffman & Schraw, 2010). Apart from this criticism, with which we agree, we experienced three specific limitations within the context of the use of instructional efficiency in research on the development of algebraic proficiency. In the following sections, we elaborate on these limitations.

## 5.4 Research context

As we have mentioned before, this study is part of a research project that aims at investigating the development of algebraic proficiency in pre-university students, see Chapter 2. In order to monitor this development, we assessed students of grades 8–11 at four different moments in time, spread out more or less evenly over the course of a calendar year. Each assessment consisted of approximately 12 open items. After each item students were asked to report their invested mental effort on the nine-point symmetrical category scale, corresponding literally to Paas (2007). Each item was allocated a performance score (0 for incorrect; 1 for correct) and a mental effort score (1 for very, very low mental effort, 2 for very low mental effort, and so on until 9 for very, very high mental effort).

For the data analysis, we planned to combine the performance and mental effort scores in the measure, instructional efficiency. We expected this combination to provide a more fine-grained analysis than performance scores alone. However, during the analysis, three methodological limitations emerged related to instructional efficiency, namely the interchangeability of performance and mental effort, the undetermined meaning of mental effort in case of an incorrect answer, and the validity of the mental effort scores.

## 5.5 First limitation: Interchangeability of performance and mental effort

Central to the construction of instructional efficiency is the assumption that points on lines parallel to the line  $E = 0$  represent the same efficiency (see Figure 5.1). De Jong (2010) already mentioned that this assumption can lead to unclear experimental situations. When we try to apply instructional efficiency as a measure to evaluate a learner's level of expertise, this assumption becomes even more problematic. We elaborate this limitation with an example.

Consider two students, student 1 and student 2, with average performance and mental effort scores as illustrated in Figure 5.2. That is, suppose that student 1 has an average performance score less than the average performance score of the whole group, and an average mental effort score less than the average mental effort score of the whole group. Further, suppose that student 2 has an average performance score higher than the average performance score of the whole group and an average mental effort score higher than the average score of the whole group. The distances to the line  $E = 0$  are equal for both student 1 and student 2. So by definition, these two students have the same instructional efficiency. Following the reasoning of Van Gog & Paas (2008), the conclusion should be that student 1 and student 2 have the same level of expertise.

In our view, this phenomenon is an expression of a fundamental problem of the measure, instructional efficiency. In the construction of this measure, mental effort counterbalances performance. This means that a higher performance is annihilated by a lower mental effort, and the other way around, due to the assumption that assigns the same instructional efficiency to all lines parallel to the line  $E = 0$ . This assumption is based on the idea that performance and mental effort are interchangeable. However, from an educational point of view, this is questionable, because higher performance is much more valuable than lower invested mental effort. In our view, performance and mental effort do not counterbalance. As a consequence, the assumption of linearity expressed by Paas & Van Merriënboer (1993), that assumes that lines that are parallel to the line  $E = 0$  in Figure 5.1 represent the same instructional efficiency, does not hold.

## 5.6 Second limitation: Undetermined meaning of mental effort in the case of an incorrect answer

The second limitation of instructional efficiency as a measure to evaluate learners' levels of expertise is the meaning and hierarchy of mental effort scores. In Section 5.3 we argued that instructional efficiency is a linear combination of performance and mental effort.

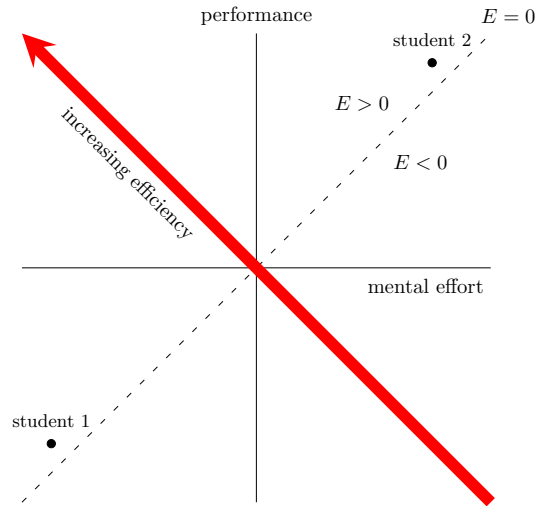


Figure 5.2: Two students with the same efficiency.

In the data from all students and all items in our project, the grand mean of performance  $P_M$  is 0.51 and the standard deviation of the performance  $P_S$  is 0.50. Note that these data are dichotomous and thus not normally distributed. The grand mean of mental effort  $R_M$  is 3.76, and the standard deviation of the mental effort  $R_S$  is 2.47. With these four variables, it is easy to calculate instructional efficiency according to the method of Paas & Van Merriënboer (1993). Based on these values, Figure 5.3 shows the instructional efficiency of different combinations of average performance and average mental effort scores of subgroups of persons and/or items as linear scales. The averages of subgroups can take many different values in between 0 and 1. In order to illustrate the shifts of the linear scales, we chose to depict, besides 0.00 and 1.00 (the extreme values for averages of subgroups), also performance scores of 0.25, 0.50 and 0.75. Analogously, for the mental effort scale we chose the values 1, 2, ..., 9 as possible averages of subgroups.

The combination of an average performance score of  $P = 0$  and an average mental effort score of  $R = 9$  yields the lowest efficiency ( $E = -2.22$ ); the combination of an average performance score of  $P = 1$  and an average mental effort score of  $R = 1$  yields the highest efficiency ( $E = 1.48$ ). From the construction of instructional efficiency, it follows that for a certain average performance score, there is a strict hierarchy in possible average mental effort scores. In the case of a correct answer, the meaning and hierarchy of the mental effort scores are clear: a correct answer combined with low invested mental effort is of higher instructional efficiency and thus represents a higher level of expertise than a cor-

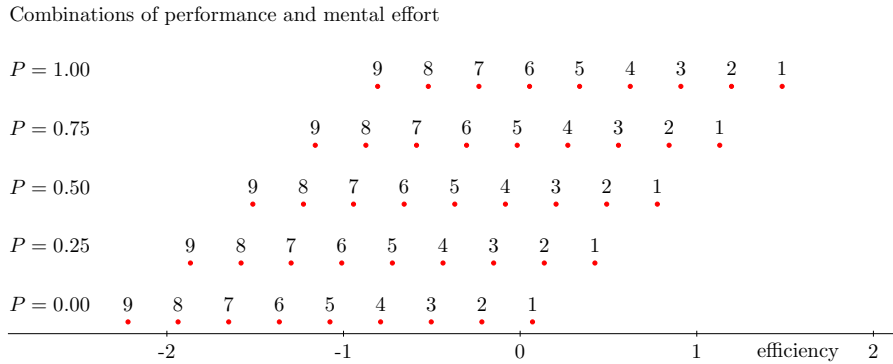


Figure 5.3: Instructional efficiency as linear scales for different values of  $P$ .

rect answer combined with high invested mental effort. However, in case of an incorrect answer, the meaning and hierarchy of the mental effort scores are not clear. We elaborate on this in the remainder of this section.

Consider the situation of low performance in combination with low mental effort. This situation is problematic for two reasons.

First, the meaning of this combination is not clear. Very low invested mental effort combined with a low performance score might on the one hand express the quick decision to not even try to solve the problem at all, because the student recognizes the problem as being beyond his or her ability. On the other hand, this combination might express only tiny errors in almost perfect computations. Both situations have the same instructional efficiency, so they should both express the same level of expertise. However, in our view, these two situations are quite different. In the first situation the student is just at the beginning of the road to mastering the problem, whereas in the second situation only a small learning effort is needed for the student to master the problem. Thus, the meaning of the mental effort scores in case of low performance is ambiguous.

Secondly, compare this situation of low performance and low mental effort to the combination of low performance and very high mental effort. This second situation can be an expression of a tiny error in an almost perfect computation that is only just within the student's capability. Compared to the situation of low mental effort and low performance, this is clearly a situation of a much higher level of expertise. Now from the linear construction of instructional efficiency, it follows that higher mental effort means lower instructional efficiency (see Figure 5.3). As a consequence, the instructional efficiency of the first situation is higher than the instructional efficiency of the second situation, exactly the opposite to what it should be.

Thus, the same combination of performance and mental effort can have multi-

ple interpretations, but lead to the same instructional efficiency. These multiple interpretations pose a problem in the transition from instructional efficiency to level of expertise. Furthermore, the inherent hierarchy of instructional efficiency is shown to be the wrong way around in certain situations. We conclude that in the case of our study the meaning and chosen fixed hierarchy of mental effort scores are not clear.

## 5.7 Third limitation: Validity of mental effort scores

The undetermined meaning of mental effort score as discussed in the previous section produces doubts as to the validity of mental effort as a measure to measure students' levels of expertise. In this section, we discuss three more reasons to question the validity of mental effort.

The first reason to doubt the validity of mental effort is the interpretation of mental effort by the student. In our study, students were asked to report on their invested mental effort after each item. Figure 5.4 shows the work of a grade 8 student. Initially, this student started to solve the equation

$$(x - 5)(x + 2)(x - 3) = 0$$

by expanding the brackets and reported very high mental effort. Then the student discovered the correct procedure, gave the (correct) solutions  $x = 5$ ,  $x = -2$  and  $x = 3$  and reported low mental effort. In our view, the reported switch in the perceived mental effort does not reflect the real invested mental effort. The working out in this example clearly illustrates that the student has invested high mental effort in order to arrive at the correct answer. However, the switch in reported mental effort on second thought, inspired by a switch of perceived difficulty, points to a validity problem. The example in Figure 5.4 suggests that for students it is hard to distinguish between perceived mental effort and perceived difficulty.

The second reason to doubt the validity of mental effort is that the question how much effort the student needed to invest to solve a problem allows for multiple interpretations. One possible interpretation of this question is: how much effort did solving this problem cost *me* in relation to other students in my class. In this case, while reporting mental effort, the student uses the class as frame of reference. Another possible interpretation is: how much effort did solving *this* problem cost in relation to other problems in the present test. In this case, the student uses the other items in the test as frame of reference. These different interpretations of how to report mental effort may influence the validity of the measurement of mental effort in our test.

Finally, in search of an alternative use for the mental effort scores, we related the mental effort scores with the ability of students that we found in the Rasch

10. Los op:

$$(x-5)(x+2)(x-3) = 0.$$

$x-5=0 \vee x+2=0 \vee x-3=0$   
 $x=5 \vee x=-2 \vee x=3$

~~$(x^3+2x-5x-10) \cdot (x-3) = 0$~~   
 ~~$x^3-3x^2+2x^2-6x-5x^2+15x-10x+30 = 0$~~   
 ~~$x^3-2x^2-3x+30 = 0$~~   
 ~~$x^3-6x^2-x+30 = 0$~~   
 ~~$x^3-6x^2-x = -30$~~

Het oplossen van het probleem hierboven kostte mij:

zeer zeer weinig moeite            zeer zeer veel moeite

↑ - toen ik = 0 zag

~~$x^3+x^2-x=5$~~   
 ~~$x^3+x^2-x-5=0$~~   
 ~~$x^3+x^2-x-5=0$~~   
 ~~$x^3+x^2-x-5=0$~~

Translation:

los op: solve

Het oplossen van het probleem hierboven kostte mij:

solving the problem above required

zeer, zeer weinig moeite: very, very low mental effort

zeer, zeer veel moeite: very, very high mental effort

toen ik = 0 zag: when I saw = 0.

Figure 5.4: Different interpretations of mental effort in the work of a student of grade 8.



analysis of the performance scores. The Rasch model is a one parameter item response model (Bond & Fox, 2007; Rasch, 1980). With this model, a linear scale has been created on which both students and items are arranged, the students in order of their ability, and the items in order of their difficulty. A key characteristic of the Rasch model is that in the computation of the ability of the students, the difficulty of the items is taken into account. Based on the Rasch measurements, we calculated for each student and each item the probability of that student's answering that item correctly. We expected students who possessed a probability of almost 1 to report very, very low mental effort, and students who possessed a probability of almost 0 to report very, very high mental effort. So, we expected a high positive correlation between the probabilities and the reported mental effort. However, the Pearson correlation coefficients between the probabilities and the reported mental effort had an average of  $-0.29$ . This clearly indicates that the relation between students' abilities and mental effort is unclear. Because of the troublesome relation between students' ability and mental effort, we argue that the use of mental effort in addition to performance to provide a more fine-grained analysis is not feasible.

The situations we described above all relate to the validity of mental effort. Literature about the validity of mental effort (Paas, 1992) refers to studies in which rating scales are used for measuring task difficulty (e.g., Bratfish et al., 1972). In our view, the relation between the validity of rating scales in measuring task difficulty and rating scales for invested mental effort is not clear. A more detailed validity study of mental effort rating scales is needed to obtain a good measurement of the invested mental effort.

## 5.8 Conclusion

In addition to problems with instructional efficiency, such as problems with reliability and group reference already identified by others (e.g. De Jong, 2010; Hoffman & Schraw, 2010), this chapter identified three methodological limitations that emerged when using instructional efficiency as a measure to evaluate students' levels of expertise in the context of a study on secondary level algebraic proficiency. The first limitation is related to the assumption that higher mental effort can be counterbalanced by higher performance and the other way around. From an educational point of view, this assumption is problematic because higher performance is considered more valuable than lower mental effort. However, this assumption of the interchangeability of performance and mental effort is crucial in the construction of instructional efficiency.

The second limitation lies in the multiple interpretations of mental effort in situations of low performance and low mental effort. In this situation, on the one hand, the combination might express a high level of expertise with some tiny mistakes. On the other hand, this combination might express a low level of expertise where the student did not even try to solve the problem.

Finally, the third limitation addresses the validity of mental effort as such. We argue that there is reason to question the validity of mental effort measured on the nine-point symmetrical category scale due to the mix of mental effort and perceived difficulty.

Based on these observed limitations, we conclude that, in the context of our study on the development of students' algebraic proficiency, instructional efficiency is not an appropriate measure for evaluating learners' levels of expertise. The suggestion by Van Gog & Paas (2008) to use instructional efficiency in this way was prompted by the fact that percentages of correct answers do not always reflect the level of learners' expertise in an appropriate manner. The limitations we experienced with the measures instructional efficiency and mental effort however are too fundamental to use these scores for a more fine-grained analysis, replacing the percentage of correct answers.

As an alternative, Hoffman & Schraw (2010) suggest using the ratio of performance and mental effort as a measure for efficiency. Because of the doubts on the validity of the mental effort scores, we decided not to use this alternative. In fact, we decided to keep the mental effort scores out of the analysis on students' algebraic proficiency.



## Chapter 6

# Contexts and models in Dutch textbook series: the case of linear relations and linear equations

In this chapter we explore what kind of support Dutch textbooks offer to help students develop conceptual algebraic proficiency, using emergent modeling (Gravemeijer, 1999) as a theoretical lens. The analysis shows that Dutch textbook series pay much attention to activities concerning exploring contexts and little attention to activities such as structuring, abstracting, formalizing, generalizing. Moreover, we find two disjunct didactical tracks: a reform mathematics track and a track of ready-made mathematics.

### 6.1 Introduction

In the previous chapters, we found that Dutch pre-university students did not develop proficiency in conceptual aspects (Chapter 2), that students had difficulties with structure sense (Chapter 3), and that operational–structural transitions could be seen as indicators of complexity (Chapter 4). Together, these chapters suggest that the majority of Dutch students do not reach a level of higher-order skills. In this chapter we focus on the question of to what extent Dutch textbooks support reaching this higher level. But first, we need to describe more precisely what we mean by this higher level.

In this chapter, we use the term *conceptual proficiency* to refer to this higher

level of mastering algebra. Conceptual proficiency is an umbrella term for skills that transcend superficial skills. The nature of these higher-order skills has been described in many different ways by many different researchers. For example, relational understanding (Skemp, 1976), conceptual understanding (Kilpatrick et al., 2001), flexible manipulation skills and the ability to read through symbols as elements of symbol sense (Arcavi, 1994, 2005), proceptual view (Tall & Thomas, 1991), reification (Sfard, 1991), and relations between relations (Van Hiele, 1986). We use the term conceptual proficiency to denote a mixture of these higher-order skills in which three aspects are essential: the ability to recognize and make flexible use of the algebraic structure, to deal with the ambiguous nature of mathematical concepts, and see the coherence between mathematical concepts.

In the previous chapters we argued that the majority of students do not show conceptual proficiency. Since textbooks play an important role in Dutch mathematics education (Hiebert et al., 2003), we decided to perform an analysis of some textbooks. For pragmatic reasons, we decided not to take into account factors such as school, class, or teacher.

In this analysis of textbooks, we investigated what kind of support Dutch textbook series offer students to help develop conceptual proficiency. We choose to focus this analysis on the topics linear relationships and linear equations. A linear relationship between two variables is a relation in which any given change in one variable will always produce a corresponding change in the other variable. The graph of a linear relationship is a straight line. By linear equation, we refer to a polynomial equation in one variable in the first degree, for example  $3x - 7 = 11$  or  $5x + 2 = -2x + 7$ .

The first reason to choose these topics is that the focus of this study is on the development of algebraic skills taught in lower secondary education. Central in the curriculum of lower secondary education is the introduction of linear relations and linear equations. The second reason is that from the results discussed in Chapter 2, it follows that linear equations are among the most difficult tasks. We split the analysis into linear relations and linear equations.

Dutch textbook series are strongly influenced by the theory of Realistic Mathematics Education. As a consequence, linear relations and linear equations are introduced by means of contexts and models. This chapter aims at investigating the role contexts and models play in concept formation. Moreover, we aim at investigating what kind of support these contexts and models offer to develop conceptual proficiency. The role of models within RME is described in the theory of emergent modeling (Gravemeijer, 1999). Below we elaborate on this theory.

## 6.2 Frame of reference

### 6.2.1 Emergent modeling

Emergent modeling (Gravemeijer, 1999) is rooted in the description of instructional design heuristics within the theory of Realistic Mathematics Education (e.g., Freudenthal, 1983). Realistic Mathematics Education (RME) is a domain-specific instruction theory that has its origins in the early 1970s, and is based on Freudenthal's view on mathematics (e.g., Freudenthal, 1968; Freudenthal, 1973). He argues that

“what humans have to learn is not mathematics as a closed system, but rather as an activity, the process of mathematizing reality and if possible even that of mathematizing mathematics” (Freudenthal, 1968, p.7).

To Freudenthal, students should not be confronted with ready-made mathematics. Rather, they should be enabled to learn mathematics by mathematizing both reality and their own mathematical activities. In RME, the focus is on teaching the activity of mathematizing instead of teaching the results of the mathematizing activities of others. The latter would lead to an “anti-didactic inversion” of the way the mathematics was invented. In the RME view, students should be enabled to construct mathematics from their own experiences and reality. In a process of progressive mathematizing, students reinvent mathematics, guided by teachers and the instructional design. In this way, students are enabled to experience mathematics as a process similar to the way mathematicians invent mathematics. It is the learning process that is seen as essential, rather than the products of the learning process (Freudenthal, 1991).

Modeling activities play an important role in mathematics. By translating a contextual problem into a mathematical problem, the problem solver makes the problem amenable to mathematical procedures. From an RME perspective, students should not be confronted with ready-made models. Rather, these models should emerge from their own mathematizing activities (Gravemeijer, 1994).

The idea of emergent modeling is that students start with modeling experientially real problem situations. Then, in the following learning process, the model gradually develops from a *model of* their own mathematical activity to a *model for* formal mathematical reasoning. The use of models in the latter form lies close to the intended use of didactical models. These models are meant to make formal mathematics more accessible to students. From a constructivist perspective, the problem with these kind of didactic models is that in order to interpret the model correctly, students should already have the mathematical knowledge at their disposal that is intended by the model (Cobb et al., 1992). As an alternative, emergent modeling concentrates on rooting the development of the model in the experiential reality of the students.

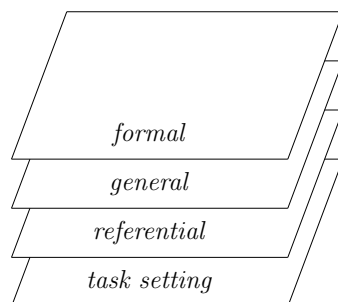


Figure 6.1: Levels of activity (Gravemeijer, 1999).

In this way, models can support mathematical growth. While working with the models, students start to come to grips with the mathematical relations involved. In this way, the model starts to function as a model for mathematical reasoning. The transition from *model of* to *model for* concerns a shift in students' thinking. This shift concerns thinking about the modeled contextual situation to thinking about mathematical relations. Gravemeijer (1999) discerns four levels of different activities: task setting, referential, general, and formal. Task setting activities concern reasoning about interpretations and solutions in a specific setting. At the referential level, the phrase “models of” refers to activities described in the task setting. General activities emerge as the students start to reason about the mathematical relations that are involved, thus constructing new knowledge. Activities at the level of formal mathematical reasoning are no longer dependent on models. Although there is a hierarchy in these levels, student activities do not take place in a strict order. Discussions about activities at a general level for instance may fold back to the referential or task setting level. In the transition from referential to general activities, models are reified, thus becoming entities on their own. At this point, Gravemeijer (1999) makes a connection to the reification of processes (Sfard, 1991), see also Chapter 4. Following this line of thought, Gravemeijer (1999) argues that it is not the model but the process of acting and reasoning with the model that is being reified.

Since we are investigating what kind of support the textbooks offer to develop conceptual proficiency, we are especially interested in activities that support the transition from the referential level to the general level.

### 6.2.2 Linear relations and linear equations from an emergent modeling perspective

From an RME perspective, and in particular from an emergent modeling perspective, students eventually construct formal mathematical knowledge in a process

of progressive mathematization of both reality and the students' mathematical activities. Following the line of thought of RME and emergent modeling, we sketch the ideal instructional sequence for linear relations and linear equations.

In our view, from an RME perspective, the introduction of linear relations should start with contextual situations of linear relations. Activities at this task setting level should invite students to reason and calculate within these contexts. Examples of such activities are discovering patterns in contextual situations and informal reasoning and calculating. In this way, the contexts can serve as models of linear relations. Next, the activities should shift to their mathematical characteristics and the mathematical relations involved. Gradually, the activities become detached from the contextual situations and the students' reasoning loses its dependency on situation-specific features. In this manner, a linear relation becomes an object. This object does not necessarily need its context, but instead has meaning in and of itself. This object is incorporated in a network of relations. At this general level, linear (word) formulas start to function as models for linear relations.

From the same perspective, in our view, the ideal introduction of linear equations starts with informal strategies at the task setting level, such as the cover-up method and the balance method, that emerge from contextual situations. Next, at the referential level, these informal strategies serve as a background for solving equations and in interpreting and giving meaning to the solution. Gradually, the balance model becomes less detached from the specific situation and emerges as model for reasoning about linear equations. At the general level, equations have become reified into objects and are part of a network of mathematical relations that justify how to operate on these objects. For example, adding and subtracting numbers and variables to both sides of the equation and multiplying both sides of the equation by the same number.

To get a handle on the way the textbooks support the students in developing conceptual proficiency in the domain of linear relations and linear equations we carried out an exploratory analysis. As a framework of reference we used the idea of emergent modeling, which offered a basis for identifying categories of instructional activities that would support the process of vertical mathematization that is aimed for in RME theory. By quantifying the number of activities in the respective categories, we would be able to create an image of in how far the textbook approach corresponds with a more or less ideal realistic instructional approach.

As we mentioned before, we focus on the topics linear relations (Section 6.3.1) and linear equations (Section 6.4).



Grade	Textbook	Chapter/Section	Edition
7	GR	4.4, 4.5, 4.6	Reichard et al. (2006)
7	MW	7, 8	De Bruijn et al. (2007)
8	GR	5.1, 5.2 5.3	Reichard et al. (2005)
8	MW	1	De Bruijn et al. (2008a)
9	GR	1.1	Reichard et al. (2010)
9	MW	1.1, 9.1	De Bruijn et al. (2009a,b)
10	GR	2	Reichard et al. (2007)
10	MW	Blok 2	Bos et al. (2007)

Table 6.1: Textbooks and chapters concerning linear relations of *Getal & Ruimte* (GR) and *Moderne Wiskunde* (MW).

## 6.3 Linear relations

### 6.3.1 Method

The data set is constructed from two textbook series, *Moderne Wiskunde* (MW) and *Getal & Ruimte* (GR), which together have an estimated market share of over 90% (cTWO, 2009). From these textbook series, chapters, and within chapters, sections, have been selected that concern linear relations. Theory on these topics is presented in the textbooks for grades 7, 8, 9, and 10. In grades 10, 11, and 12, the student population is split into a social stream and a science stream. In this chapter we focus on students in the science stream because especially for these students it is important to develop conceptual proficiency. Table 6.1 provides an overview for both textbook series of the chapters and sections for each grade in which the subject of linear relations is treated.

As units of analysis, we chose tasks that are denoted with a separate number and fragments of texts on theory that belong to the main sections of the chapters. Besides these main sections, chapters of both textbook series also contain sections concerning pre-requisite knowledge, information technology, diagnostic tests, summaries, etc. We decided not to take these extra sections into account, because they do not directly contribute to the introduction of new concepts. For both textbook series, grade 10 is the last grade in which the topic of linear relations is explicitly paid attention to.

Drawing on the theory of emergent modeling, and on the description of the “ideal curriculum” as we view it, we categorized tasks and fragments of theory in the framework of emergent modeling. We discuss this analysis of linear relations and linear equations in subsections 6.3.1 and 6.4 respectively.

In subsection 6.2.2, we briefly described what we see as the ideal instructional sequence of linear relations from an emergent modeling perspective. In our view,

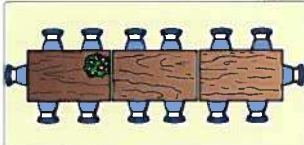
the investigation of contextual situations can be seen as an activity at the task setting level. The shift to the referential level occurs when (word) formulas are introduced and the focus of the activities shifts to the meaning of the formulas, while maintaining a strong relation to the context. Ideally, the instructional sequence then gradually shifts to the general level when the students have developed a sense of the properties of linear relations and the ways of operating with them. Activities that contribute to this shift should focus on these properties and ways of operating and less on what the models signify.

In this analysis of these textbooks, we chose not to make a strict connection between the four levels of activity described in the theory of emergent modeling and activities in the textbooks for two reasons. The first reason is that the levels describe the students' mental activity, not the tasks or the models per se. The second reason is that in the theory of emergent modeling, the shifts between levels is seen as important. This aspect would not be done justice to if we had classified activities at the four levels. Rather, the way this theory described the construction of formal mathematical knowledge gave rise to distinguishing three categories of activities for linear relations as well as linear equations.

The first category consists of activities that ask students to reason and calculate within contexts. Activities at this level concern for instance: searching for patterns in given contextual situations, constructing word formulas, informal reasoning about the context, drawing a graph from a context, and drawing conclusions about the context from a graph. In activities at this category, the constant and proportional factor in the word formula derive their meaning directly from the context.

The second category consists of activities in which the focus shifts from the context itself to the mathematical properties and characteristics involved. Examples of such activities are: searching for patterns in the relation, reasoning and calculating with formulas, comparing properties of relations, e.g., which one is steeper, and investigating the rise. In activities at this level, it becomes clear that the same rule holds for different inputs and that this rule involves something constant plus something that is proportional to the input.

The third category consists of activities that help students to develop conceptual proficiency. These activities offer support to seeing and treating the linear relationship as an object that does not necessarily derive its meaning from the context, but has a meaning by itself and by its relation to other mathematical objects. Seeing linear relations this way requires formalization and generalization to algebraic contexts. For example, a shift from the discrete to the continuous, a shift from only the first quadrant to all four quadrants, and a clear meaning for  $x = 0$ . The algebraic context concerns the connection between the formula and the intercept and the slope of linear formulas.

<p><b>4</b> Een caféhouder heeft rechthoekige tafels. Hiernaast zie je hoe hij ze neerzet.</p> <p><b>a</b> Hoeveel mensen kunnen er zo bij vier tafels zitten? En hoeveel bij vijf tafels?</p> <p><b>b</b> Hoeveel stoelen staan er steeds per tafel, de uiteinden niet meegerekend?</p> <p><b>c</b> Hoeveel stoelen komen er nog bij aan de uiteinden van de rij?</p> <p><b>d</b> Schrijf de volgende regel voor het berekenen van het aantal stoelen op: 'Het aantal tafels keer ... plus ... is gelijk aan het aantal stoelen.'</p> <p><b>e</b> Bereken hoeveel stoelen er in deze opstelling moeten staan bij zeven tafels.</p>	
---	--

*Translation:*

*4 A café owner has rectangular tables. On the side, you can see how he places them.*

*a How many people can be seated at four tables? And how many at five?*

*b How many chairs are there per table, the ends excluded?*

*c How many chairs do the ends of the row have?*

*d Write the following line down for the calculation of the number of chairs: "The number of tables times ... plus ... equals the number of chairs."*

*e Calculate how many chairs are needed in this setting with seven tables.*

Figure 6.2: Example of a task setting activity in the first category of a grade 7 textbook. Reprinted from De Bruijn et al. (2007, p. 11) with permission of the publisher.

### 6.3.2 Analysis

In the previous subsection, we described three categories of activities that fit into the ideal curriculum sketched in subsection 6.2.2. In this subsection, we describe these categories in greater detail by providing examples of activities from both textbook series.

#### First category

Activities in the first category concern the investigation of contextual situations. Figure 6.2 represents an example of such an activity of a grade 7 textbook (De Bruijn et al., 2007). In this task, students have to reason about the number of chairs depending on the number of tables. Each table provides four chairs, excluding the ends. The relationship between the number of tables and the number of chairs is linear, but the term 'linear' is not introduced in this grade 7 textbook. In this stage, the focus is on capturing the relation in a word formula and using this word formula to calculate the number of chairs when the number of tables is given, see part **d** and **e** of the task in Figure 6.2. In this example, students do not have to search for patterns or characteristics of the context. All calculations are straightforward and stay within the context.

**A 90** Twee kaarsen worden tegelijk aangestoken. De lengte van kaars I in cm is te berekenen met de formule  $\text{lengte} = 15 - 3u$ . Hierin is  $u$  het aantal branduren.

**a** Teken de grafiek van de lengte van kaars I.

Kaars II is 12 cm lang. Als deze kaars brandt, wordt hij per uur 2 cm korter.

**b** Vul in. Bij kaars II hoort de formule  $\text{lengte} = \dots$

**c** Teken in de figuur van vraag a de grafiek van de lengte van kaars II.

**d** Na hoeveel uur zijn de kaarsen even lang?

**e** Welke kaars brandt het langst?

*Translation:*

*Two candles are lit at the same time. The length of candle I in cm can be calculated with the formula  $\text{length} = 15 - 3u$ . The  $u$  represents the number of hours it has been burning.*

*a Draw the graph of the length of candle I.*

*The length of candle II is 12 cm. If this candle is burning, it shortens by 2 cm an hour.*

*b Fill in. The formula  $\text{length} = \dots$  belongs to candle II.*

*c Draw the graph of the length of candle II in the figure of subquestion a.*

*d After how many hours are the lengths of the candles equal?*

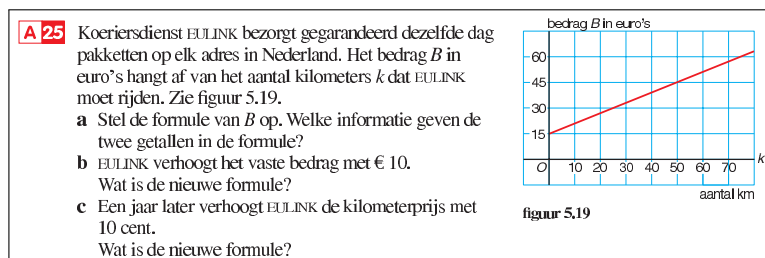
*e Which candle burns longer?*

Figure 6.3: Example of a first category activity of a grade 7 textbook. Reprinted from Reichard et al. (2006, p. 156) with permission of the publisher.

Another example of activities in the first category are tasks in which students have to reason and calculate with a given formula, see the example in Figure 6.3. In this example, the length of a candle (length) depends on the number of hours ( $u$ ) the candle has been burning via the formula  $\text{length} = 15 - 3u$ . Students have to graph this relation and compare the length with another candle. We classified this task as being of the first category because the focus of this task is within the context, and not on the linear structure of the relation or on properties of the linear relation.

### Second category

The second category consists of activities in which the focus shifts from the context itself, to properties and characteristics of the context. A representative example of an activity at this level is presented in Figure 6.4. In this example, the graph is given of the linear relation between the costs of a courier service and the number of kilometers traveled. Based on this graph, students are asked to construct the corresponding formula. Next, students have to explain the meaning of the numbers in the formula. At this point, students have to make the connection between characteristics of the context (the fixed costs and the variable costs) and the numbers in the formula. In this way, the focus shifts



*Translation:*

*Courier service EULINK guarantees to deliver packages in one day to any address in the Netherlands. The cost  $B$  in euros depends on the number of kilometers  $k$  that sc eulink has to drive. See Figure 5.19.*

*a Construct the formula for  $B$ . What information do the two numbers in the formula provide?*

*b EULINK raises the fixed costs by €10. What is the new formula?*

*c After a year, EULINK raises the costs per kilometer by 10 cents. What is the new formula?*

Figure 6.4: Representative example of a second category activity of a grade 8 textbook. Reprinted from Reichard et al. (2005, p. 179) with permission of the publisher.

from the context itself to its properties.

### Third category

In the third category, we included activities that help students to develop higher-order thinking, which we defined in Section 6.1 as conceptual proficiency. Activities at this level require students to see and treat the linear relationship as an object that does not necessarily derive its meaning from the context. In activities in this category, the linear relation has become an independent entity, an object, that has meaning on its own. Figure 6.5 provides a representative example of such an activity, taken from a grade 8 textbook (Reichard et al., 2005). In this example, three formulas of linear relations are given,  $y = 0.5x - 4$ ,  $y = 2x + 2$ , and  $y = 0.5x + 2$ . The question is which of these formulas intersect the  $y$ -axis in the same point, and how you can see this from the formulas. In this way, the linear relations are seen as objects with certain properties. These properties, such as the intersection with the  $y$ -axis, are the subject of investigation.

### Additional fourth category

During the analysis we found that the activities in Dutch textbook series do not fit nicely into the above categories. This is due to the fact that, in both textbook

<b>10</b>	<p>Gegeven zijn de formules <math>y = 0,5x - 4</math>, <math>y = 2x + 2</math> en <math>y = 0,5x + 2</math>.</p> <p><b>a</b> Van welke formules snijden de grafieken de <math>y</math>-as in hetzelfde punt? Hoe kun je dat aan de formules zien?</p> <p><b>b</b> Van welke formules zijn de grafieken evenwijdig? Hoe kun je dat aan de formules zien?</p>
-----------	---

*Translation:*

Given the formulas  $y = 0.5x - 4$ ,  $y = 2x + 2$ ,  $y = 0.5x + 2$ .

**a** From which formulas do the graphs intersect the  $y$ -axis in the same point? How can you see this from the formulas? **b** For which formulas are their graphs parallel? How can you see this from the formulas?

Figure 6.5: Representative example of a third category activity of a grade 8 textbook. Reprinted from Reichard et al. (2005, p. 173) with permission of the publisher.

series, the introduction of the general formula for a straight line is not related to the contextual situations discussed in the preceding chapters. Instead, the definition of a ‘linear formula’ (GR) is given by means of examples of formulas, and that of ‘linear relation’ (MW), by means of equal growth of variables. In both textbook series, the introduction of the general formula for a straight line  $y = ax + b$  seems just a new track of activities. From an emergent modeling perspective, since there is no link in the textbooks between the contexts and the general formula, the chain of the development of contexts as models of linear relations to contexts as models for linear relations is broken.

Because of the missing link between the contextual situations of linear relations and the introduction of the general formula of a linear relation, activities in which students get acquainted with this formula do not fit in the emergent modeling approach. These activities are not part of a process of progressive mathematization. Rather, these activities seem to be a new track of activities, not grounded in contextual activities in the first and second categories. Therefore, we introduce a residual category of activities concerning the ‘bare introduction’ of the linear formula  $y = ax + b$ .

An example of an activity in this residual category is presented in Figure 6.6. In this fragment of text, a definition of linear relation is given, based on the equal growth of variables. In preceding tasks, students have had to calculate the equal growth from tables. So superficially, there seems to be a good connection between this fragment of theory and preceding tasks. However, a more profound analysis showed that in this fragment, no connection is made between this equal growth of the variables and the structure of the linear formula. The latter is presented in the example on the left-hand side by means of examples and non-examples. In these examples, no connection is made to the text about the equal growth. In the example on the right, the equal growth is demonstrated. However, the formula  $b = 4n + 3$  is not given. Also, the relation between the equal growth and

**THEORIE**

Als in een tabel in de bovenste rij opeenvolgende gehele getallen staan en in de onderste rij de toename steeds hetzelfde is, dan is er sprake van een **lineair verband**. De toename in de onderste rij getallen kan positief of negatief zijn. De bijbehorende grafiek is een rechte lijn. Zo'n grafiek heet een **lineaire grafiek**. Een formule waarvan de grafiek een rechte lijn is, noem je een **lineaire formule**.

**Voorbeeld**

Lineaire formules zijn bijvoorbeeld  
 $7x - 90 = y$  en  $B = 34 + 75t$ .  
 Niet-lineaire formules zijn bijvoorbeeld  
 $h = t^2 - 6$ ,  $g = 3a^2 + 7a$  en  $m - m^2 = d$ .

**Voorbeeld**

In de bovenste rij van de tabel hiernaast staan opeenvolgende gehele getallen.  
 In de onderste rij van de tabel is de toename steeds hetzelfde.  
 Tussen  $n$  en  $b$  bestaat een lineair verband.

		+1	+1	+1	+1	+1	
$n$	0	1	2	3	4	5	
$b$	3	7	11	15	19	23	
		+4	+4	+4	+4	+4	

*Translation:*

*If the upper row of a table contains subsequent integers and in the bottom row, the increase is constant, then there is a linear relation. The increase in the bottom row of numbers can be positive or negative. The corresponding graph is a straight line. Such a graph is called a linear graph. A formula whose graph is a straight line, is called a linear formula.*

*Example*

*Examples of linear formulas are  $7x - 90 = y$  and  $B = 34 + 75t$ . Examples of non-linear formulas are  $h = t^2 - 6$ ,  $g = 3a^2 + 7a$ , and  $m - m^2 = d$ .*

*Example*

*In the upper row of the adjoining table are subsequent integers. In the bottom row of the table, the increase is constant. Between  $n$  and  $b$  there exists a linear relation.*

Figure 6.6: Representative example of a fourth category activity of a grade 8 textbook. Reprinted from De Bruijn et al. (2008a, p. 15) with permission of the publisher.

Grade	<i>Getal &amp; Ruimte</i>					<i>Moderne Wiskunde</i>				
	C1	C2	C3	C4	Total	C1	C2	C3	C4	Total
7	26	5	–	–	31	63	1	–	9	73
8	1	8	2	27	38	3	12	6	27	48
9	2	4	–	33	39	2	1	–	10	13
10	–	–	–	15	15	–	–	–	6	6
Total	29	17	2	75	123	68	14	6	52	140

Table 6.2: Numbers of units of analysis in each category in the Dutch textbook series *Getal & Ruimte* and *Moderne Wiskunde*.

the formula is not discussed. In our view, the lack of these connections make this definition of a linear relation one that dangles in the air.

Because this text is not related to contextual situations, it does not fit into the first and second category of activities. Also, this text does not contribute to the development of conceptual proficiency due to the lacking connections. Therefore, we consider this text as an example of the fourth category.

Summarizing, in the analysis we distinguished four categories of activities. The first category consists of activities in the context; the second category consists of activities in which the focus shifts from the context itself to properties and characteristics of the context; and the third category consists of tasks that help students to develop conceptual proficiency.

### 6.3.3 Findings

Using the categories we defined above, we classified units of analysis in GR and MW. Table 6.2 provides an overview of the numbers of units of analysis in GR and MW in each category for grades 7, 8, 9, and 10. We recall that we took as units of analysis tasks that are denoted with a separate number and fragments of texts. The selected chapters and sections consisted of 123 (GR) and 140 (MW) units of analysis. From these units, 29 of GR and 68 of MW concerned activities from the first category; 17 of these units of GR and 14 of these units of MW concerned activities from the second category; 2 of these units of GR and 6 of these units of MW concern activities from the third category. These results show that both textbook series include a considerable amount of activities of the first and second category with an emphasis on the first category. So, both textbook series pay attention to the development of contexts to models of linear relations. However, both textbook series have only a small number of activities in the third category. This means that neither series provides many activities that offer students support in reifying linear relations.

However, in our view, the most important finding is the need for a fourth



category that emerged during the analysis. We argued that the instructional sequence of linear relations can be seen as two different tracks: one track started with activities in contextual situations and gradually, in a process of progressive mathematization, more formal mathematical knowledge is constructed. The other track introduced mathematical concepts by means of giving the formal definition, without linking this definition explicitly to the theory discussed in earlier chapters of the textbooks. In both textbook series, these two tracks of activities can be distinguished.

In addition, the number of activities in both textbook series in the third category was low, see Table 6.2. This means that the number of activities that help students to develop conceptual proficiency with linear relations was low. As a consequence, the growth from specific contextual situations into paradigms for linear relations is not adequately supported by the textbooks. Consequently, the contexts do not serve as a basis for developing paradigms of linear relations in the sense of Freudenthal (1991). Moreover, the low number of activities that support building on what is learned in the contexts illustrates the gap between the contextual track and the formal track.

The main differences between both textbook series are that MW had relatively more activities in the first category (68 of 140 in MW against 29 of 123 in GR), and GR had relatively more activities in the fourth category (75 of 123 in GR against 52 of 140 in MW).

## 6.4 Linear equations

### 6.4.1 Method

Analogously to Section 6.3.1, the data set is constructed from *Moderne Wiskunde* (MW) and *Getal & Ruimte* (GR). From these textbook series we selected chapters and sections concerning linear equations. Table 6.3 provides an overview of the selected chapters and sections for each grade.

Analogously to the categories of activities concerning linear relations, we initially distinguished three categories of activities, based on our sketch of the ideal instructional sequence described in subsection 6.2.2.

The first category consists of activities concerning the investigation of contextual situations. Examples of such activities are the development of informal strategies such as the cover-up method, and the use of the balance model to solve linear equations. Numbers in the linear equations represent different weights on the balance: standardized units and weights with a fixed, but unknown, weight.

For activities in the second category, informal strategies serve as the background. Strategies at this level involve adding and subtracting an equal number to the left-hand side and the right-hand side of the equation, and multiplying the equation by the same number.

Grade	Textbook	Chapter/Section	Edition
7	GR	–	–
7	MW	11	De Bruijn et al. (2007)
8	GR	5.4, 5.5, 5.6	Reichard et al. (2005)
8	MW	9	De Bruijn et al. (2008a)
9	GR	1.1	Reichard et al. (2010)
9	MW	9.1	De Bruijn et al. (2009b)
10	GR	2.1	Reichard et al. (2007)
10	MW	Blok 2	Bos et al. (2007)

Table 6.3: Textbooks and chapters concerning linear equations of *Getal & Ruimte* (GR) and *Moderne Wiskunde* (MW).

The third category consists of activities that no longer depend on the balance model. In this way, the balance model has become a model for reasoning about linear equations. Adding and subtracting the same multiple of the independent variable to the left-hand side and to the right-hand side of the equation is seen as a toggle of the sign. Activities in this category involve a broad view of linear equations. In this view, linear equations do not necessarily reduce to ‘ $x = \text{number}$ ’; equations can also reduce to  $6x + 4 = 6x + 4$  and  $6x + 4 = 6x + 2$  (see also Chapter 4). Equations and solutions can be interpreted in terms of lines and their intersections.

We took these three categories as the point of departure in the analysis of the Dutch textbook series.

### 6.4.2 Analysis

Based on the description of categories we described above, we divided the units of analysis into these categories. Below, we elaborate on the description of these categories by providing examples of activities in each category.

#### First category


Activities in the first category focus on the investigation of contextual situations. An example of such an activity is presented in Figure 6.7. In this task, students have to solve a contextual problem about bags with marbles. Four bags with an unknown, but equal, number of marbles, and eleven loose marbles equal six bags and three loose marbles. Students are invited to reason about this contextual problem in three subquestions. In our view, this activity is a first category activity because it is the context that is investigated by exploring informal strategies as a way to mathematize the problem.

**1** Daan heeft een aantal zakjes met in elk zakje evenveel knikkers. Hij geeft Erik vier zakjes en elf losse knikkers. Zelf pakt hij zes zakjes en drie losse knikkers. Daan zegt tegen Erik: 'We hebben nu allebei evenveel knikkers. Weet jij hoeveel knikkers er in een zakje zitten?' Hiernaast zie je de tekening die Erik maakt bij het probleem.

**a** Erik zegt: 'Dan is twee zakjes en drie losse knikkers evenveel als elf losse knikkers'. Leg uit waarom Erik gelijk heeft.

**b** Hoeveel knikkers zitten er dus in twee zakjes?

**c** Hoeveel knikkers heeft Daan in een zakje?



*Translation:*

*Daan has a number of bags: each bag has the same number of marbles. He gives Erik four bags of marbles and eleven loose marbles. He himself takes six bags and three loose marbles. Daan says to Erik: "We both have the same number of marbles. Do you know how many marbles there are in one bag?" On the side, you see the drawing Erik made from this problem.*




*a Erik says: "In that case, two bags and three loose marbles equals eleven loose marbles."*

*Explain why Erik is right.*

*b How many marbles do two bags contain?*

*c How many marbles does Daan have in one bag?*

Figure 6.7: Representative example of a first category activity of a grade 8 textbook. Reprinted from De Bruijn et al. (2008b, p. 74) with permission of the publisher.

Balans <span style="color: green;">■</span> = $x$ kg <span style="color: red;">■</span> = 1 kg	Zo schrijf je het op.
 <p>Haal van beide schalen 1kg af.</p>	$3x + 1 = 7$ <span style="color: red;">(-1)</span> <span style="color: red;">(-1)</span>
 <p>Zorg dat je op beide schalen het derde deel overhoudt, dus deel door 3.</p>	$3x = 6$ <span style="color: red;">:3</span> <span style="color: red;">:3</span>
 <p>Dus <math>x = 2</math>.</p>	$x = 2$

*Translation:*

*You write it up in this way. Remove 1 kg from both scales. Make sure you keep one third in both pans, so divide by 3. So  $x = 2$ .*


Figure 6.8: Representative example of a second category activity of a grade 8 textbook. Reprinted from Reichard et al. (2005, p. 183) with permission of the publisher.

### Second category

In the second category, we included activities in which informal strategies such as the balance model gradually shift to the background. An example of such an activity is presented in Figure 6.8. The equation presented in the right-hand side of the Figure,  $3x + 1 = 7$ , is solved. The balance model at the left-hand side supports the manipulations of the equation in the right-hand side. In this example, the balance model legitimates the different steps in the procedure of solving the equation.

### Third category

The third category consists of activities that contribute to obtaining a broad view of equations. Linear equations in these activities become objects that are part of a network of knowledge about how to operate on these objects. An example of such an activity is given in Figure 6.9. In this task, students have to solve the equations  $3x + 1 = 3(x + 2) - 5$  and  $3(x + 4) + 2(x - 1) = 5x + 8$ . Expanding the brackets in the first equation yields  $3x + 1 = 3x + 1$ , which holds for every  $x \in \mathbb{R}$ . Expanding the brackets in the second equation yields  $5x + 10 = 5x + 8$ , which does not hold for every  $x \in \mathbb{R}$ . According to the textbook GR, the medal



**U 15** **Bijzondere lineaire vergelijkingen**

Floris moet de vergelijking  $3x + 1 = 3(x + 2) - 5$  oplossen. Volgens zijn vader is elk getal een oplossing van deze vergelijking, dus de vergelijking heeft oneindig veel oplossingen.

**a** Ga na dat de vader van Floris gelijk heeft.

**b** Wat weet je van het aantal oplossingen van de volgende vergelijkingen?

$6(x + 1) = 3(2x + 2)$	$2(3x + 5) - 6(x - 1) = 16$
$5x - 3(x + 2) = 2(x - 3)$	$3(x + 4) + 2(x - 1) = 5x + 8$

Gegeven is  $3(x + 1) + 2x = 5x + p$ .

Kies je  $p = 6$ , dan krijg je de vergelijking  $3(x + 1) + 2x = 5x + 6$ .

Kies je  $p = -4$ , dan krijg je de vergelijking  $3(x + 1) + 2x = 5x - 4$ .

Welk getal moet je voor  $p$  kiezen om een vergelijking te krijgen

**c** met oneindig veel oplossingen

**d** die geen oplossingen heeft?

*Translation:*

*Special linear equations*

*Floris has to solve the equation  $3x + 1 = 3(x + 2) - 5$ . According to his father, every number is a solution to this equation, so the equation has infinitely many solutions. **a** Check that Floris's father is right. **b** What do you know about the number of solutions of the following equations? Given is  $3(x + 1) + 2x = 5x + p$ .*

*If you choose  $p = 6$  then you get the equation  $3(x + 1) + 2x = 5x + 6$ . If you choose  $p = -4$  then you get the equation  $3(x + 1) + 2x = 5x - 4$ . What number do you have to choose for  $p$  to get an **c** equation with infinitely many solutions? **d** equation that has no solutions?*

Figure 6.9: Representative example of a third category activity of a grade 9 textbook. Reprinted from Reichard et al. (2010, p. 14) with permission of the publisher.

below the number of the task indicates that this is a challenging task, which goes beyond the demands of the textbooks. In our view, this task provides students the opportunity to broaden the scope of linear equations, thus contributing to the development of conceptual proficiency.

### Additional fourth category

Similar to the analysis of activities concerning linear relations, during the analysis we found that both textbook series contain activities that do not fit nicely in the categories described above. We elaborate on this by a fragment of theory presented in Figure 6.10. In this fragment from the grade 9 GR textbook, the rule “change sides, change signs” is presented. By carrying the terms, something “striking” happens: minus signs disappear from one side of the equation and appear as plus signs on the other side of the equation. However, this rule is not related to preceding theory such as, for example, the balance model. Due to this missing link, we consider this activity as from a different category, because it

Het is handig als je voortaan in één keer van  $10x - 4 = 7x + 20$  naar  $10x - 7x = 20 + 4$  gaat. We zeggen dat de termen  $-4$  en  $7x$  naar het andere lid zijn overgebracht.

Bij dat **overbrengen van termen** gebeurt iets opvallends.

- $-4$  is uit het linkerlid verdwenen en als  $+4$  in het rechterlid terechtgekomen.
- $7x$  is uit het rechterlid verdwenen en als  $-7x$  in het linkerlid terechtgekomen.

**In een vergelijking mag je termen van het ene naar het andere lid overbrengen, maar dan moet je - vervangen door + en + vervangen door -.**

*Translation:*

*It is handy if from now on you go at once from  $10x - 4 = 7x + 20$  to  $10x - 7x = 20 + 4$ .*

*We say that the terms  $-4$  and  $7x$  have been carried over to the other side.*

*With the carrying over of terms something striking happens.*

- *$-4$  has disappeared from the left-hand side and has appeared as  $+4$  on the right-hand side.*
- *$7x$  has disappeared from the right-hand side and has appeared as  $-7x$  on the left-hand side.*

*In an equation it is allowed to carry terms over from one side to the other, but then you have to replace  $-$  by  $+$  and replace  $+$  by  $-$ .*

Figure 6.10: Example of a third category activity of a grade 9 textbook. Reprinted from (Reichard et al., 2010, p. 10) with permission of the publisher.

does not fit into the approach of progressive mathematization. Instead, the rule “change sides, change signs” is brought in as a mere fact.

### 6.4.3 Findings

We classified the units of analysis into the categories we described above. As we have mentioned before, the units of analysis are tasks that are denoted with a separate number. In case subquestions are classified at different activity levels, the whole task was given the highest level. Table 6.4 provides an overview of the numbers of activities in each category per grade for both textbook series. The selected Chapters and Sections consisted of 86 (GR) and 98 (MW) units of analysis. The topic of linear equations in GR is introduced in grade 8. As a consequence, the number of units of analysis is zero in grade 7. In grade 9 of MW, there is only one section in one chapter concerning linear relations and linear equations.

In GR, nine units of analysis consisted of activities in the first category; 40 in the second category; one in the third category; and 36 in the fourth category. In MW, 24 units of analysis consisted of activities of the first category; 51% of

Grade	<i>Getal &amp; Ruimte</i>					<i>Moderne Wiskunde</i>				
	C1	C2	C3	C4	Total	C1	C2	C3	C4	Total
7	–	–	–	–	0	16	23	–	–	39
8	9	40	–	–	49	8	26	4	8	46
9	–	–	1	15	16	–	–	3	–	3
10	–	–	–	21	21	–	–	4	6	10
Total	9	17	1	36	86	24	49	11	14	98

Table 6.4: Numbers of units of analysis in each category in Dutch textbook series *Getal & Ruimte* and *Moderne Wiskunde*.

the second category; 13% of the third category; and 11% of the fourth category. In both textbook series, the largest part of the activities concern the second category. In both textbook series, manipulations of equations are illustrated with adding and removing marbles (MW) or blocks (GR) to the pans of the balance. In this way, the balance model is used to legitimate manipulations such as adding 4 to both sides of the equation. However, in our view, this a narrow interpretation of the balance model. For example, the balance in both textbook series is always steady, thus suggesting that equations have at least one solution. Instead, the balance model could be used for equations such as  $6x + 4 = 6x + 2$  to illustrate that whatever  $x$  is,  $6x + 4$  does not equal  $6x + 2$ , so the equation does not have a solution.

From an emergent modeling perspective, the balance model initially emerges from activities in contextual situations to a model of linear equations. Then, ideally, by a process of progressive mathematization, the balance model gradually develops into a model for reasoning about linear equations. In the analysis, in both textbook series, we found a small number of activities in the third category that offer students support in developing conceptual proficiency in linear equations. Activities in the third category are important to obtain a broad view of equations. Part of this broad view is the connection between linear equations and linear relations with representations such as graphs and tables. This connection is not obvious. For example, Chapter 5 of GR (Reichard et al., 2005) concerns linear relations and linear equations. The sections concerning linear relations contain graphs. However, sections in the same chapter concerning linear equations do not contain any graph. In our view, the connection between linear relations and linear equations is not supported by the textbook in this way.

Another part of a broad view of linear equations is the notion of parameters in equations, and what role these parameters play in determining the relation. Obtaining such a broad view is sometimes even hindered. In MW, students are confronted with the equation  $3a + 7 = -5a + p$ , a linear equation in two variables (Figure 6.11). The first question is “why can’t you solve this equation?” So, in

- 16** Gegeven is de vergelijking  
 $3a + 7 = -5a + p$
- a** Waarom kun je deze vergelijking niet oplossen?
- b** Neem voor  $p$  de waarde 8.  
 Hoe ziet de vergelijking er nu uit? Los deze vergelijking op.
- c** Los de vergelijking ook voor  $p = -5$  op.
- d** Welke waarde heeft  $p$  als  $a = 22$ ? En als  $a = -\frac{1}{2}$ ?

*Translation:*

*Given is the equation  $3a + 7 = -5a + p$ . a Why can't you solve this equation? b Take for  $p$  the value 8. What does the equation look like then? Solve this equation. c Solve the equation for  $p = -5$ . d Which value has  $p$  if  $a = 22$ ? And if  $a = -\frac{1}{2}$ ?*

Figure 6.11: Example of a grade 8 task concerning unknown and variable, reprinted from De Bruijn et al. (2008b, p. 77) with permission of the publisher.

this task, the student is explicitly taught that it is not possible to solve equations with two variables, instead of being asked to solve the equation for  $a$  and then for  $p$ .

Analogously to linear relations, the instructional sequence of linear equations can be seen as consisting of two tracks, the first track starting with informal strategies that gradually become more formal, and the second track introducing theory without making the connection to the theory discussed in earlier chapters. For example, in the first chapter of GR in grade 9 the rule “change sides, change signs” is presented without any support for why these signs change, see Figure 6.10. Here, the obvious link to the balance model is not made. Instead, a new track, with the new rule “change sides, change signs” is presented.

Another example of activities in the second track is the introduction of systems of linear equations. Both textbook series introduce linear systems such as

$$\begin{cases} 4x - 2y = -2 \\ 3x + 5y = 31 \end{cases}$$

in grade 10 (Reichard et al., 2007; Bos et al., 2007). In GR, these systems are solved by multiplying, subtracting, and adding equations to eliminate one of the variables. In MW, in one equation, one variable is isolated and then substituted into the other equation. Neither textbook series underpins these manipulations by, for example, referring to the way linear equations are solved in preceding chapters. In addition, they do not make a connection between solving these systems of linear equations and determining the intersection points of lines presented in the grade 8 and grade 9 textbooks. In this way, these activities belong to the second track.



## 6.5 Conclusion and discussion

In this chapter, we investigated what kind of support Dutch textbook series offer to students to develop conceptual proficiency. Using emergent modeling (Gravemeijer, 1999) as the theoretical background, we classified activities of two Dutch textbook series into four categories. The analysis showed that both textbook series include activities in the first and second category. However, in the third category, for activities that offer students support in developing conceptual proficiency, both series have only a very small number of activities.

We do realize that the quantitative data obtained in this manner have their limitations. Other researchers might make other categories and arrive at different results. We maintain, however, that the data do point to both an imbalance in the instructional set up and two disjoint instructional tracks. Moreover, this observation is corroborated by a more qualitative inspection, which also revealed additional instances of the lack of attention for a conscious generalization and formalization of the more contextualized knowledge and understanding that the students may have gained from the informal explorations. In relation to this we mentioned the lack of connection between tables and graphs, the lack of attention for parameters, the ad hoc introduction of the rule, “changing sides, changing signs”, and the fact that multiplying, adding equations is not grounded in the earlier work on contextual problems.

From an emergent modeling perspective, the lack of attention for the development of mathematical relations and mathematical objects is problematic. The most important step in the development of a concept is the shift from activities at the referential level to activities at the general level. Models are supposed to support this switch by gradually developing from models of to models for. We conclude that the small number of activities at the general level is an impediment to the development of models in the case of linear relations and equations.

As regards linear relations, we argued that the switch from the referential to the general level involves a switch from contexts as models of linear relations to contexts as models for linear relations. At the general level, contexts serve as paradigms. From an RME perspective, paradigms are seen as important for learning universals as linear relationships (Freudenthal, 1991).

As regards linear equations, we found that both textbook series have a narrow interpretation of the balance model. Therefore we argued that the balance model does not contribute its full potential to a broad view of linear equations. The balance model functioned as a *model of* linear equations and did not grow to be a *model for*. Once it would have become a *model for*, situations with negative numbers would no longer be problematic, because at this level, the model no longer derives its meaning from the context. The need for the model to grow is also discussed by Vlassis (2002), who showed that the balance model can be helpful in giving meaning to the equal sign in equations with the unknown on both sides of the equation, but failed at being meaningful in more general

situations. Once equations become more complex and involve subtraction and negative numbers, and the balance model no longer applies, it may then become an impediment (De Lima & Tall, 2008). In our view, this impediment only occurs in case the balance model is still seen as a model of a linear equation.

Summarizing, both textbook series pay extensive attention to phenomenological exploration, with a considerable amount of activities that focus on the explanation of contextual problems. However, the most crucial step in the formation of concepts, in which students have to construct mathematical relations and mathematical objects, is hardly supported. Both textbook series seem to be torn between two ideas on how to teach mathematics. On the one hand an approach influenced by RME, starting with contexts. On the other hand a more classical approach of teaching mathematics by providing definitions and examples. In our view, balancing between these two different views on how to teach mathematics is unfortunate, because in this way, the instructional sequence becomes incoherent and many links are never made. In this way, the development of higher order skills, such as structure sense (see Chapter 3), and flexible skills as discussed in Chapter 4, are not supported by the textbook series.



# Chapter 7

## Conclusions

The studies described in this thesis aim to get a more detailed view of the student level of algebraic proficiency, and to get more insight into what causes the problems students experience in mastering algebra. In this final chapter, we first discuss the results of the studies described in this thesis. Next, we reflect upon the findings, and provide some recommendations.

### 7.1 Main findings

Based on the discussion in the Netherlands of the student level of algebraic proficiency, this thesis aimed at the following. We wanted to obtain more insight into the actual level of student algebraic proficiency. In addition, we aimed at obtaining insight into the possible determinants why this level is as it is.

#### 7.1.1 The development of algebraic proficiency

In the discussion on the student level of algebraic proficiency, several considerations played a role. The problems indicated by educators of secondary and higher education concerned the skills taught in lower secondary education. Part of the concerns in higher education emerged from the fact that students were perceived as being unable to solve simple mathematical problems and unable to do simple mathematical calculations. Examples of topics that were put forward concerned students' difficulties were solving linear equations, expanding brackets, calculating with square roots, and calculating with fractions. In part, these topics even concerned subjects that should already have been mastered in primary education. Furthermore, a call was heard from higher education for skills exceeding procedural skills. The subsequent discussion on the student level of algebraic

proficiency readily shifted to a debate concerning the proposed solutions and the classic dichotomy procedural fluency–conceptual understanding.

These considerations served as input for a study of the development of algebraic proficiency (Chapter 2). In order to obtain a more detailed view of the student level of algebraic proficiency, we constructed four tests for students of grades 8–12. As our theoretical background, we used two aspects of algebraic proficiency. The first aspect concerned the relation between basic skills and the deeper understanding that also plays a role in the aforementioned discussion. There is general agreement that basic skills have to go hand in hand with more conceptual aspects of algebraic proficiency (e.g. Kilpatrick et al., 2001), although different researchers provide different views of the nature of these more conceptual aspects. Arcavi (1994) introduced *symbol sense*, that he describes in terms of the behaviors related to skills that transcend basic manipulations, such as seeing the communicativity and power of symbols. Other behaviors relate to the ability to manipulate or to read through symbolic expressions depending on the problem at hand, and to flexible manipulation skills, such as the ability to cleverly select and use a particular symbolic representation. In the construction of the tests, we followed Drijvers (2006a, 2010) in viewing algebraic proficiency as a sliding scale ranging from basic skills to symbol sense. In the tests we aimed at including items that cover the full range from basic skills to symbol sense. An example of an item that we see as basic is: expand the brackets,  $-4(3a + b)$ , which yields  $-12a - 4b$ . An example of a symbol sense item is: is there any  $x$  for which  $\frac{2x+3}{4x+6} = 2$ ? If so, calculate  $x$ ; if not, explain why such an  $x$  does not exist. Arcavi (1994) argues that the ability to withstand the invitation to solve this equation directly, and instead perform an a priori inspection and conclude that the quotient equals  $\frac{1}{2}$  except for  $x = -1\frac{1}{2}$ , is an expression of symbol sense.

The second aspect concerned the transition from arithmetic to algebra, which is seen as crucial for the development of algebraic proficiency (e.g. Linchevski & Herscovics, 1996). We decided to include topics such as calculating with negative numbers, fractions, and square roots, that are related to the transition from arithmetic to algebra. Examples of items we included in the tests are: simplify  $2\sqrt{5} + 4\sqrt{5}$  (with correct answer  $6\sqrt{5}$ ); and, Jantine claims that  $2\frac{1}{2} \times 3\frac{1}{3} = 6\frac{1}{6}$ —explain why this is incorrect. An example of a correct answer to this item would be  $2\frac{1}{2} \times 3$  equals  $7\frac{1}{2}$  which is already greater than  $6\frac{1}{6}$ . For the complete list of items, refer to Appendix A.

Based on these two aspects, we constructed four sequential paper and pencil tests in a partly cross-sectional and partly longitudinal design. A thousand-odd students joined at least one of the sessions of the assessments, which were used to investigate how student algebraic proficiency developed from both a cross-sectional and a longitudinal perspective, and in terms of basic skills and symbol sense. For the analysis of our data, we used the Rasch model, a one parameter item response model (Rasch, 1980; Bond & Fox, 2007). With a Rasch analysis, one linear scale is created on which both persons are situated according

to their ‘ability’ and items according to their ‘difficulty.’ For each assessment, we created its corresponding Rasch scale, resulting in four Rasch scales. These four scales were connected via anchor items, resulting in one Rasch scale for all four assessments. As a consequence, items of all four assessments were placed on one scale. Also, students of different assessments have a Rasch measure on this very scale.

We found that students made progress, both cross-sectionally and longitudinally. In the cross-sectional development of students, the positive effect might in part be attributed to a curriculum change, whereas the new curriculum paid considerably more attention to algebraic skills in the lower grades. The longitudinal analysis showed that for the majority of students, the level of algebraic proficiency increased after one year. However, for the majority of students, this progression was not significant. With respect to different streams, we found that students from the science stream performed better than students from the social stream, but again the difference was not significant.

Relating student development to the individual tasks in the assessment, we found that students mastered simple tasks, but tasks became too complicated for them rather quickly. For example, the task “simplify:  $2\sqrt{5} + 4\sqrt{5}$ ” is mastered by the majority of the students, whereas the task “simplify:  $\frac{\sqrt{12}}{\sqrt{3}}$ ” is mastered by less than 10% of the students. In addition, the task “expand the brackets:  $-4(3a + b)$ ” is mastered by almost all students, whereas the task “simplify:  $(3a^2 + 2a + 7)(a + 8)$ ” is mastered by less than half of the students.

As a result, few tasks were mastered by the majority of students. This indicates that there is a severe problem in Dutch secondary education with the mastering of procedural skills. Furthermore, we saw that the range in which the development of the majority of students took place did not incorporate the tasks related to important conceptual aspects.

To get a better understanding of what the difficulties were and where they came from, we used various lenses. We used the notion of structure sense (Hoch & Dreyfus, 2004, 2006; Novotná & Hoch, 2008) and Sfard’s theory of reification (Sfard, 1991) to come to grips with the conceptual aspect. In addition, we used Van Gog & Paas’s construct of instructional efficiency to get a handle on fluency. Below, we elaborate on these analyses.

### Structure sense

In Chapter 3 we used the notion of structure sense (Linchevski & Livneh, 1999) to get a better handle on the conceptual aspects that seem to be problematic. Structure sense in high school algebra consists of several abilities including the ability to see the algebraic structure of a formula and the ability to see a part of an expression as a unit (Hoch & Dreyfus, 2004, 2006; Novotná & Hoch, 2008). More precisely, Novotná & Hoch (2008) define that students display structure sense if they can 1. recognize a familiar structure in its simplest form; 2. deal

with a compound term as a single entity and through an appropriate substitution recognize a familiar structure in a more complex form; 3. choose appropriate manipulations to make best use of a structure. We argued that structure sense can be seen as a part of symbol sense, which is a complex feel for symbols that includes a positive attitude towards symbols and a global view of expressions (Arcavi, 1994, 2005). From the assessments described above, we selected tasks that appeal to structure sense and focused the analysis on these tasks. We concluded that the majority of the students did not master tasks that require structure sense in order to be solved.

In addition we saw that the relations between the students' strategies, their symbol sense, and their positions on the Rasch scale, were not as clear as we expected based on the literature. In our view, this means that either the relation between strategy and symbol sense is not as strong as the literature suggests, or the relation between symbol sense and the underlying latent variable of the Rasch scale is troublesome. Too much attention to clever procedures may lead to a procedural approach towards very conceptual notions. Further research on the relations between strategies, symbol sense, and algebraic proficiency seems appropriate. In our view this touches a more fundamental problem with the notion of structure sense: it is both difficult to acquire and hard to measure.

Many researchers argued that flexibility, that can be seen as part of structure sense, are an important aspect of algebraic proficiency (Tall & Thomas, 1991; Sfard & Linchevski, 1994b; Kilpatrick et al., 2001; Stacey et al., 2004). This flexibility is especially of importance for students in higher education, for mathematics as well as for associated disciplines such as science and economics. In Chapter 3, we argued that one of the problems with Dutch mathematics education is that students do not develop structure sense. Arcavi (2005) argues that flexible skills include the ability to switch between applying rules and carrying out procedures on one hand, and reflecting upon these activities, and analyzing them on the other hand. An example in which this ability is useful is the simplification of  $\frac{5x^2+10-2(2x^2+4)}{x^2+2}$ . In this example, it helps to see that there is a relation between the numerator and the denominator. Also, it helps to see that the numerator equals 5 times  $x^2 + 2$  minus 4 times  $x^2 + 2$ . In our view, seeing this structure reduces the chance of errors. Novotná & Hoch (2008) argue that teaching mathematics only as a set of algorithms to be learned by rote might lead to a lack of understanding of the mathematical structures of the subjects. From this point of view, the notion of structure sense provides explanations for possible causes of students' difficulties with algebra. The results of the assessments suggest that Dutch mathematics education does not pay sufficient attention to this kind of skill. Therefore we call for more attention to structure sense in Dutch mathematics lessons.

### Sfard's theory of reification

Whereas Arcavi (2005) and Novotná & Hoch (2008) stress the strategies students use to solve problems, Sfard (1991) focuses on the conceptual barriers behind the difficulties which students experience.

In her theory of reification, Sfard (1991) presents a theoretical framework for advanced mathematical thinking based on what she calls the dual nature of mathematics. She argues that in mathematics, abstract concepts can be conceived of in two fundamentally different ways: structurally as objects, and operationally as processes. She bases this observation on an analysis of the historical development of algebra, which can be characterized as a long sequence of transitions from operational to structural conceptions. Again and again, processes performed on abstract objects at one level turn into new objects that in turn serve as objects of new higher level processes.

For example, a mathematical function can be seen as a method of computing values by performing some computational process. In this view, a function is a kind of recipe to get the output of a given input like “multiply the input by 3 and then add 2.” On the other hand, a function is an object with certain characteristics such as continuity and differentiability. Sfard (1991) argues that although abstract objects and computational processes look very different, in fact, they are “different sides of the same coin”.

In the process of concept formation, three stages are distinguished: interiorization, condensation, and reification. In the stage of interiorization, the learner gets acquainted with a process, for instance a function as a recipe. A process is interiorized if this process does not necessarily have to be carried out in order to be analyzed. In the second phase, condensation, a person becomes more capable of thinking about a process as a whole. Whereas interiorization and condensation are gradual processes, reaching the phase of reification is not gradual, but “a sudden ability to see something familiar in a totally new light” Sfard (1991, p. 19). Reification is the point where the object turns into input for interiorization on a higher level.

Sfard & Linchevski (1994b) identified two crucial transitions: the transition from the purely operational algebra to the structural algebra, and the transition from algebra of an unknown to the functional algebra of a variable. To Sfard & Linchevski (1994a), an important component of algebraic proficiency is a flexible approach in which a student is able to switch between an operational approach in which the focus is on the processes, and a structural approach in which the focus is on objects. These switches play a role in the process–product duality and the fixed–value–variable duality. We used these transitions to try to explain the student performance on the assessments described in Chapter 2. Scrutinizing the test items from Sfard's perspective showed a remarkable difference in performance of tasks concerning calculating with square roots. The task “simplify  $2\sqrt{5} + 4\sqrt{5}$ ” (with the correct answer  $6\sqrt{5}$ ) is mastered by almost 90% of the students, whereas the task “do you or do you not agree that  $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ ?”



is mastered by less than 10% of the students of grade 12. Further, the task in which students had to solve the linear equation  $2(3x + 2) = 3(2x - 1) + 7$  was mastered correctly by less than 10% of the grade 12 students. Based on the student performance on these tasks, we hypothesized that the problems grow in complexity according to the number of transitions that have to be made between an operational and a structural approach (or the reverse).

Testing this hypothesis among 92 grade 11 students yielded a confirmation of this hypothesis. In the case of square roots, the majority of the incorrect answers seemed to be due to a lack of the connection between the process of extracting the square root and a root as an object. As for the linear equations, the majority of students were able to expand the brackets correctly. Here, the problem was in drawing the correct conclusion.

The confirmation of the hypothesis led us to conclude that the majority of Dutch students do not have the ability to deal with the operational–structural duality in a flexible way. From the results discussed in Chapter 4 we conclude that the process–object duality is inherently difficult for Dutch students and that Dutch mathematics education apparently does not pay enough attention to fostering the process of reification.

### **Instructional efficiency**

Part of the discussion in The Netherlands on student algebraic proficiency concerned student fluency in solving algebraic tasks. In order to obtain a more detailed view of the student level of fluency in algebra, we decided to measure mental effort. Based on the principles of the cognitive load theory, students were asked to report their invested mental effort on a nine point symmetrical rating scale. We expected the measure instructional efficiency, which combines mental effort and performance, to provide a more detailed view of the student level of algebraic proficiency than performance alone. Although this measure was originally developed to compare instructional designs, we followed the suggestion of Van Gog & Paas (2008) to use this measure to evaluate the student level of expertise in algebraic proficiency. Trying to use the measure in this way, however, we found that this measure brings with it three methodological limitations.

The first limitation concerns the assumption in the construction of instructional efficiency that higher mental effort is counterbalanced by a fixed unit of higher performance and vice versa. The second limitation concerns ambiguous meanings of mental effort scores in situations of low performance and low mental effort. The third limitation concerns the validity of mental effort scores. Our data gave rise to questioning the validity of ‘mental effort’ since students seemed to mix mental effort and perceived difficulty.

Based on these limitations, we concluded that instructional efficiency is not an appropriate measure for investigating the student level of algebraic expertise. Due to our concern with the validity of the mental effort scores, we eventually

decided not to use the mental effort scores to determine the student level of expertise. We conclude that instructional efficiency did not help in obtaining a view of the student level of algebraic proficiency.

### 7.1.2 Analysis of Dutch textbook series

Because so many students do not reach the area of important conceptual aspects, and because textbooks play an important role in Dutch mathematics education Hiebert et al. (2003), we performed an analysis of two most used Dutch textbook series in Chapter 6. This analysis aimed at investigating which means of support Dutch textbooks offer to foster conceptual understanding. Dutch textbooks are influenced by the theory of Realistic Mathematics Education (RME). In line with RME, we used the theory of emergent modeling (Gravemeijer, 1999) as a framework to analyze the two most used textbook series: *Getal & Ruimte* and *Moderne Wiskunde*.

From an RME perspective, students should not be confronted with ready-made mathematics, because this would lead to an anti-didactical inversion in comparison to the way mathematics has been invented (Freudenthal, 1983). Rather, students should mathematize reality, and, if possible, mathematics (Freudenthal, 1968). Modeling activities play an important role in this process. By translating a contextual problem into a mathematical problem, the problem solver makes the problem amenable to mathematical procedures. From an RME perspective, students should not be confronted with ready-made models either. Rather, these models should emerge from their own mathematizing activities (Gravemeijer, 1994), thus supporting mathematical growth. In this way, the model emerges as a *model for* mathematical reasoning. The transition from *model of* to *model for* concerns a shift in student thinking. This shift passes from thinking about the modeled contextual situation to thinking about mathematical relations. Related to the development of the models, Gravemeijer (1999) discerns four levels of different activities: task setting, referential, general, and formal.

Task setting activities concern reasoning about interpretations and solutions in a specific setting. At the referential level, models derive their meaning from the activities in the task setting they refer to. At the general level the models start to derive their meaning from the mathematical relations that the students develop while reflecting on their activities with the models. Activities at the level of formal mathematical reasoning are no longer dependent on models. Although there is a hierarchy in these levels, a student's activities do not take place in a strict order. Discussions about activities at the general level for instance may fold back to the referential or task setting level. In the transition from referential to general activities, models are reified, thus becoming entities on their own. At this point, Gravemeijer (1999) makes a connection to the reification of processes (Sfard, 1991). Following this line of thought, Gravemeijer (1999)

argues that it is the activity that is being reified.

We focused the analysis on two topics introduced in lower secondary education, linear relationships and linear equations, because in Chapter 2 we found that these topics were among the most difficult tasks. A linear relationship between two variables is a relation in which any given change in one variable will always produce a proportional change in the other variable. The graph of a linear relationship is a straight line. By linear equation, we refer to a polynomial equation in one variable in the first degree, for example  $4x - 2 = 19$  or  $8x - 14 = -5x + 12$ .

Using emergent modeling as our theoretical background, we sketched the ideal instructional sequence of linear relations and linear equations from this point of view. As for linear relations, the introduction should start with contextual situations of linear relations that invite students to reason and calculate within these contexts. Next, the activities should shift to their mathematical characteristics and the mathematical relations. As the attention shifts to the mathematical relations involved, the model becomes detached from the contextual situations, and the students' reasoning loses its dependence on situation-specific features. In this way, a linear relation becomes an object. This object does not necessarily need its context, but instead has meaning in and of itself. This object is incorporated in a network of relations. At this general level, linear (word) formulas start to function as 'models for' reasoning about linear relations.

From the same perspective, we argued that the ideal introduction of linear equations starts with informal strategies at the task setting level, such as the cover-up method and the balance method, that fit contextual situations. Next, at the referential level, these informal strategies serve as the background for solving equations and for interpreting and giving meaning to the solution. Gradually, the balance model becomes more detached from its specific situation and emerges as a model for reasoning about linear equations. At the general level, equations have become reified to objects and are part of a network of mathematical relations that justify how to operate on these objects. For example, adding and subtracting numbers and variables on both sides of the equation and multiplying both sides of the equation by the same number.

In the analysis, we chose not to make a strict connection between the four levels of activity described in the theory of emergent modeling and activities in the textbooks for two reasons. The first reason is that the levels describe the students' mental activity, not the tasks or the models per se. The second reason is that in the theory of emergent modeling, the shifts between levels is seen as important. This aspect would not be done justice to if we had classified activities at the four levels. Rather, the way this theory described the construction of formal mathematical knowledge gave rise to distinguishing three categories of activities for linear relations as well as linear equations.

In the first categories, we included activities that ask students to reason and calculate within contexts. Examples of such activities are searching for

patterns in given contextual situations and constructing word formulas (linear relations), and developing informal strategies such as the cover-up method, and using the balance model (linear equations). The second categories consisted of activities in which the focus shifts from the context itself to its properties and characteristics. Examples of such activities are searching for patterns in the relation, reasoning and calculating with formulas, and comparing properties of relations (linear relations). As regards linear relations, the second category consisted of activities in which informal strategies serve as background. Activities in the third category offer support to seeing and treating linear relationships as objects that do not necessarily derive their meaning from the context, but have meaning by themselves and by their relations to other mathematical objects. As for linear equations, we included activities that help students in obtaining a broad and flexible view of equations. In this view, linear equations do not necessarily reduce to ' $x = \text{number}$ '; equations can also reduce to  $6x + 4 = 6x + 4$  and  $6x + 4 = 6x + 2$  (see also Chapter 4).

The data set is constructed from two textbook series, *Moderne Wiskunde* (MW) and *Getal & Ruimte* (GR), that together have an estimated market share of over 90% (cTWO, 2009). From these textbook series, chapters, and within chapters, sections, were selected that concern linear relations. Theory on these topics is presented in the textbooks for grades 7, 8, 9, and 10. In grades 10, 11, and 12, the student population is split into a social stream and a science stream. We focused on students in the science stream because especially for these students it is important to develop conceptual proficiency.

The analysis indicated that Dutch textbook series pay relatively much attention to activities in the first and the second category. In addition, both textbook series hardly contain activities that guide students from the context to more formal mathematics. Moreover, we found two disjunct didactical tracks: one track that follows the RME approach, and one track with a more traditional approach, in which new concepts are introduced as ready-made mathematics. The RME track mainly contains activities at the task setting and referential levels. We found only a few activities that support the switch from contexts to formal mathematics. The other track does not build on the contexts and models, thus missing the link to the student's prior knowledge. Because of these two disjunct tracks, we concluded that these two Dutch textbook series do not offer a consistent instructional sequence and therefore hamper the development of conceptual proficiency.

The consequence of this is that students are not supported in reaching the stage of reification, in the words of Sfard (1991), which is a prerequisite for a flexible approach in which the student is able to switch between an operational perspective in which the focus is on the processes, and a structural perspective in which the focus is on objects (Sfard & Linchevski, 1994b). This higher level is known and described by many researchers, using just as many names, such as procept (Gray & Tall, 1994); encapsulation (Dubinsky, 1991); generalized

substitution principle (Wenger, 1987); processes at the lower level become objects at a higher level (Freudenthal, 1991); network of relations between relations (Van Hiele, 1973); and construction of some new mathematical reality via emergent modeling (Gravemeijer, 1999).

Gray & Tall (1994) argued that an alternative approach could be to start with processes rather than with ready-made algebraic objects. Curricula often assume a structural approach even though the process–object duality is not grasped by the students (Sfard & Linchevski, 1994b). For example, a letter is introduced as a variable and not just as an unknown, even though from a historical perspective, the latter precedes the former, which is more advanced. In this way, the curriculum reverses the order of concept formation.

### 7.1.3 Overall findings

This thesis aimed at obtaining more insight into the actual level of Dutch student algebraic proficiency, and at obtaining insight into possible causes of this level. From the results we described above, we may argue that in general, the student level of algebraic proficiency is disappointing. We found that students made progress, both cross-sectionally and longitudinally, but this progress is small. Furthermore, we saw that students master simple tasks, but tasks become too complicated for them rather quickly, and we saw that the range in which the development of the majority of students took place did not incorporate the tasks involving important conceptual aspects.

A more detailed view of the student level of algebraic proficiency was obtained by using the notion of structure sense (Linchevski & Livneh, 1999; Novotná & Hoch, 2008) and Sfard’s theory of reification (Sfard, 1991). Using structure sense as a theoretical lens provided insight into the kind of ability Dutch students lack. These abilities include the ability to see the algebraic structure of a formula, and the ability to see a part of an expression as a unit. We argued that one of the problems with Dutch mathematics education is that students do not develop structure sense. Using the theory of reification (Sfard, 1991) provided insight into the conceptual deficiencies that explain why students experience difficulties while solving mathematical problems. We found that flexibility with the process–object duality of mathematical entities is at the core of algebraic proficiency and that this flexibility is hard to achieve for many students.

From these results, we concluded that the majority of students did not develop *conceptual proficiency*, which we defined as a mixture of higher-order skills in which three aspects are essential: the ability to recognize the algebraic structure and use it flexibly, the ability to deal with the ambiguous nature of mathematical concepts, and the ability to recognize coherence between mathematical concepts. Because the majority of Dutch students do not develop conceptual proficiency and because textbooks play an important role in Dutch mathematics education, we decided to perform an analysis of two Dutch textbook series in

order to investigate what kind of support the textbooks offer for developing conceptual proficiency. The analysis, with emergent modeling (Gravemeijer, 1999) as our theoretical background, showed that textbooks hardly provide this kind of support. Moreover, we found two disjunct tracks of teaching mathematics: one track that follows the RME approach, and one track with a more traditional approach, in which new concepts are introduced as ready-made mathematics. The analysis showed that there are hardly any activities that help students to bridge the gap between these two tracks. We concluded that Dutch textbook series do not have a consistent instructional sequence at their disposal, and thus do not support students in the development of conceptual proficiency.

The theoretical frames we used in this thesis have served as the underpinning of our findings. At the same time, this thesis in turn contributes to the theory of mathematics education. In Chapter 3, we brought the notion of structure sense (Linchevski & Livneh, 1999; Novotná & Hoch, 2008) into practice and discussed the relation of structure sense to algebraic proficiency. Chapter 4 builds on Sfard's theory of reification (Sfard, 1991). We brought this theory into practice and defined some indicators of complexity in relation with this theory. The theory of emergent modeling (Gravemeijer, 1999) served as a theoretical lens in the analysis of some textbooks in Chapter 6. In Chapter 5, we argued that in the context of our study on the development of student algebraic proficiency, instructional efficiency is not an appropriate measure for evaluating a learner's level of expertise.

## 7.2 Discussion

### 7.2.1 Limitations

The studies described in this thesis aim at obtaining insight into the actual level of student algebraic proficiency in pre-university education. And, if this level is disappointing, to obtain insight into the deeper causes of the problems students experience in the transition from pre-university education to higher education.

These difficulties have a complicated nature because many factors play a role. The studies reported in this thesis all have their limitations because we simply cannot take into account all factors. The most natural limitations concern the number of schools and students that participated in the assessments. Four schools participated, together with more than 1000 students. In the fourth and final assessment, one school withdrew, so the fourth assessment concerned students of the three remaining schools. After two assessments, the summer started, after which students were situated one grade higher, sometimes in different classes. As a consequence, it was difficult to find students because they were mixed with students of other classes, or had even left school. As a result, only 277 of the more than 1000 students completed all four assessments. This reduced the number of students in the longitudinal analysis.

Although we do have to be careful with generalizing the results we found, we consider the findings described in this thesis representative because they are in line with the findings of similar research among students of the same age (Bruin-Muurling, 2010) and among university students (Kraker & Sauren, 2004; Caspers, 2007).

The assessments described in Chapter 2 are constructed to cover several different topics within algebra (expanding brackets, simplifying, and solving equations) and arithmetic (negative numbers, square roots, and fractions). The explorative character of this research is demonstrated by the choice of assessing different mathematical processes. An advantage of this approach is that we search for possible causes of problems in a larger part of the curriculum. A disadvantage is the small number of items per theme, which has as a consequence that the relation to the underlying theoretical concepts is less clear. We have overcome this by focusing the analysis on a single theoretical concept in Chapters 3 (structure sense) and 4 (theory of reification). It is left for future research to construct more extensive tasks more directly related to the underlying concepts, although we would like to note that the relation between tasks and concepts is complicated.

The purpose of Chapter 2 was to obtain a view of the student level of algebraic proficiency. To obtain this view, more than 1000 students participated at least once in this research project. The written answers of the students served as input in the analysis. In this way, we draw conclusions on student thinking from their written answers. However, the written answers do not always reflect what students have been thinking. In order to better get to grips with their thinking, we asked students to reflect on their written answers in an interview, see Chapter 4.

The difficulties Dutch students have with algebra in secondary education can not be fully interpreted without taking into account the role of the teachers and what actually happens in the classroom. Therefore, future research should focus on our findings in combination with a more participationist approach (Sfard, 1998) in which, for example, the social norms and the discourse in the classroom are taken into account.

Below, we reflect on the above findings and place them in a broader perspective.

### 7.2.2 Reflection

This thesis provides insight into student development of algebraic proficiency. We have seen that the number of tasks that is mastered by 75% of the students is low. Examples of tasks that are mastered by the majority of the students are

- Expand the brackets:  $-4(3a + b) =$
- Simplify:  $3\sqrt{7} + 2\sqrt{7} =$

- Calculate:  $\frac{18-5}{7-20} + 4 =$

Examples of tasks that are mastered by less than 25% of the students are

- Simplify:  $\frac{8}{21} - \frac{2}{7} + \frac{1}{14}$
- Simplify:  $\frac{\sqrt{12}}{\sqrt{3}} =$
- Solve:  $2(3x + 2) = 3(2x - 1) + 7$
- Solve:  $(x - 5)(x + 2)(x - 3) = 0$ .

We consider this to be problematic; students should be more proficient in procedural skills. However, we think that the problem is also on conceptual aspects of algebraic proficiency. In the longitudinal analysis we found that the majority of students do not make significant progress during one year mathematics education. These problems and difficulties concern students of all grades, and in upper-secondary education, they concern students of both social and science streams. Therefore in our view Dutch mathematics education does not sufficiently address and develop students' algebraic proficiency. At the same time, the impression emerges that mathematics education makes suboptimal use of students' talents, as it is hard to believe that understanding why  $2\frac{1}{2} \times 3\frac{1}{3}$  does not equal  $6\frac{1}{6}$  is too difficult for more than half of the grade 12 students in pre-university education.

In addition we argue that there is inadequate to almost no student development in the area of the conceptual aspects of algebraic proficiency such as structure sense and the process-object duality. These conceptual aspects include flexible skills which are important in different sectors of higher education. The following examples (McCallum, 2010) illustrate that the demands in higher education exceed the level of superficial procedural fluency:

- recognizing that  $P \cdot \left(1 + \frac{r}{12}\right)^{12n}$  is linear in  $P$  (finance);
- identifying  $\frac{n(n+1)(2n+1)}{6}$  as being a cubic polynomial with leading coefficient  $\frac{1}{3}$  (calculus);
- observing that  $L_0\sqrt{1 - \left(\frac{v}{c}\right)^2}$  vanishes when  $v = c$  (physics);
- understanding that  $\frac{\sigma}{\sqrt{n}}$  halves when  $n$  is multiplied by 4 (statistics).

The skills needed for these demands require flexibility and feeling for the mathematical structure of formulas and expressions. However, we have argued that students hardly develop structure sense and symbol sense. Attention to structure is seen as an important part of mathematics and of the learning of algebra in particular (Novotná & Hoch, 2008). In our view, mathematics teachers should pay more attention to the structure and the coherence of mathematics. This could be arranged by stimulating classroom dialogues on the structure of formulas and on generalizations.



For example, when treating quadratic equations, students learn that the equation  $(x - 1)(x + 3) = 0$  implies  $x - 1 = 0$  or  $x + 3 = 0$ , which is an application of the mathematical rule

$$A \cdot B = 0 \Rightarrow A = 0 \vee B = 0.$$

In our view, at this point, students should come to see that a generalization can be made to products of more than two factors, enabling them to solve more complex problems such as  $(x - 1)(x + 3)(x - 4) = 0$ . Since we have seen in Chapter 2 that less than 10% of the grade 12 students masters this equation, it is clear that students are not aware of the generalization. Discussing mathematical structures and generalizations could contribute to establishing connections between different approaches. Here we may note that we do not argue in favor of a transformation of the current curriculum with tasks that focus only on symbol sense, because in our view, tasks themselves do not embody symbol sense. Rather, it is the activity of solving the problem and the discussion on the problem that determines whether the problem supports the construction of symbol sense.

Another conceptual aspect is the ambiguous nature of mathematics. Sfard (1991) argues that flexibility with the process–object duality of mathematical entities is at the core of algebraic proficiency and that this flexibility is hard to achieve for many students. In addition she argues that reaching the stage of reification is necessary in order to flexibly switch between an operational and a structural approach. In Chapter 4 we found that dealing with the process–object duality and the fixed-value–variable duality is indeed difficult for students. In our view, teachers should pay attention to the ambiguous nature of mathematics by, for example, discussing the different roles of letters as fixed-values or as variables. Another example is the ambiguous nature of fractions and square roots that can be seen as processes as well as objects. Discussing these dualities might help students to accept  $\sqrt{2}$  as an answer instead of searching for an appropriate approximation because they consider  $\sqrt{2}$  as a process. In addition, teachers should stimulate students to search for reification. For example, teachers can stimulate students not to be satisfied with performing tasks only, but instead to ask themselves to find the bigger picture of which the task is part of. In this way, the focus is not on performing tasks, but on searching for the mathematical concepts that are supposed to be expressed in the task, and the coherence between tasks and concepts.

From the analysis of two Dutch textbook series we found that contexts in these textbook series often concern oversimplified situations. Although we agree that situations have to be simplified because problems otherwise become too complicated, in our view, teachers should discuss these oversimplifications because students know very well that these kinds of contexts do not relate to problems they might experience in their daily life. In addition, students have to develop a sense for which details of the context are seen as relevant and which are seen as irrelevant. For example, *Moderne Wiskunde* (De Bruijn et al., 2008b) introduces

linear equations with pictures of a balance with bags and marbles (see also Chapter 6). Each bag is filled with a fixed, but unknown, number of marbles. Four bags with 11 loose marbles are in equilibrium with six bags and 3 loose marbles. The question is how many marbles there are in each bag. In this example, the weight of the bags is ignored without discussing whether this is reasonable. This means that there is an implicit understanding among the authors of the textbooks about which issues can be ignored and which cannot. However, it is unclear how students are supposed to know this (Gravemeijer, 2004).

Dutch textbook series are influenced by the theory of Realistic Mathematics Education. From this perspective, students are supposed to use their own situated informal knowledge to construct more formal mathematical knowledge in a process of guided reinvention and progressive mathematization. Analysis in Chapter 6 of the two most-used textbook series indicates that Dutch textbook series hardly contain activities that guide students from the context to more formal mathematics. Moreover, we found a double disjunct track: one track RME and one track more formal traditional mathematics in which new concepts are introduced by definitions. This means that formal strategies are not built on in context-rooted, semi-formal strategies. Ideally, contexts serve as the starting point and enable students to start a process of progressive mathematization of both the context and their own informal mathematical knowledge. In this view, students construct their mathematical knowledge. In the current situation, however, Dutch students are supposed to discover and construct formal mathematical knowledge with little support from the textbooks. In Dutch mathematics education, students are used to working independently on the tasks presented in the textbooks. Also the introduction of new content is often accomplished by students working with textbooks (Hiebert et al., 2003). From this point of view, the role of Dutch textbook series is relatively large, which increases their impact.

We have argued that Dutch textbooks pay relatively too much attention to exploring activities within contexts, whereas in our view, these activities do not require the most attention. It is not the contextual situation itself that is difficult to get to grips on, it is the underlying mathematical structure of the situation. Given the present textbook series, teachers could help students to use the contexts and models to construct formal mathematics. For example, teachers could discuss the limitations of contexts and how these limitations might be overcome in more general situations. Furthermore, teachers could help students to bridge the gap between the two disjunct didactical tracks by offering activities that help to make the shift towards formal mathematics.

Publishers should reconsider the way contexts and models are used in the present textbooks. More specifically, they should reconsider the way students can build more formal mathematical knowledge based on context-rooted, semi-formal strategies. In addition, they should include activities that help students to develop conceptual proficiency, which we defined as a mixture of higher-order thinking skills in which three aspects are essential: the ability to recognize and

make flexible use of the algebraic structure, to deal with the ambiguous nature of mathematical concepts, and see the coherence between mathematical concepts.

We have argued that more attention should be paid to procedural fluency, mathematical structure, coherence, and ambiguity. Arcavi (2005) argued that algebraic proficiency requires students to flexibly switch between meaningless actions (automatic application of rules and procedures) to sense making in the way of “symbol sense.” In other words, students have to be able to postpone meaning in order to apply procedures quickly. And also, to “unclog an automatism” as Freudenthal (1983, p. 469) put it, to reconsider, to reflect, or to question, if necessary.

As a caveat we want to stress that structure sense has to be paired to, and depends on, procedural fluency. Algebraic proficiency asks for both fluency and conceptual understanding, as also is expressed in Arcavi’s notion of symbol sense and Sfard’s theory of reification.

In addition we have argued that teachers could help students to bridge the gap between the two didactical tracks in the textbook series. We are well aware that these changes require a different view on what is important in mathematics and in mathematics education. In relation to this, Yackel & Cobb (1996) speak of sociomathematical norms. These norms concern “normative understandings of what counts as mathematically different, mathematically sophisticated, mathematically efficient, and mathematically elegant in a classroom” (Yackel & Cobb, 1996, p. 461). Changing these norms requires a renegotiation of the didactical contract (Brousseau, 1990), which is described by Gravemeijer (1997) as “the set of reciprocal expectations and obligations between the teacher and the students that has evolved in their ongoing interaction.” In this new didactical contract, the focus should come to lie on the mathematics behind the tasks instead of on performing tasks. In this way, mathematics is no longer a collection of disconnected facts and procedures; rather, it is about searching for underlying structures, searching for abstractions and relations. This requires a shift in a student’s thinking as described by Yackel & Cobb (1996, p. 471):

“This shift in students’ thinking is analogous to the shift between process and object that Sfard (1991) describes for mathematical conceptions. In the same way that being able to see a mathematical entity as an object as well as a process indicates a deeper understanding of the mathematical entity, taking an explanation as an object of reflection indicates a deeper understanding of what constitutes explanation.”

We believe that in this way, the strength and power of mathematics becomes more visible for students. This will make mathematics more challenging for the students, and mathematics as a discipline more interesting. In addition, the attitude that simply performing the tasks is not sufficient mathematical

preparation, is an attitude that is fitting for students who are to eventually enter higher education.



# Bibliography

- Arcavi, A. (1994). Symbol sense: Informal sense-making in formal mathematics. *For the Learning of Mathematics*, 14(3), 24–35.
- Arcavi, A. (2005). Developing and using symbol sense in mathematics. *For the Learning of Mathematics*, 25(2), 42–47.
- Ayres, P. (1993). Why goal-free problems can facilitate learning. *Contemporary Educational Psychology*, 18(3), 376–381.
- Ayres, P. (2001). Systematic mathematical errors and cognitive load. *Contemporary Educational Psychology*, 26(2), 227–248.
- Bond, T. G. & Fox, C. M. (2007). *Applying the Rasch Model. Fundamental Measurement in the Human Sciences. Second Edition*. Mahwah, NJ, USA: Lawrence Erlbaum.
- Boon, P., Van Kervel, E., Lagerwaard, K., Limpens, G., Severins, E., & Stroomer, G. (2002). Wiskunde-examens 2002, 1e tijdvak. *Euclides*, 78(1), 4–18.
- Bos, D., Bouw, J., Doekes, W., Knippers, W., Van der Linden, T., & Roorda, G. (2007). *Moderne Wiskunde VWO B deel 1, 9e editie*. Groningen: Wolters-Noordhoff.
- Bratfisch, O., Borg, G., & Dornic, S. (1972). Perceived item-difficulty in three tests of intellectual performance capacity. *Reports from the Institute of Applied Psychology, University of Sweden*, 29, 1–12.
- Brousseau, G. (1990). Le contrat didactique: le milieu. [The didactical contract: The setting]. *Recherches en Didactique des Mathématiques*, 9(3), 309–336.
- Bruin-Muurling, G. (2010). *The development of proficiency in the fraction domain. Affordances and constraints in the curriculum*. PhD thesis, Eindhoven School of Education, Eindhoven University of Technology.

- Byers, W. (2007). *How mathematicians think: Using ambiguity, contradiction, and paradox to create mathematics*. New Jersey: Princeton University Press.
- Caspers, W. (2007). *Wi 1000, Instaptoets en Opfrisonderwijs aan de TU Delft 2006 (Opzet, Resultaten en Analyse)*. Technical report, TU Delft.
- CEVO (2006). *Syllabus Wiskunde C vwo (definitief concept, 31 - 05 - 2006)*. Technical report, CEVO commissie herziening examenprogramma 2007 wiskunde C vwo.
- Cobb, P., Yackel, E., & Wood, T. (1992). A constructivist alternative to the representational view of mind in mathematics education. *Journal for Research in Mathematics Education*, 23(1), 2–33.
- Cooper, G. & Sweller, J. (1987). Effects of schema acquisition and rule automation on mathematical problem-solving transfer. *Journal of Educational Psychology*, 79(4), 347–362.
- cTWO (2009). *Rapport tussenevaluatie van de 2007-programma's wiskunde havo/vwo*. Technical report, cTWO, Retrieved 2 September 2010 from <http://www.ctwo.nl/>.
- De Bruijn, I., Doekes, W., Van der Eijk, E., Greefkens, S., Koens, T., Kok, D., Nauta, E., Van Proosdij, G., Roll, C., Schaberg, G., & Sinkeldam, R. (2007). *Moderne Wiskunde vwo deel 1B, 9e editie*. Groningen: Wolters-Noordhoff.
- De Bruijn, I., Van der Eijk, E., Eikelboom, A., Greefkens, S., Hietbrink, H., Koens, T., Kok, D., Nauta, E., Van Proosdij, G., Roll, C., Schaberg, G., & Sinkeldam, R. (2009a). *Moderne Wiskunde VWO deel 3A, 9e editie*. Groningen: Wolters-Noordhoff.
- De Bruijn, I., Van der Eijk, E., Eikelboom, A., Greefkens, S., Hietbrink, H., Koens, T., Kok, D., Nauta, E., Van Proosdij, G., Roll, C., Schaberg, G., & Sinkeldam, R. (2009b). *Moderne Wiskunde VWO deel 3B, 9e editie*. Groningen: Wolters-Noordhoff.
- De Bruijn, I., Van der Eijk, E., Greefkens, S., Koens, T., Kok, D., Nauta, E., Van Proosdij, G., Roll, C., Schaap, W., Schaberg, G., & Sinkeldam, R. (2008a). *Moderne Wiskunde vwo deel 2A, 9e editie*. Groningen: Wolters-Noordhoff.
- De Bruijn, I., Van der Eijk, E., Greefkens, S., Koens, T., Kok, D., Nauta, E., Van Proosdij, G., Roll, C., Schaap, W., Schaberg, G., & Sinkeldam, R. (2008b). *Moderne Wiskunde VWO deel 2B, 9e editie*. Groningen: Wolters-Noordhoff.
- De Jong, T. (2010). Cognitive load theory, educational research, and instructional design: Some food for thought. *Instructional Science*, 38(2), 105–134.

- De Lima, R. & Tall, D. (2008). Procedural embodiment and magic in linear equations. *Educational Studies in Mathematics*, 67(1), 3–18.
- Drijvers, P. (2003). *Learning algebra in a computer environment. Design research on the understanding of the concept of parameter*. PhD thesis, Utrecht: CD- $\beta$  press.
- Drijvers, P. (2006a). Context, abstractie en vaardigheid in schoolalgebra. *Nieuw Archief voor Wiskunde*, 7(3), 198–203.
- Drijvers, P., Ed. (2006b). *Wat a is, dat kun je niet weten*. Freudenthal Instituut, Universiteit Utrecht.
- Drijvers, P., Ed. (2010). *Secondary algebra education. Revisiting topics and themes and exploring the unknown*. Rotterdam: Sense.
- Dubinsky, E. (1991). Reflective abstraction in advanced mathematical thinking. In D. Tall (Ed.), *Advanced Mathematical Thinking* (pp. 95–123). Dordrecht: Kluwer Academic Publishers.
- Filloy, E. & Rojano, T. (1989). Solving equations: The transition from arithmetic to algebra. *For the Learning of Mathematics*, 9(2), 19–25.
- French, D. (2002). *Teaching and Learning Algebra*. London: Continuum International Publishing Group.
- Freudenthal, H. (1968). Why to teach mathematics so as to be useful. *Educational Studies in Mathematics*, 1(1–2), 3–8.
- Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics*, 3(3–4), 413–435.
- Freudenthal, H. (1973). *Mathematics as an Educational Task*. Dordrecht, The Netherlands: Reidel.
- Freudenthal, H. (1983). *Didactical Phenomenology of Mathematical Structures*. Dordrecht: Reidel.
- Freudenthal, H. (1991). *Revisiting mathematics education: China lectures*. Dordrecht: Kluwer.
- Gerjets, P., Scheiter, K., & Cierniak, G. (2009). The scientific value of cognitive load theory: A research agenda based on the structuralist view of theories. *Educational Psychology Review*, 21(1), 43–54.
- Gravemeijer, K. (1990). Globaal kijken, een kenmerk van algebraïsche deskundigheid. *Nieuwe Wiskrant, Tijdschrift voor Nederlands Wiskundeonderwijs*, 10(2), 29–33.



- Gravemeijer, K. (1994). *Developing Realistic Mathematics Education*. PhD thesis, Utrecht: CD- $\beta$  press.
- Gravemeijer, K. (1997). Commentary solving word problems: a case of modelling? *Learning and Instruction*, 7(4), 389–397.
- Gravemeijer, K. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning*, 1(2), 155–177.
- Gravemeijer, K. (2004). Creating opportunities for students to reinvent mathematics. In *Regular lecture at the 10th International Congress on Mathematical Education (ICME 10), Copenhagen, Denmark, Copenhagen*.
- Gray, E. & Tall, D. (1992). Success and failure in mathematics: The flexible meaning of symbols as process and concept. *Mathematics Teaching*, 142, 6–10.
- Gray, E. & Tall, D. (1994). Duality, ambiguity, and flexibility: A proceptual view of simple arithmetic. *Journal for Research in Mathematics Education*, 25(2), 116–140.
- Harper, E. (1987). Ghosts of Diophantus. *Educational Studies in Mathematics*, 18(1), 75–90.
- Herscovics, N. & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, 27(1), 57–78.
- Hiebert, J., Gallimore, R., Garnier, H., Bogard Givvin, K., Hollingsworth, H., Jacobs, J., Miu-Ying Chui, A., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries. Results from the TIMSS 1999 Video Study*. U.S. Department of Education, National Center for Education Statistics. Washington DC.
- Hiebert, J. & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and Procedural Knowledge: The Case of Mathematics* (pp. 1–27). Hillsdale, NJ, USA: Lawrence Erlbaum.
- Hoch, M. & Dreyfus, T. (2004). Structure sense in high school algebra: The effect of brackets. In M. J. Høines & A. B. Fuglestad (Eds.), *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education, Vol. 3* (pp. 49–56).
- Hoch, M. & Dreyfus, T. (2006). Structure sense versus manipulation skills: An unexpected result. In J. Novotná, H. Moraová, M. Krátká, & N. Stehlíková (Eds.), *Proceedings 30th Conference of the International Group for the Psychology of Mathematics Education, Prague, Vol. 3* (pp. 305–312).

- Hoffman, B. & Schraw, G. (2010). Conceptions of efficiency: Applications in learning and problem solving. *Educational Psychologist*, 45(1), 1–14.
- Janssen, J., Van der Schoot, F., & Hemker, B. (2005). *PPON, Balans [32] van het reken-wiskundeonderwijs aan het einde van de basisschool 4. Uitkomsten van de vierde peiling in 2004*. Uitgave Stichting Cito Instituut voor Toetsontwikkeling.
- Janvier, C. (1996). Modeling and the initiation into algebra. In N. Bednarz et al. (Ed.), *Approaches to Algebra* (pp. 225–236). Dordrecht: Kluwer Academic Publishers.
- Kester, L., Kirschner, P., & Van Merriënboer, J. (2006). Just-in-time information presentation: Improving learning a troubleshooting skill. *Contemporary Educational Psychology*, 31, 167–185.
- Kilpatrick, J., Swafford, J., & Findell (Eds.), B. (2001). *Adding It Up*. Washington, D.C.: National Academy Press.
- Kirshner, D. & Awtry, T. (2004). Visual salience of algebraic transformations. *Journal for Research in Mathematics Education*, 35(4), 224–257.
- Kraker, B. & Sauren, F. (2004). *Eerstejaars Studenten over VWO → WO. Een Inventarisatie van Meningen van Eerstejaarsstudenten Werktuigbouwkunde en Biomedische Technologie van de Technische Universiteit Eindhoven*. Technical report, Technische Universiteit Eindhoven.
- Linacre, J. M. (2009). *A User's Guide to Winsteps*. Free to achieve from <http://www.winsteps.com/winsteps.htm>.
- Linchevski, L. & Herscovics, N. (1996). Crossing the cognitive gap between arithmetic and algebra: Operating on the unknown in the context of equations. *Educational Studies in Mathematics*, 30(1), 39–65.
- Linchevski, L. & Livneh, D. (1999). Structure sense: The relationship between algebraic and numerical contexts. *Educational Studies in Mathematics*, 40(2), 173–196.
- Matz, M. (1982). Towards a process model for high school algebra errors. In D. Sleeman & J. Brown (Eds.), *Intelligent Tutoring Systems* (pp. 25–50). Academic Press.
- McCallum, W. (2010). Restoring and balancing. In Z. Usiskin, K. Andersen, & N. Zotto (Eds.), *Future curricular trends in school algebra and geometry: Proceedings of a conference* (pp. 277–286). Charlotte, Information Age Publishing Inc.
- Miller, G. (1956). The magical number seven, plus or minus two; some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.

- Ministerie van Onderwijs, Cultuur en Wetenschap (2010). *Kerncijfers 2005-2009*. Technical report, Ministerie van OCW Zoetermeer.
- National Mathematics Advisory Panel (2008). *Foundations for Success: The Final Report of the National Mathematics Advisory Panel*. Technical report, Washington, U.S. Department of Education.
- Novotná, J. & Hoch, M. (2008). How structure sense for algebraic expressions or equations is related to structure sense for abstract algebra. *Mathematics Education Research Journal*, 20(2), 93–104.
- OECD (2003). *PISA 2003 Technical Report*. Technical report, Organisation for Economic Co-operation and Development.
- OECD (2007). *PISA 2006: Science Competencies for Tomorrow's World Executive Summary*. Technical report, Organisation for Economic Co-operation and Development.
- Olson, J. F., Martin, M. O., & I.V.Mullis, Eds. (2008). *TIMSS 2007 Technical Report*. TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Owen, E. & Sweller, J. (1989). Should problem solving be used as a learning device in mathematics. *Journal for Research in Mathematics Education*, 20(3), 322–328.
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429–434.
- Paas, F. (2007). *Neem eens een voorbeeld aan anderen*. *Nieuwe impulsen voor onderzoek naar leren en instructie*. Inaugural lecture, Open Universiteit Nederland.
- Paas, F., Renkl, A., & Sweller, J. (2003a). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, 38(1), 1–4.
- Paas, F., Tuovinen, J., Tabbers, H., & Van Gerven, P. (2003b). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist*, 38(1), 63–71.
- Paas, F., Van Gerven, P., & Wouters, P. (2007). Instructional efficiency of animation: Effects of interactivity through mental reconstruction of static key frames. *Applied Cognitive Psychology*, 21(6), 783–793.
- Paas, F. & Van Merriënboer, J. (1993). The efficiency of instructional conditions: An approach to combine mental effort and performance measures. *Human Factors*, 35(4), 737–743.

- Paas, F. & Van Merriënboer, J. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review*, 6(4), 351–371.
- Paas, F. & Van Merriënboer, J. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive load approach. *Journal of Educational Psychology*, 86(1), 122–133.
- Piaget, J. (1970). *Genetic Epistemology*. New York: Columbia University Press.
- Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*. University of Chicago Press.
- Reichard, L., Rozemond, S., Dijkhuis, J., Te Vaarwerk, G., Verbeek, J., De Jong, G., Brokamp, N., Houwing, H., De Vroome, R., Kuis, J., Ten Klooster, F., Van Leeuwen, F., De Waal, S., & Van Braak, J. (2005). *Getal en Ruimte 2 vwo 1, eerste druk, vierde oplage*. Houten: EPN.
- Reichard, L., Rozemond, S., Dijkhuis, J., Te Vaarwerk, G., Verbeek, J., De Jong, G., Brokamp, N., Houwing, H., De Vroome, R., Kuis, J., Ten Klooster, F., Van Leeuwen, F., De Waal, S., & Van Braak, J. (2006). *Getal en Ruimte 1 vwo 1, eerste druk, zesde oplage*. Houten: EPN.
- Reichard, L., Rozemond, S., Dijkhuis, J., Te Vaarwerk, G., Verbeek, J., De Jong, G., Brokamp, N., Houwing, H., De Vroome, R., Kuis, J., Ten Klooster, F., Van Leeuwen, F., De Waal, S., & Van Braak, J. (2007). *Getal en Ruimte vwo B deel 1, eerste druk, vijfde oplage*. Houten: EPN.
- Reichard, L., Rozemond, S., Dijkhuis, J., Te Vaarwerk, G., Verbeek, J., De Jong, G., Brokamp, N., Houwing, H., De Vroome, R., Kuis, J., Ten Klooster, F., Van Leeuwen, F., De Waal, S., & Van Braak, J. (2010). *Getal en Ruimte 3 vwo 1, eerste druk, eerste oplage*. Houten: EPN.
- Rosnick, P. (1981). Some misconceptions concerning the concept of variable. *Mathematics Teacher*, 74(6), 418–420.
- Schnotz, W. & Kürschner, C. (2007). A reconsideration of cognitive load theory. *Educational Psychological Review*, 19(4), 469–508.
- Schoenfeld, A. & Arcavi, A. (1988). On the meaning of variable. *Mathematics Teacher*, 81(6), 420–427.
- Schutte, M. (2010). *Inschatten van het algebranimeau in het voortgezet onderwijs door docenten*. Master's thesis, Eindhoven School of Education, Eindhoven University of Technology, Retrieved 25 May 2011 from <http://alexandria.tue.nl/extra2/afstversl/esoe/690053.pdf>.

- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational Studies in Mathematics*, 22(1), 1–36.
- Sfard, A. (1995). The development of algebra: confronting historical and psychological perspectives. *Journal of Mathematical Behavior*, 14(1), 15–39.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4–13.
- Sfard, A. & Linchevski, L. (1994a). Between arithmetic and algebra: In the search of a missing link, the case of equations and inequalities. *Rendiconti Del Seminario Matematico*, 52(3), 279–307.
- Sfard, A. & Linchevski, L. (1994b). The gains and pitfalls of reification: The case of algebra. *Educational Studies in Mathematics*, 26(2-3), 191–228.
- Skemp, R. R. (1976). Relational understanding and instrumental understanding. *Mathematics Teaching*, 77, 1–7.
- SLO (2008). *ReAL Leerlijnen 'van rekenen naar algebra'*. *Wiskunde in de onderbouw van het havo/vwo*. Technical report, Enschede: SLO/Fi.
- Stacey, K., Chick, H., & Kendal, M., Eds. (2004). *The future of the teaching and learning of algebra: The 12th ICMI study*. Dordrecht: Kluwer.
- Sterk, H. & Perrenet, J. (2005). Kunnen (wij op) onze kinderen rekenen? Gewenste rekenvaardigheden voor aanstaande bètawetenschappers. *Euclides*, 81(2), 63–65.
- Sweller, J. (1989). Cognitive technology: Some procedures for facilitating learning and problem solving in mathematics and science. *Journal of Educational Psychology*, 81(4), 457–466.
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285.
- Sweller, J. & Low, R. (1992). Some cognitive factors relevant to mathematics instruction. *Mathematics Education Research Journal*, 4(1), 83–94.
- Sweller, J., Van Merriënboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251–296.
- Tall, D. & Thomas, M. (1991). Encouraging versatile thinking in algebra using the computer. *Educational Studies in Mathematics*, 22(2), 125–147.
- Tempelaar, D. (2007). Onderwijzen of bijspijkeren? *Nieuw Archief voor Wiskunde*, 5/8(1), 55–59.

- Trigueros, M. & Ursini, S. (1999). Does the understanding of variable evolve through schooling? In O. Zaslavsky (Ed.), *Proceedings of the 23rd Conference of the International Group for the Psychology of Mathematics Education*, volume 4 (pp. 273–280). Haifa, Israel: Technion.
- Tuovinen, J. E. & Sweller, J. (1999). A comparison of cognitive load associated with discovery learning and worked examples. *Journal of Educational Psychology*, 91(2), 334–341.
- Tweede Fase Adviespunt (2005a). *Met een Tweede Fase-diploma naar het Hoger Onderwijs. Deelonderzoek: Ervaringen en Opvattingen van Opleiders in het Hoger Onderwijs*. Technical report, Tweede Fase Adviespunt.
- Tweede Fase Adviespunt (2005b). *Zeven Jaar Tweede Fase, een Balans*. Technical report, Tweede Fase Adviespunt.
- Ursini, S. & Trigueros, M. (2001). A model for the use of variable in elementary algebra. In M. van den Heuvel-Panhuizen (Ed.), *Proceedings of the 25th Conference of the International Group for the Psychology of Mathematics Education*, volume 4 (pp. 327–334). Utrecht: Freudental Institute.
- Van Gastel, L., Cuypers, H., Grift, Y., Kaper, W., Van der Kooij, H., Tempelaar, D., & Van de Vrie, E. (2010). *Aansluitmonitor Wiskunde VO-HO*. Consortium NKBW p/a Universiteit van Amsterdam, Amstel Instituut.
- Van Gastel, L., Cuypers, H., Jonker, V., Van de Vrie, E., & Van der Zanden, P. (2007). *Eindrapport Nationale Kennisbank Basisvaardigheden Wiskunde*. Technical report, Consortium NKBW, Universiteit van Amsterdam.
- Van Gog, T. & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, 43(1), 16–26.
- Van Gog, T., Paas, F., & Van Merriënboer, J. (2008). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction*, 18(3), 211–222.
- Van Hiele, P. (1973). *Begrip en inzicht, werkboek van de wiskundedidactiek*. [Meaning and insight, workbook of didactics of mathematics]. Purmerend: Muusses.
- Van Hiele, P. (1986). *Structure and Insight, A Theory of Mathematics Education*. Orlando: Academic Press.
- Van Merriënboer, J. & Sweller, J. (2005). Cognitive load theory and complex learning: Recent developments and future directions. *Educational Psychology Review*, 17(2), 147–177.

- Vlassis, J. (2002). The balance model: hindrance or support for the solving of linear equations with one unknown. *Educational Studies in Mathematics*, 49(3), 341–359.
- Vos, P. (2007). Algebra-prestaties van tweedeklassers. Zijn ze voor- of achteruit gegaan? *Euclides*, 82(4), 129–132.
- Wagner, S. (1983). What are these things called variables? *Mathematics Teacher*, 76, 474–479.
- Wenger, R. (1987). Cognitive science and algebra learning. In A. Schoenfeld (Ed.), *Cognitive science and mathematics education* (pp. 217–251). Hillsdale: Lawrence Erlbaum.
- Whitehead, A. (1924). *An introduction to mathemamtics*. London: Williams and Norgate.
- Yackel, E. & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education*, 27(4), 458–477.
- Zwaneveld, B. (2004). Algebra: verloren zaak of uitdaging? *Euclides*, 80(2), 42–47.

# Appendix A: Test items

Tasks with an A in the measure are tasks that have served as anchor items in the Rasch analysis. Tasks are arranged by increasing measure. The Rasch measure of a task correspond to a probability of 0.50 to answer that task correctly.

Measure (logits)	Task	Assessment
-3.46 A	Expand the brackets: $-4(3a + b) =$	March 2008
-3.46 A	Expand the brackets: $-5(2p + q) =$	May 2008
-3.46 A	Expand the brackets: $-3(4p + q) =$	October 2008
-3.46 A	Expand the brackets: $-4(5p + q) =$	February 2009
-2.42 A	Simplify: $2\sqrt{5} + 4\sqrt{5} =$	March 2008
-2.42 A	Simplify: $3\sqrt{7} + 2\sqrt{7} =$	May 2008
-2.42 A	Simplify: $4\sqrt{11} + 3\sqrt{11} =$	October 2008
-2.42 A	Simplify: $2\sqrt{3} + 5\sqrt{3} =$	February 2009
-1.92	Calculate: $\frac{18-5}{7-20} + 4 =$	May 2008
-1.83	Calculate: $3 \cdot (-2) \cdot 5 - 2 \cdot 5 =$	March 2008
-1.59	Calculate: $2 \cdot (-3) \cdot 5 - 2 \cdot 5 =$	February 2009
-1.47 A	Simplify: $-2(3x - y) + 3(-4y - 2) =$	March 2008
-1.47 A	Simplify: $-3(2a - b) + 4(-2b - 3) =$	May 2008
-1.47 A	Simplify: $-4(2x - y) + 2(-3y - 4) =$	October 2008
-1.47 A	Simplify: $-2(4x - y) + 3(-2y - 4) =$	February 2009
-1.24	You know the operations plus, minus, multiplication and division. We introduce a new operation, diamond, and $a \diamond b$ is defined as follows. For two numbers $a$ and $b$ , we say $a \diamond b = a^2 - a \cdot b$ . Does $a \diamond b = b \diamond a$ hold for all numbers $a$ and $b$ ?	October 2008
-1.13	Calculate $(4\frac{1}{2})^2$ . Show your working.	May 2008
-0.88	Simplify: $\frac{4}{15} - \frac{2}{5} + \frac{3}{10} =$	February 2009
-0.85 A	Simplify $(3a^2 + 2a + 7)(a + 8)$ . Show your working.	March 2008
-0.85 A	Simplify $(2x^2 + 3x + 4)(x + 6)$ . Show your working.	May 2008
-0.85 A	Simplify $(2a^2 + 4a + 3)(a + 5)$ . Show your working.	October 2008
-0.85 A	Simplify $(3a^2 + 4a + 5)(a + 7)$ . Show your working.	February 2009



Measure (logits)	Task	Assessment
-0.69	Simplify: $\frac{7}{15} - \frac{2}{5} + \frac{1}{10} =$	May 2008
-0.68	By which number do you have to divide 4 to obtain 12?	March 2008
-0.62	Fleur claims that $3\frac{1}{3} \times 2\frac{1}{2} = 6\frac{1}{6}$ . Explain why this is incorrect.	February 2009
-0.60	Jantine claims that $2\frac{1}{2} \times 3\frac{1}{3} = 6\frac{1}{6}$ . Explain why this is incorrect.	March 2008
-0.60	Solve: $(x - 1)(x + 3)(x - 4) = 0$ .	February 2009
-0.58	Willem claims that $2\frac{1}{2} \cdot 4\frac{1}{4} = 8\frac{1}{8}$ . Explain why this is incorrect.	October 2008
-0.50	Simplify: $\frac{7x^2+7-3(2x^2+2)}{x^2+1} =$	February 2009
-0.48	Solve: $(x - 5)(x - 7) = 3$ .	March 2008
-0.31	A classmate asks for your help in solving $\frac{15}{4+\frac{6}{1+x}} = 3$ . He does not know how to start. Describe what you would do to help your classmate.	March 2008
-0.22	Simplify: $\frac{5x^2+10-2(2x^2+4)}{x^2+2} =$	May 2008
-0.04	The sum of two numbers is 93 and their difference is 21. Show how you can calculate these numbers.	October 2008
0.01	Solve: $\frac{21}{6+\frac{5}{1+x}} = 3$ .	October 2008
0.07	Martijn claims that $Q = \sqrt{P-2}$ implies $P = Q^2 + 2$ . Explain why you do or do not agree with Martijn.	May 2008
0.15 A	Calculate: $-7 - (4 - 3) \cdot (-8) - 2 =$	March 2008
0.15 A	Calculate: $-6 - (5 - 4) \cdot (-8) - 3 =$	May 2008
0.15 A	Calculate: $-5 - (3 - 2) \cdot (-8) - 4 =$	October 2008
0.15 A	Calculate: $-7 - (5 - 4) \cdot (-8) - 4 =$	February 2009
0.33	Solve: $x^2 - 5x + 6\frac{1}{4} = 0$	March 2008
0.33	Simplify: $\frac{1}{14} + \frac{4}{7} - \frac{10}{21} =$	October 2008
0.37	Simplify: $\frac{\sqrt{28}}{\sqrt{7}} =$	March 2008
0.38	Willem claims that $P = \frac{1}{Q} + 5$ implies $Q = \frac{1}{P} - 5$ . Explain why you do or do not agree with Willem.	February 2009
0.42	Simplify: $\frac{8}{21} - \frac{2}{7} + \frac{1}{14}$	March 2008
0.68	Solve: $(x - 5)(x + 2)(x - 3) = 0$ .	May 2008
0.72	Substitute $a = -1$ and $b = -2$ in $-(ab^2)^3 - 2(a^2b)^2$ .	October 2008
0.91	Substitute $a = -2$ and $b = -1$ in $-(a^2b)^3 - 2(ab^2)^2$ .	March 2008

Measure (logits)	Task	Assessment
1.00	Martijn claims that $\sqrt{a^2} = a$ holds for all numbers $a$ . Explain why you do or do not agree with Martijn.	March 2008
1.19	Simplify: $\frac{\sqrt{12}}{\sqrt{3}} =$	February 2009
1.34	Is there any $x$ for which $\frac{2x+1}{4x+2} = 2$ ? If so, calculate $x$ ; if not, explain why such an $x$ does not exist.	February 2009
1.37	Is there any $x$ for which $\frac{2x+3}{4x+6} = 2$ ? If so, calculate $x$ ; if not, explain why such an $x$ does not exist.	March 2008
1.43	Rewrite the formula $P = \frac{1}{Q} + 5$ as a formula of the form $Q = \dots$ something with $P \dots$	March 2008
1.55	Martijn claims that $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ . Explain why you do or do not agree with Martijn?	May 2008
2.01	Solve: $2(3x + 2) = 3(2x - 1) + 7$ .	October 2008
3.94	Solve: $a\sqrt{2} = 1 + 2a\sqrt{3}$ . The square roots may remain.	October 2008
6.13	If $a\sqrt{b} = 1 + 2a\sqrt{1+b}$ , then $a =$	March 2008



## Appendix B: Structure sense test items

Tasks with an A behind the Rasch measure have served as anchor items. The Rasch measure of a task correspond to a probability of 0.50 to answer that task correctly.

Measure (logits)	Task	Assessment
-0.31	A1 A classmate asks for your help in solving $\frac{15}{4+\frac{6}{1+x}} = 3$ . He does not know how to start. Describe what you would do to help your classmate.	March 2008
0.01	A2 Solve: $\frac{21}{6+\frac{5}{1+x}} = 3$ .	October 2008
-0.22	A3 Simplify: $\frac{5x^2+10-2(2x^2+4)}{x^2+2} =$	May 2008
-0.50	A4 Simplify: $\frac{7x^2+7-3(2x^2+2)}{x^2+1} =$	February 2009
0.68	A5 Solve: $(x-5)(x+2)(x-3) = 0$ .	May 2008
-0.60	A6 Solve: $(x-1)(x+3)(x-4) = 0$ .	February 2009
1.37	A7 Is there any $x$ for which $\frac{2x+3}{4x+6} = 2$ ? If so, calculate $x$ ; if not, explain why such an $x$ does not exist.	March 2008
1.34	A8 Is there any $x$ for which $\frac{2x+1}{4x+2} = 2$ ? If so, calculate $x$ ; if not, explain why such an $x$ does not exist.	February 2009
3.94	A9 Solve: $a\sqrt{2} = 1+2a\sqrt{3}$ . The square roots may remain.	October 2008
6.13	A10 If $a\sqrt{b} = 1 + 2a\sqrt{1+b}$ , then $a =$	March 2008

Measure (logits)	Task	Assessment
-1.83	N1 Calculate: $3 \cdot (-2) \cdot 5 - 2 \cdot 5 =$	March 2008
-1.59	N2 Calculate: $2 \cdot (-3) \cdot 5 - 2 \cdot 5 =$	February 2009
0.15 A	N3 Calculate: $-7 - (4 - 3) \cdot (-8) - 2 =$	March 2008
0.15 A	N4 Calculate: $-6 - (5 - 4) \cdot (-8) - 3 =$	May 2008
0.15 A	N5 Calculate: $-5 - (3 - 2) \cdot (-8) - 4 =$	October 2008
0.15 A	N6 Calculate: $-7 - (5 - 4) \cdot (-8) - 4 =$	February 2009

# Summary

## The development of algebraic proficiency

Discussions of students' level of algebraic proficiency are taking place worldwide. In the Netherlands, the discussion has focused on the level of basic algebraic skills in the transition from secondary education to higher education. This discussion has served as input for the study described in this thesis. This study has aimed at obtaining insight in the actual level of student algebraic proficiency in pre-university education. A second aim is to obtain insight into the deeper causes of the problems students experience in pre-university education in case the level turned out to be disappointing.

In the discussion on the student level of algebraic proficiency, several considerations played a role. The problems indicated by educators of secondary and higher education concerned the skills taught in lower secondary education and even primary education. Furthermore, a call has been heard from higher education for skills exceeding procedural skills. The subsequent discussion on the student level of algebraic proficiency readily has shifted to a debate concerning the proposed solutions and the classic dichotomy procedural fluency–conceptual understanding. These two aspects served as input in a study of the development of algebraic proficiency described in Chapter 2. In order to obtain a detailed view of the student level of algebraic proficiency, we constructed four tests for students of grades 8–12. As theoretical background, we used two aspects of algebraic proficiency.

The first aspect concerned the relation between basic skills and the deeper understanding that also played a role in the aforementioned discussion. There is general agreement that basic skills have to go hand in hand with conceptual aspects of algebraic proficiency, although different researchers provide different views of the nature of these conceptual aspects. We used symbol sense (Arcavi, 1994) which is described as a complex feel for symbols that transcend basic manipulations. It includes the ability to manipulate or to read through symbolic expressions depending on the problem at hand, and flexible manipulation skills, such as the ability to cleverly select and use a particular symbolic representation. In the construction of the tests, we followed Drijvers (2006a, 2010) in viewing

algebraic proficiency as a sliding scale ranging from basic skills to symbol sense. In the tests we aimed at including items that cover this full range. The second aspect concerned the transition from arithmetic to algebra, which is seen as crucial for the development of algebraic proficiency. We decided to include topics such as calculating with negative numbers, fractions, and square roots, which are related to the transition from arithmetic to algebra.

Based on these two aspects, we constructed four sequential paper and pencil tests in a partly cross-sectional and partly longitudinal design. Around one thousand students joined at least one of the sessions of the assessments; 277 students participated in all four assessment. The written answers were used to investigate how student algebraic proficiency developed from both a cross-sectional and a longitudinal perspective, and in terms of basic skills and symbol sense. For the analysis of our data, we used the Rasch model (Rasch, 1980; Bond & Fox, 2007).

The results of the analysis showed that the students made progress, both cross-sectionally and longitudinally. In the cross-sectional development of the students, the positive effect might in part be attributed to a curriculum change, as the new curriculum pays considerably more attention to algebraic skills in the lower grades. The longitudinal analysis showed that for the majority of the students, the level of algebraic proficiency increased after one year. However, for the majority of the students, this progression was not statistically significant. With respect to different streams, we found that the students from the science stream performed better than the students from the social stream, but again the difference was not statistically significant.

Relating the student development to the individual tasks in the assessment, we found that the students mastered simple tasks, but tasks became too complicated for them rather quickly. For example, the task “simplify:  $2\sqrt{5} + 4\sqrt{5}$ ” is mastered by the majority of the students, whereas the task “simplify:  $\frac{\sqrt{12}}{\sqrt{3}}$ ” is mastered by less than 10% of the students. In addition, the task “expand the brackets:  $-4(3a + b)$ ” is mastered by almost all students, whereas the task “simplify:  $(3a^2 + 2a + 7)(a + 8)$ ” is mastered by less than half of the students. As a result, few tasks were mastered by the majority of the students. This indicates problems in Dutch secondary education with the mastery of procedural skills. Furthermore, we saw that the range of tasks which the majority of the students mastered did not incorporate the tasks related to important conceptual aspects.

In Chapter 3 we used the notion of structure sense (Linchevski & Livneh, 1999; Novotná & Hoch, 2008) to get a better handle on the conceptual aspects that seem to be problematic. Structure sense in high school algebra is seen as consisting of several abilities including the ability to see the algebraic structure of a formula and the ability to see a part of an expression as a unit. We argued that structure sense can be seen as a part of symbol sense. From the assessments described above, we selected tasks that appeal to structure sense and focused the analysis on these tasks. We concluded that the majority of the students did

not master tasks that require structure sense.

In addition we saw that the relations between the students' strategies, their success on tasks that appeal to structure sense, and their positions on the Rasch scale, were not as clear as we expected on the basis of the literature. In our view, this means that either the relation between strategy and structure sense is not as strong as one would expect, or the relation between structure sense and the underlying latent variable of the Rasch scale is not straightforward. Too much attention to success on tasks may lead to a procedural approach towards conceptual notions. Further research on the relations between strategies, structure sense, and algebraic proficiency seems appropriate. In our view this touches on the complexity of the notion of structure sense: it is both difficult to acquire and not easy to measure.

Many researchers argue that flexible manipulation skills, which include structure sense, are an important aspect of algebraic proficiency. This flexibility is especially of importance for students in higher education, for mathematics as well as for associated disciplines such as science and economics. The results of the assessments suggest that Dutch mathematics education does not pay sufficient attention to these kinds of skills. We therefore call for more attention to structure sense in Dutch mathematics education.

In Chapter 4, we focused on the conceptual barriers that cause the difficulties which students experience by using Sfard's theory of reification (Sfard, 1991). In that theory, it is argued that abstract concepts can be conceived of in two fundamentally different ways: structurally as objects, and operationally as processes. Sfard bases this observation on an analysis of the historical development of algebra, which can be characterized as a long sequence of transitions from operational to structural conceptions. Again and again, processes performed on abstract objects at one level turn into new objects that serve as objects of new higher level processes. In the process of concept formation, three stages are distinguished: interiorization, condensation, and reification. Whereas interiorization and condensation are gradual processes, reaching the phase of reification is not gradual, but "a sudden ability to see something familiar in a totally new light" Sfard (1991, p. 19). Reification is the point where the object turns into input for interiorization on a higher level.

Sfard & Linchevski (1994b) identified two crucial transitions: the transition from the purely operational algebra to the structural algebra, and the transition from algebra of an unknown to the functional algebra of a variable. To Sfard & Linchevski (1994a), an important component of algebraic proficiency is a flexible approach in which a student is able to switch between an operational approach, in which the focus is on the processes, and a structural approach in which the focus is on objects. These switches play a role in the process-object duality and the fixed-value-variable duality. We used these transitions to try to explain the student performance on the assessments described in Chapter 2. Scrutinizing the test items from Sfard's perspective showed a remarkable difference in per-



formance of tasks concerning calculating with square roots. The task “simplify  $2\sqrt{5} + 4\sqrt{5}$ ” (with the correct answer  $6\sqrt{5}$ ) is mastered by almost 90% of the students, whereas the task “do you or do you not agree that  $\sqrt{12} + \sqrt{3} = 3\sqrt{3}$ ?” is mastered by less than 10% of the students of grade 12. Further, the task in which students had to solve the linear equation  $2(3x + 2) = 3(2x - 1) + 7$  (which actually is an identity) was mastered by less than 10% of the grade 12 students. Based on the student performance on these tasks, we hypothesized that the problems grow in complexity according to the number of transitions that have to be made between an operational and a structural approach (or the reverse).

Testing this hypothesis among 92 grade 11 students yielded a confirmation. In the case of square roots, the majority of the incorrect answers seemed to be due to a lack of the connection between the process of extracting the square root and a root as an object. As for the linear equations, the majority of students was able to expand the brackets correctly. Here, the problem was in drawing the correct conclusion. The confirmation of the hypothesis led us to conclude that the majority of Dutch students were not able to deal with the operational–structural duality in a flexible way. We argued that the process–object duality is inherently difficult for Dutch students and that Dutch mathematics education apparently does not pay enough attention to fostering the process of reification.

In Chapter 5, we discussed methodological issues concerning the measure ‘instructional efficiency’, used in the cognitive load theory. Because part of the discussion in the Netherlands on student algebraic proficiency concerned student fluency in solving algebraic tasks, we decided to measure the invested mental effort. Based on the principles of the cognitive load theory, students were asked to report their invested mental effort on a nine point symmetrical rating scale. We expected the measure of instructional efficiency, which combines mental effort and performance, to provide a more detailed view of the student level of algebraic proficiency than performance alone. Although this measure was originally developed to compare instructional designs, we followed the suggestion of Van Gog & Paas (2008) to use this measure to evaluate the student level of expertise in algebraic proficiency. Trying to use the measure in this way, however, we found that it brought with it three methodological limitations.

The first limitation concerned the assumption in the construction of instructional efficiency that higher mental effort is counterbalanced proportionally by higher performance and vice versa. The second limitation concerned ambiguous meanings of mental effort scores in situations of low performance and low mental effort. The third limitation concerned the validity of mental effort scores. Our data gave rise to questioning the validity of ‘mental effort’ since students seemed to mix mental effort and perceived difficulty. Based on these limitations, we concluded that instructional efficiency is not an appropriate measure for investigating the student level of algebraic expertise. Due to our concern with the validity of the mental effort scores, we eventually decided not to use these scores to determine the student level of expertise.

Because so a large number of students do not reach the area of important conceptual aspects, and because textbooks play an important role in Dutch mathematics education (Hiebert et al., 2003), we performed an analysis of two most widely used Dutch textbook series in Chapter 6. This analysis aimed at investigating which means of support Dutch textbooks offer to foster conceptual understanding. Dutch textbooks are influenced by the theory of Realistic Mathematics Education (RME). In line with RME, we used the theory of emergent modeling (Gravemeijer, 1999) as a framework to analyze the most used textbook series. From an RME perspective, students should not be confronted with ready-made models. Rather, these models should emerge from their own mathematizing activities (Gravemeijer, 1994), while supporting mathematical growth. In this way, the models emerge as *models for* mathematical reasoning. The transition from *model of* to *model for* concerns a shift in student thinking. This shift passes from thinking about the modeled contextual situation to thinking about mathematical relations.

Based on the theory of emergent modeling, we classified activities in four categories varying from activities in the context to formal mathematical reasoning. The data set was constructed from two Dutch textbook series, *Moderne Wiskunde* (MW) and *Getal & Ruimte* (GR) from which entire chapters and sections within chapters concerning linear relations and linear equations were selected. The analysis indicated that the textbooks pay relatively great attention to activities such as searching for patterns in given contextual situations, constructing word formulas (linear relations), and developing informal strategies such as the cover-up method and using the balance model (linear equations). In addition, we found that both textbook series hardly contain activities that guide students from the contexts and semi-formal mathematics to formal mathematics. Moreover, we found that the textbooks start with an RME approach, but later on shift to a track with a more traditional approach, in which new concepts are introduced as ready-made mathematics. The RME track mainly contains activities at the task setting and referential levels. We found few activities that support the switch towards formal mathematics. The other track does not build on the semi-formal mathematical understanding grounded in the contexts and models, thus missing the link to the student's prior knowledge. Because of these two disjunct tracks, we concluded that these two Dutch textbook series do not offer a consistent instructional sequence and therefore hamper the development of conceptual proficiency.

In Chapter 7 we summarized and reflected upon our findings. We argued that Dutch students appear to make little progress in algebraic proficiency during pre-university secondary education. We stressed that this is not just a matter of too little fluency with basic algebraic skills; it also reflects the fact that the students in our research hardly developed conceptual proficiency in areas such as symbol sense, structure sense, and the process-object duality. These conceptual aspects include flexible skills which are important in different sectors of

higher education. In our view, mathematics teachers should pay more attention to the structure and the coherence of mathematics. This could be arranged by stimulating classroom dialogues on the structure of formulas and on generalizations. Discussing mathematical structures and generalizations could contribute to establishing connections between different approaches.

We recommend teachers to pay attention to the ambiguous nature of mathematics by discussing different perspectives on mathematical entities. In addition, we advise teachers to foster reification by, for example, stimulating their students not to be satisfied with performing tasks only. Instead, students must be stimulated to look for the bigger picture of which the tasks and solution procedures are part. In this way, the focus is not on performing tasks, but on searching for the mathematical concepts that are supposed to be expressed in the tasks, and on the coherence between solution procedures and concepts.

We have argued that Dutch textbooks pay relatively too much attention to exploring activities within contexts, whereas in our view, these activities do not require the most attention. It is not the contextual situation itself that is difficult to grasp; it is the underlying mathematical structure of the situation. Given the present textbook series, teachers could help students to use the contexts and models in a process that eventually leads to the construction of formal mathematics. For example, by discussing contexts and models, and the role they play in the learning trajectory, teachers could help students to generalize and formalize. In this way, teachers help students to bridge the gap between the two now disjunct didactical tracks.

Textbook authors are advised to reconsider the way contexts and models are used in the present textbooks. More specifically, they are advised to reconsider the way students can build more formal mathematical knowledge based on context-rooted, semi-formal strategies. In addition, they should include activities that help students to develop conceptual proficiency, which we defined as a mixture of higher-order thinking skills in which three aspects are essential: the ability to recognize and make flexible use of the algebraic structure, the ability to deal with the ambiguous nature of mathematical concepts, and the ability to see the coherence between mathematical concepts.

# List of publications

## Conference proceedings

Van Stiphout, I.M., Perrenet, J., Gravemeijer, K. (2009). Algebra at risk in Dutch education: a pilot study. Short oral communication presented at PME 33. In Tzekaki, M., Kaldirmidou, M. & Sakonidis, H. (Eds.). *Proceedings of the 33rd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 1, p 483. Thessaloniki, Greece: PME.

Van Stiphout, I.M. (2009). *Structural conceptions and the cognitive load theory*. Paper presented at Joint Researcher Training: Berlin Summer School.

Van Stiphout, I.M., Perrenet, J.C., Gravemeijer, K. & Jochems, W. (2009). *Cognitive load theory and the development of algebraic skills*. Poster presented at 3rd International Cognitive Load Theory Conference.

Van Stiphout, I.M. (2008). *Cognitive load theory and the development of algebraic skills*. Poster presented at joint researcher training Velbert Langenberg, Essen.

Van Stiphout, I.M., & G. Bruin-Muurling. (2008). *Dimensies in opvattingen over wiskundeonderwijs*. Group discussion organized at the Onderwijs Research Dagen, Eindhoven.

Van Stiphout, I.M., Perrenet, J.C., Gravemeijer, K. & Jochems, W. (2008). *Cognitive load theory and the development of algebraic skills*. Paper presented at ICO-toogdag.

## Journal articles

Van Stiphout, I.M., & G. Bruin-Muurling. (2009). *Verslag van de Onderwijs Research Dagen*. *Nieuwe Wiskrant* 28 (3), 39-40.

**Oral presentations (without proceedings)**

Van Stiphout, I.M. & Bruin, G. (2010). *Doorlopende leerlijnen in het onderwijs: het wiskundeonderwijs als een paradigmatisch voorbeeld*. Praktijkonderzoeksdag Fontys Lerarenopleiding Tilburg.

Bruin, G. & Van Stiphout, I.M. (2010). *Wat & Hoe*. Voordracht op afscheidssymposium.

Van Stiphout, I.M. (2010). *Algebraïsch vaardig?* Gastcollege algebraïsche vaardigheden FLOT voor masterstudenten. Fontys Lerarenopleiding Tilburg.

Bruin, G. & Van Stiphout, I.M. (2010). *Een stevige basis?* Voordracht op NKBW nascholing.

Bokhove, C., Van Stiphout, I.M. & Bruin-Muurling, G. (2009). *Wiskunde op de drempel. Onderzoeken naar algebraïsche vaardigheden*. Voordracht op symposium Wiskunde op de drempel van NKBW.

Bruin, G. & Van Stiphout, I.M. (2009). *Aansluitingsproblematiek in de doorlopende leerlijn rekenen en algebraïsche vaardigheden*. Voordracht op wiskundecolloquium VU Amsterdam.

Van Stiphout, I.M. (2009). *Algebraïsche vaardigheden nader onderzocht*. Workshop voor de Nederlandse Vereniging van Wiskundeleraren, november 2009.

Van Stiphout, I.M. & Bruin, G. (2009). *Onderzoek van algebraïsche vaardigheden van eerstejaars op TU/e*. Lezing InterTU studiedag.

Van Stiphout, I.M. (2009). Gastcollege algebraïsche vaardigheden FLOT voor masterstudenten. Fontys Lerarenopleiding Tilburg.

Van Stiphout, I.M. & Bruin, G. (2009). *Resultaten en discussie van de scores van eerstejaars studenten op algebraïsche vaardigheden*. Presentatie colloquium informatica TU/e.

Bruin, G. & Van Stiphout, I.M. (2009). *Is er een nieuwe kijk op de aansluitingsproblematiek nodig? Eerste resultaten breuken- en algebraïsche vaardigheden eerstejaars TU/e studenten*. Presentatie op aansluitingsbijeenkomst TU/e.

# Curriculum Vitae

Irene van Stiphout was born on 9th August 1969 in Breda in the Netherlands. After finishing general education (h.a.v.o.) in 1986 at Newman College in Breda, she studied one year secondary vocational laboratory education (M.L.O.) at the Dr. Struycken Institute in Etten-Leur. In 1987, she completed the foundational year at the Moller teacher training institute in Tilburg. Subsequently, she studied mathematics at the Catholic University of Nijmegen. In 1998 she graduated in algebraic number theory on a thesis entitled “Indices in cyclic cubic fields”.

In the period 1998–1999 she worked as an actuary trainee at Heijnis en Koelman in Eindhoven. From 2000 till 2007, she worked as a mathematics teacher at several secondary schools, the last five years at the Merewade College in Gorinchem. Meanwhile, she studied for her teaching degree which she obtained in 2002 at the teacher training institute at Eindhoven University of Technology. In January 2007 she started a PhD project at the Eindhoven School of Education, a collaborative institute of the TU/e and Fontys University of Applied Sciences.

In January 2012 she will teach mathematics at Ds. Pierson College in Den Bosch in the Netherlands.



# ESoE dissertation series

Van de Sande, R.A.W. (2007). *Competentiegerichtheid en scheikunde leren: Over metacognitieve opvattingen, leerresultaten en leeractiviteiten.*

Rajuan, M. (2008). *Student teachers' perceptions of learning to teach as a basis for the supervision of the mentoring relationship.*

Hooreman, R.W. (2008). *Synchronous coaching of trainee teachers: An experimental approach.*

Raessens, B.A.M. (2009). *De E-kubus: Een analysemodel voor curricula.*

Rohaan, E.J. (2009). *Testing teacher knowledge for technology teaching in primary schools.*

Oemar Said, E.W.R. (2009). *De Da Vinci case: Een onderzoek naar de relaties tussen vernieuwende leeromgevingen en de motivatie en de regulatievoorkeuren van leerlingen in het MBO.*

Koopman, M. (2010). *Students' goal orientations, information processing strategies and knowledge development in competence-based pre-vocational secondary education.*

Mittendorf, K.M. (2010). *Career conversations in senior secondary vocational education.*

Crasborn, F.J.A.J., Hennissen, P.P.M. (2010). *The skilled mentor. Mentor teachers' use and acquisition of supervisory skills.*

Van Bragt, C.A.C. (2010). *Students' educational careers in Higher Education: a search into key factors regarding study outcome.*

De Bakker, G.M. (2010). *Allocated online reciprocal peer support via instant*



*messaging as a candidate for decreasing the tutoring load of teachers.*

Vos, M.A.J. (2010). *Interaction between teachers and teaching materials. On the implementation of context-based chemistry education.*

Bruin-Muurling, G. (2010). *The development of proficiency in the fraction domain; affordances and constraints in the curriculum.*

Cornelissen, L.J.F. (2011). *Knowledge processes in School-University research networks.*

Van Stiphout, I.M. (2011). *The development of algebraic proficiency.*