On Mathematical Reasoning
- being told or finding out

Mathias Norqvist

Department of Mathematics and Mathematical Statistics
Umeå 2016

ISBN: 978-91-7601-525-4
ISSN: 1102-8300
Electronic version available at http://umu.diva-portal.org/
Printed by: Print & Media, Umeå University
Umeå, Sverige 2016

I love deadlines.
I like the whooshing sound
they make as they fly by.
- Douglas Adams

Table of Contents

Table of Contents **i**
Abstract **iii**
Sammanfattning **v**
Acknowledgements **vii**
List of papers **ix**
1 Introduction **1**
1.1 Aim 2
2 Mathematics in school **3**
2.1 Syllabi 3
2.2 Rote learning 3
2.2.1 Textbooks 4
2.2.2 Teaching 5
3 Theoretical frameworks **7**
3.1 Theory of didactical situations in mathematics 7
3.2 Productive struggle 8
3.3 Problem solving 10
3.4 A framework for mathematical reasoning 12
3.4.1 Reasoning sequences 12
3.4.2 Creative Mathematically Founded Reasoning 13
3.4.3 Algorithmic Reasoning 14
3.5 Cognitive demand 16
4 Memory and cognition **17**
4.1 Individual variation in cognition 17
4.1.1 Working memory 17
4.1.2 Fluid intelligence 18
4.1.3 Cognitive tests 18
4.2 Observing brain activity 19
4.2.1 The technique behind an fMRI image 20
4.2.2 A downside to fMRI 21
4.2.3 fMRI and Mathematics 21
4.2.4 Brain functions connected to mathematics 22
4.2.4.1 Memory retrieval 22
4.2.4.2 Novel and complex tasks 23
4.2.4.3 Calculation and number sense 23
4.3 Summary 24
5 Method **25**
5.1 Task design 25
5.1.1 Design by researchers 25
5.1.2 Authentic, design by teachers 29
5.2 Data collection 29
5.2.1 Behavioral-experiments 29

5.2.2 fMRI-experiment *30*
5.2.3 Observations *31*
5.3 Methods of analyses 31
5.3.1 Statistical *32*
5.3.2 Qualitative analysis *32*
6 Summary of the articles/Result **34**
6.1 Study 1 – Learning mathematics through algorithmic and creative reasoning 34
6.2 Study 2 – The affect of explanations on mathematical reasoning tasks 35
6.3 Study 3 – Learning mathematics without a suggested method: Durable effects on performance and brain activity 36
6.4 Study 4 – Unraveling students' reasoning: analyzing small-group discussions during task solving 37
6.5 Additional result 39
7 Discussion **41**
7.1 How will the task design influence students' solutions process, mathematical reasoning, and brain activity? 41
7.2 How will students' cognitive variation affect their solution rate and does task design matter? 43
7.3 How can these results influence teaching practice? 43
7.4 Limitations and implications for further studies 45
7.5 Conclusion 47
8 References **49**

Abstract

Background
Research has for many years pointed out the inefficiency of rote-learning and the importance of regarding concepts and mathematical properties while struggling with mathematics tasks (e.g., Hiebert, 2003; Schoenfeld, 1985; Stein, Grover, & Henningsen, 1996). From a theoretical viewpoint, Brousseau (1997) suggested that students have to consider such important aspects while constructing solutions by themselves and that teachers have to develop situations where this is possible for the students. The added effort needed from the students could however be cognitively demanding to the point that it will be overwhelming, in particular for cognitively less proficient students. Therefore, students' cognitive abilities are important to consider when constructing tasks or didactical situations. The aim of this thesis is therefore to examine how task design and students' cognitive abilities will influence students' mathematical reasoning, student outcome and students' brain activity.

Methods
Three of the four included studies are done with a between-groups design where data is analyzed statistically to search for significant differences in test results between the different practice conditions. In these studies, practice tasks were designed by researchers to promote special types of reasoning (algorithmic reasoning (AR) and creative mathematically founded reasoning (CMR)) and in one study an explanation on why the solution method is working was also provided. The practice data from these three studies are also analyzed as an additional result, not part of the included studies. The last study was based on observations of students work on tasks designed by teachers to unravel how student reasoning evolves during the solution process. Here we transcribed audio recordings from four student-groups' when they solved tasks constructed by their teacher. We then coded the solution process by utilizing the framework on mathematical reasoning suggested by Lithner (2008).

Results
The overall results suggest that creative tasks are more effective than algorithmic tasks when it comes to memory retrieval and reconstructing practiced solution methods. There are also clear indications that AR is more taxing on cognitive abilities during the test than creative tasks (where practice performance seems to be more important). During practice the dependence of cognitive abilities is however higher when working with creative tasks. Furthermore, task design is important for which type of reasoning that the student will use. However, student-group characteristics (i.e., motivation and

persistence) are also important both when choosing reasoning type and for task-progression.

Conclusion

Since mathematics students spend a lot of time doing tasks it is important to study these tasks from a learning perspective. The results in this thesis points to a few important issues regarding task design and the result of different types of reasoning. *First*, since creative reasoning seems to be more effective than algorithmic reasoning, it would be good for students to encounter more of this type of task in textbooks as well as in teacher presentations. *Second*, cognitive abilities are important for mathematics but there is a difference where the student's cognitive abilities are taxed (i.e., algorithmic reasoning will put higher strain on cognition during the test while creative reasoning will be highly demanding during practice). This difference in cognitive strain seems to be related to a deeper encoding during creative practice than during algorithmic practice. CMR also seems to be more beneficial than AR for cognitively less proficient students. While the teacher can reduce students' cognitive load by for instance directing focus to the important properties during practice, this may not be done during tests (at least not to the same extent). *Third*, even though algorithmic tasks do not prohibit the use of creative reasoning, it is much less likely to occur than algorithmic reasoning. To ensure that creative reasoning will take place, the task must be designed for this purpose.

Since creative tasks can put focus on one or more important mathematical properties and provide deeper understanding than algorithmic tasks, implementation in school practice can be essential if we want students to become mathematically literate.

Sammanfattning

Bakgrund

Forskning har under många år pekat på ineffektiviteten av utantill-lärande och vikten av att dels reflektera över koncept och matematiska egenskaper och dels bli tvungen att kämpa med matematikuppgifter (se t.ex. Hiebert, 2003; Schoenfeld, 1985; Stein, Grover, & Henningsen, 1996). Brousseau (1997) föreslog från ett teoretiskt perspektiv att elever måste överväga sådana viktiga aspekter medan de konstruerar egna lösningar och att lärare bör utforma situationer där detta är möjligt för eleverna. Den extra ansträngning som eleverna måste lägga ner kan dock vara kognitivt belastande till den grad att det blir överväldigande, speciellt för elever med lägre kognitiv kapacitet. Därför är elevers kognitiva kapacitet viktigt att beakta när man konstruerar uppgifter eller didaktiska situationer. Syftet med denna avhandling är därför att undersöka hur uppgiftsdesign och kognitiva färdigheter kan inverka på elevers matematiska resonemang, testresultat och hjärnaktivitet.

Metoder

Tre av de fyra inkluderade studierna är genomförda med en mellangrupps-design där datamaterialet analyserats statistiskt för att hitta signifikanta skillnader i testresultat mellan olika träningsgrupper. I dessa studier hade forskarna designat uppgifter som skulle främja olika typer av matematiska resonemang (algoritmiska resonemang (AR) och kreativa matematiskt grundade resonemang (CMR)). I en studie fanns även en förklaring tillgänglig som beskrev varför lösningsmetoden fungerade. Träningsdata från dessa tre studier har också analyserats som ett ytterligare resultat, utanför de inkluderade studierna. Den sista studien baseras på elevers arbete med uppgifter i klassrummet för att reda ut hur elevers resonemangssekvens utvecklas under lösningsprocessen. Vi transkriberade ljudinspelningar från fyra elevgrupper när de löste uppgifter som konstruerats av deras lärare. Sedan kodade vi lösningsprocessen genom att tillämpa Lithner's (2008) ramverk om matematiska resonemang.

Resultat

Det övergripande resultatet visar att kreativa uppgifter är effektivare än algoritmiska vad gäller minne och rekonstruktion av tränade lösnings-metoder. Det finns också klara indikationer på att AR ställer högre krav på kognitiva färdigheter under testen än vad CMR gör (där träningsresultatet är viktigare). Under träningen är däremot kravet på kognitiva färdigheter större när man arbetar med kreativa uppgifter. Vidare är uppgiftsdesignen viktig för vilken typ av resonemang studenterna kommer att använda. Studentgruppens egenskaper (d.v.s. motivation och envishet) är också viktigt, både för vilken resonemangstyp eleverna väljer samt för progressionen i uppgiften.

Slutsats

Eftersom matematikelever använder mycket tid till att lösa uppgifter är det viktigt att studera dessa uppgifter från ett inlärningsperspektiv. Resultatet i denna avhandling pekar ut några viktiga saker vad gäller uppgiftsdesign och resultat av olika typer av resonemang: 1) Eftersom CMR verkar vara effektivare än AR så vore det bra om eleverna mötte mer kreativa uppgifter i såväl läroböcker som i lärares presentationer. 2) Kognitiva färdigheter är viktiga för matematik, men det är skillnad när elevernas kognitiva färdigheter belastas, d.v.s. AR är mer belastande under testen medan CMR ger högre belastning under träningen. Denna skillnad i kognitiv belastning verkar bero på en djupare inkodning under den kreativa träningen än under algoritmisk träning. Dessutom verkar de lågpresterande eleverna dra mer nytta av CMR (jämfört med AR) än de högpresterande eleverna. Även om läraren kan reducera den kognitiva belastningen genom att exempelvis rikta fokus mot de viktiga egenskaperna under träning så kan läraren inte göra detta under ett test (åtminstone inte i samma utsträckning). 3) Även om algoritmiska uppgifter inte förhindrar CMR är det mindre sannolikt att detta skulle förekomma än AR. För att säkerställa att CMR ska ske måste uppgiften vara designad för det.

Eftersom kreativa uppgifter kan sätta fokus på en eller flera viktiga matematiska egenskaper, samt ge en djupare förståelse än algoritmiska uppgifter så är det nödvändigt att omsätta dem i skolpraktiken om vi vill att våra elever ska bli förtrogna med matematik.

Acknowledgements

Året var 1971. En ung kvinna klev in i en taxi för att åka till förlossningen, ca 130 km bort, för att få sitt första barn. Barnet i fråga beslutade att en så lång resa inte alls kom på fråga, så taxichauffören blev tvungen att svänga in vid den lilla sjukstugan i det lilla samhället i Västerbottens inland, där en liten pojke sedan föddes. Ungefär 39 år senare påbörjade den pojken, nu själv far till en son och med en till på väg, en intellektuell resa som han varken hade planerat eller packat för. Precis nu, cirka sex och ett halvt år senare, börjar den första delsträckan av denna akademiska resa närma sig sitt slut. När jag tänker tillbaka på mina år som doktorand finns det förstås en massa människor som hjälpt och stöttat mig på många sätt och jag ska, så gott jag förmår, försöka tacka er alla här.

Redan innan jag blev antagen som doktorand fick jag vara med i uppstarten av projektet *Att lära matematik med imitativa och kreativa resonemang* och fick umgås med två av mina blivande handledare. Det slog mig då att båda två tog vara på mina idéer även om jag var en nykomling. När jag sedan blev antagen blev handledarna tre och de har kompletterat varandra på ett utmärkt sätt. Jag har lärt mig massor av er och uppskattar verkligen att ni tagit er tid att följa, ledsaga och stötta denna inbitne slöfock så att avhandling och forskningsprojekt blivit av och blivit bättre. Johan Lithner, Bert Jonsson och Nina Rudälv, jag tackar er från djupet av mitt hjärta.

När forskarskolan startade mötte jag ett antal människor i samma läge som jag. Tillsammans har vi läst, funderat, analyserat och rest världen runt. Jag vill tacka er alla för ett trevligt resesällskap i såväl flyg och vulkanbuss som avhandlingsresan. Jag vill här även passa på att tacka den dåvarande ledningen i forskarskolan, P-O Erixon och Carina Rönnqvist, som stöttade alla oss vimsiga höns, nej doktorander, genom de första åren.

Jag blev även en del av Umeå Forskningscenter för Matematikdidaktik (UFM) och fick lära känna alla de trevliga och roliga människorna där. Hela UFM har varit oerhört värdefulla men jag vill rikta ett speciellt tack till Tomas Bergqvist som snabbt informerade mig om att jag absolut inte ska jämföra mig med Johan vad gäller effektivitet. Alla ni övriga i UFM ska veta att ni också gjort ett intryck på mitt liv och min forskning. Tack ska ni ha!

As I started working at the math-department I shared workroom with a woman that also had an impact on my view on research. Therefore, I would like to thank Gulden Karakok for her encouragements during my first year as a PhD-student and for her patience with my, at the time, quite rusty English.

När Gulden lämnade Umeå så flyttade ett blivande extrasyskon in i arbetsrummet. Vi har under många år delat både rum och målsättning och nu, så här på upploppet, så har vi båda slitit vårt hår över kappor och artiklar. I vårt rum har vi avhandlat allt från barn, månggifte, resfeber, träning, matlagning och undervisning, till avhandlingar och artiklar. Vi har både läst och undervisat kurser tillsammans och jag hoppas verkligen att vi kan fortsätta med det i många år framöver. Anneli Dyrvold, du har varit en klippa och jag är djupt tacksam för att jag fick göra denna resa med dig.

Jag tackar även 'fikagänget' och alla andra på institutionen som kommit med glada tillrop under min forskarutbildning. Det är ni som är skälet till att jag trivts så bra med mitt jobb de senaste 6 ½ åren. Jag vill dock rikta ett speciellt tack till Lisa Hed och Berit Bengtson som varit ett extra stort stöd under slutet av mitt avhandlingsarbete.

Hela projektgruppen förtjänar också ett stort tack. Förutom Johan och Bert vill jag speciellt tacka Carina Granberg, Janne Olsson och Tony Qwillbard för många nyttiga diskussioner och ett bra samarbete som jag också hoppas ska fortsätta. Jag vill även tacka Moa Eirell för hennes värdefulla hjälp med datainsamlingen till min egen studie och Hendrik van Stenbrugge som jag stångats med under den långa process som mynnade ut i avhandlingens sista artikel. Nu skulle man kunna tro att jag har glömt Yvonne Liljekvist, min sol i Karlstad som har varit en mycket god vän och hjälp under min doktorandtid, men så är inte fallet. Yvonne, du har varit en värdefull stöttepelare under dessa år och vi kommer säkert att fortsätta vårt samarbete när jag kliver på nästa delsträcka (vad det än må vara).

Den unga kvinnan som jag nämnde i början och den unge man som då var delvis ansvarig för situationen förtjänar givetvis ett enormt tack. Britt-Marie och Börje, mamma och pappa, utan er hade inte denna bok blivit skriven. Jag passar även på att skicka en liten hälsning till min syster Anna och hennes familj. Ni tillför glädje i mitt liv. Ett tack riktas också till grabbarna i EyeWood som varit mitt mentala andningshål under alla dessa år.

Slutligen vill jag tacka de som är viktigast i mitt liv, min älskade Maria och våra söner, Andreas och Viktor, som knappt träffat mig under denna intensiva sommar. Maria, jag kan inte på något sätt tacka dig nog för att du 1) drog med mig till Umeå och in på en lärarutbildning, 2) tyckte att det var ok att jag gick ner i lön samt förlorade min ferietjänst när jag blev doktorand och 3) tagit ansvar för och roddat allt under senvåren och sommaren medan jag lätt panikartat jobbat för att få denna bok i tryckfärdigt skick. Jag lovar att gottgöra dig på något sätt…om inte annat om ca. fem år när det är din tur. 😉 Du är underbar!😘

Om du som läsare tycker att jag har glömt bort dig så är det inte med mening. Jag skyller i så fall på tillfällig afasi och hoppas att även du tar del av min stora tacksamhet. Tack allihop!

Sävar, augusti 2016

I may not have gone where I intended to go,
but I think I have ended up where I needed to be.
Douglas Adams

List of papers

I. Jonsson, B., Norqvist, M., Liljekvist, Y., & Lithner, J. (2014). Learning mathematics through algorithmic and creative reasoning. *The Journal of Mathematical Behavior, 36*, 20-32. doi:10.1016/j.jmathb.2014.08.003

II. Norqvist, M. (2016). The affect of explanations on mathematical reasoning tasks. Manuscript in preparation.

III. Wirebring, L. K., Lithner, J., Jonsson, B., Liljekvist, Y., Norqvist, M., & Nyberg, L. (2015). Learning mathematics without a suggested solution method: Durable effects on performance and brain activity. *Trends in Neuroscience and Education, 4(1-2)*, 6-14. doi:10.1010/j.tine.2015.03.002

IV. Van Steenbrugge, H., & Norqvist, M. (2016). Unraveling students' reasoning: analyzing small-group discussions during task solving. Manuscript in preparation.

1 Introduction

Mathematics is one of the few school subjects that to its content does not differ that much across the world. The sum of two and three is always five[1] and if you differentiate the function $f(x) = x^2 \cdot \sin(x)$ you will end up with the derivative $f'(x) = 2x \cdot \sin(x) + x^2 \cdot \cos(x)$ regardless of which country you are educated in. There is however much difference regarding the way this uniform subject is taught. Teachers are more or less directed by syllabuses, textbooks are used to differing extent, and more or less responsibility is left to the student. What is common for all mathematics teaching is however the use of mathematics tasks to practice and hopefully learn the mathematics that will be used further on in both higher mathematics and everyday life. In this thesis the starting point will be the tasks and relating to them both design, learning outcome, and influence on student work and brain activity will be studied. As mathematics tasks are used to such a large extent, the impact and efficiency of these tasks are important to study and understand.

In this thesis I will begin with a short description of school-mathematics by discussing the content of the syllabi and comparing this to teachers' presentations and textbook content. This is done to set the stage for a discussion about task design and its influence on the mathematical reasoning that students choose or are directed to. Students use a lot of time practicing by doing tasks and their reasoning during this task solving process will affect their learning. Learning by rote is quite common in schools all over the world, and might, if dominating, be one problem when trying to learn mathematical theories or heuristics as it implies that things are learned without reflection. When discussing mathematics teaching and the classroom work, Brousseau's (1997) Theory of Didactical Situations in Mathematics offers insight in which roles students and teachers could have in the classroom to increase problem solving activities and encourage students' own knowledge construction. How this should be acted out in the classroom can of course be discussed extensively, but I will give an overview on some previous findings that will lead to the framework of mathematical reasoning by Lithner (2008), which is at the center of all the papers in the thesis.

The framework defines two major types of reasoning, imitative and creative. Imitative reasoning is closely related to rote learning while creative reasoning is based on students' own construction of solutions and therefore more connected to understanding relations and justifying choices on a more mathematically fundamental level. Both types will be discussed more extensively later on in this thesis.

When introducing tasks that are more cognitively demanding, individual variation in cognitive abilities can be crucial. There are many cognitive

[1] Provided that the calculations are made in a base ≥ 6 and modulo 5 or higher.

abilities that could influence mathematics learning but working memory and non-verbal problem solving ability are two that has been proven to be closely related to mathematical achievement (e.g., Primi, Ferrão, & Almeida, 2010; Swanson & Alloway, 2012). These two cognitive constructs are used to match participants into equally proficient groups before our experiments, and in the following analyses the cognitive measures will contribute to the results. The cognitive constructs will be discussed later on in the thesis and at this time functional magnetic resonance imaging (fMRI) will also be addressed. fMRI has been used to help make connections between active brain regions and cognitive processes and this is also what it is used for in one of the papers in this thesis. The connection between different cognitive processes and mathematical reasoning is important for the results of our experiments and the conclusions that can be drawn from them.

1.1 Aim

The overall aim for the thesis is to tie all these perspectives together and build on previous research to extend the knowledge on how task design, mathematical reasoning and cognitive abilities can affect the learning of mathematics. This will be done by comparing and combining the results from the four included papers to discuss the following questions.

1) How will the task design influence students' solution process, mathematical reasoning, and brain activity?
2) How will students' cognitive variation affect their solution rate and does the task design matter?
3) How could these results influence teaching practice?

2 Mathematics in school

Mathematics is a global school subject that has become one of the measures for school achievement, with tests like PISA and TIMSS. Newspapers have reported on decline or rise in results of these tests, both over time and between countries, and this has put some focus on mathematics education worldwide. In Sweden, decline in PISA results intensified the political debate on how education should be governed and executed. This debate included the syllabus for mathematics as well as for education in general both in primary and secondary school.

2.1 Syllabi

In Swedish syllabi, up until the early 90's, the aim of school mathematics was to prepare students for the every-day life with a focus on ability to perform necessary calculations (e.g., Skolöverstyrelsen, 1970, 1980). The last three syllabi, from 1994, 2000 and 2011, have changed this focus from pure calculation to include other abilities as well (e.g., Skolverket, 2000; Skolverket, 2011a, 2011b; Utbildningsdepartementet, 1994). Since 2000, seven competencies (i.e., procedures, reasoning, problem solving, modelling, communication, conceptual understanding, and relating mathematics to the surrounding world) have been explicitly defined in the syllabus. This change towards mathematical competence has been seen in other countries as well. The international movement towards a more competence focused curricula was in large initiated by the National Council of Teachers of Mathematics (NCTM, 1989, 2011). The Danish KOM-project defined eight mathematical competencies that education in mathematics should enhance (Niss & Jensen, 2002). The shift of focus from pure calculating skills to a broader mathematical competence could be important for mathematical proficiency in every-day life, as many occupations demand other competencies as well (e.g., problem-solving skills and modelling). However, these competencies are not mutually excluding. Calculating skills and rote learning of certain facts or rules are also important for an efficient problem solving process, since the focus can be put on the problem at large instead of each small item that need to be processed or calculated.

2.2 Rote learning

A lot of things in mathematics are memorized for quick and effortless retrieval and application when needed. For example, the multiplication table and the order of operations can be two important things to have quick access to. But this memorization can also become a hindrance if rote learning is dominating and if it occurs without understanding. Rote learning can be defined as a mechanical and habitual repetition of the learning object. If the student lacks understanding of the memorized rules and methods (i.e., why and how the rule works and is valid) it can be difficult when trying to get a more conceptual view on mathematics. Also, when rules become too numerous to keep track of there can also be difficulties if no connection to mathematical properties is made (e.g., learning integration only by rules is hard as sometimes many basic techniques have to be used simultaneously). When rote learned rules or facts become the main knowledge, students will not be able to solve tasks with the slightest variation from the ordinary (Hiebert, 2003). Hiebert (2003) went so far as to compare students that had learned mathematics mainly by rote to robots with poor memory, expressing that one could predict their errors just by erroneously recall rules or algorithms. This view is shared with Schoenfeld (1985) that states that the earlier focus on mechanical skills produced dismal results when students were challenged by more complex problems. Schoenfeld (1985) found that students tend to use only a small proportion of their total solution time on analyzing the problem. Instead they rush into a solution process without good strategies and towards a certain failure. Experts are more flexible in their problem solving. They use more time for analyzing the problem and are more prone to revising their choice when they get stuck (Schoenfeld, 1985). Boaler (1998) demonstrated in a longitudinal study that students that were taught mathematics with more emphasis on rote learning did not view mathematics as important in their daily life. These students also expressed that mathematics was boring, complicated and useless. Boaler's other group in this study, which were taught mathematics in a more project based way (i.e., with more emphasis on mathematical concepts and construction), expressed a more positive view on mathematics as a useful and important subject.

Although syllabi have shifted towards a more competence oriented aim, mathematics textbooks, as we will see in the next section, often stress algorithmic skills.

2.2.1 Textbooks

Textbooks are used throughout the world to provide students of all ages with tasks to solve and often also instruction on how this should be done. The high proportion of classroom time spent reading or solving tasks from textbooks might vary to some degree, but in most mathematics classrooms, textbooks are used as a source of information and practice tasks (Mullis, Martin, Foy, & Arora, 2012; Wakefield, 2006). However, textbooks often tend to send an implicit message that the focus of mathematics is to swiftly perform mechanical computations of correct answers rather than to encourage a conceptual learning of mathematical structures (Lithner, 2004; Newton & Newton, 2007; Shield & Dole, 2013). Many textbooks also include more tasks then is reasonable for a student to complete within the frame of the course

(Jäder, Lithner, & Sidenvall, 2014). It is also apparent that many of the more demanding tasks are located at the end of each chapter (Jäder et al., 2014). This implies that a selection must be done and if students are supposed to make this selection of tasks they tend to choose the first task sets (Sidenvall, Lithner, & Jäder, 2015), which increases the proportion of routine tasks further. In a cross-national study of textbooks from twelve countries Jäder et al. (2014) concluded that 79% of the tasks could be solved completely by imitating or following given instructions while only 9% of the tasks required more extensive conceptual knowledge and justification. If the textbooks mainly promote algorithms and rote learning then the mathematical foundation and conceptual knowledge will, most probably, not be developed. Here the teacher has an important mission to fill the gap between mechanical calculations often presented in the textbooks and the conceptual understanding and competencies that the syllabi often calls for.

2.2.2 Teaching

In a study on 200 teachers' implementation of reform-based syllabi, Boesen et al. (2013) found that there was an emphasis on procedural activities during mathematics lessons both in teacher presentations and during students individual (or small group) work. Focusing on the teacher presentations, Bergqvist and Lithner (2012) concluded that routine tasks and simplifying explanations were commonly used. Teachers often used quite thorough explanations when presenting new topics but often without verifications or connections to intrinsic mathematics (Bergqvist & Lithner, 2012). It seems logical, that if we want students to become skilled at solving novel tasks and at justifying their solution methods, teachers must demonstrate this during presentations. Studies have shown that students often motivate their choice of solution methods by looking at superficial properties and poorly memorized algorithms (Bergqvist, Lithner, & Sumpter, 2008; Hiebert, 2003; Lithner, 2000, 2003). Stein et al. (1996) also saw this behavior among the students that they studied. Many students preferred to use known procedures even if the procedures did not fit the task at hand. Stein et al. (1996) also concluded that many tasks lost their challenging quality due to poor help from the teacher or a shifted focus towards the correctness of the answer.

An important factor for the the task to maintain its complexity and its conceptual challenge is the type of help the teacher provides the students with. Henningsen and Stein (1997) conclude that teachers that select appropriate and worthwhile tasks, press for justifications, and "support students' cognitive activity without reducing the complexity" will help students to reach further. The teacher also has an important role in showing the class what high-level performance should look like and in giving appropriate time constraints to the tasks (Henningsen & Stein, 1997; Stein et al., 1996).

If teachers want students to become problem solvers and students expect to learn an algorithm or simple rules there is a problem. Add to this that algorithms, although effective and secure, according to Brousseau (1997) are designed to avoid meaning and there won't be that much room left for the mathematical concepts and properties that teachers probably wish to communicate.

3 Theoretical frameworks

This thesis has the aim to analyze how task design and mathematical reasoning can affect the learning of mathematics. To do this there is a need to frame it with theories and frameworks that are relevant for the coming analysis. I will do so by starting off with Brousseau's (1997) theory of didactical situations, which lurks at the center of most of the included studies in this thesis. Brousseau's thought that students need to take responsibility of the task solving process to learn the intended knowledge has also been discussed by others (e.g., Bjork & Bjork, 2011; Hiebert & Grouws, 2007; Jonsson, Kulaksiz, & Lithner, 2016) as the importance to struggle with central mathematical concepts. This struggle can of course be accomplished in different ways and engaging in problem solving is one of the ways that have been studied and discussed extensively ever since Pólya (1945) wrote his famous book 'How to solve it'. However, OECD (2015) concluded that teachers in average have seven hours a week to spend on lesson preparation and as Blum and Niss (1991) infer, mathematics teachers are afraid that problem solving will take too much time. They also mention that problem solving can be viewed as a challenging and slightly overwhelming project to embark on for many teachers, since additional non-mathematical knowledge is necessary. A slightly more reasonable effort could instead be put into constructing or adjusting ordinary tasks so that they require more justification and conceptual understanding rather than procedural skills. One way of doing this is to apply Lithner's (2008) framework on mathematical reasoning to task design. This framework will be described and the types of reasoning that Lithner suggests will be defined, together with a new type that is more connected to task design than to student reasoning. Later on Lithner's framework will also be connected to student cognition, as tasks' cognitive demand may differ as Stein et al. (1996) suggests. This will then lead us into chapter 7 that describes cognition and its influence on mathematics education.

3.1 Theory of didactical situations in mathematics

Learning mathematical concepts and strategies, to be able to construct or reconstruct solution methods if the old ones are insufficient or forgotten, is a major goal in mathematics learning, but as algorithms often are provided the aim instead becomes to memorize and make use of them (Hiebert, 2003)

In Brousseau's (1997) Theory of Didactical Situations (TDS) one of the central ideas is that if students are to learn mathematics they have to construct the key concepts by themselves. Brousseau (1997) argued for a task or problem design where the students have to construct at least some part of the solution by themselves as this is central for a learning that goes beyond mere memorization of a method.

However, Brousseau argues that much of the work in classrooms is done interacting with teaching materials, peers, and the teacher. The interaction with this *milieu* can, if done poorly, destroy the learning opportunity, for example by having the solution method given in the textbook or by having a peer telling the answer. Thus, Brousseau took into account that the milieu was important to relate to when discussing the learning of mathematics.

For the student to be able to go beyond given algorithms the teacher must arrange the *devolution of a good problem* (Brousseau, 1997). This requires that the student gradually takes responsibility for the solution process which ends with the construction of a justified solution. If this devolution occurs the student will enter an *a-didactical situation* devoid of the teacher's didactical intentions and where the teacher is separated from both the student's progress and learning. Here the responsibility for solving the task falls completely on the student and the teacher has to release control of the solution and the learning to the student. For this devolution to take place a mutual relationship that states what responsibility the student will have during this process and what the teacher's duty is will implicitly be agreed upon (e.g., what effort the student should give the problem before calling for help or what kind of support the teacher will give when called for). This informal and often non-spoken agreement is called a *didactical contract* (Brousseau, 1997). Brousseau underlines the shared responsibility of teacher and student if the contract is broken. The teacher has the responsibility for the results by designing solvable problems that give rise to a natural a-didactical situation. Simultaneously, the student has to accept a problem solving situation where the solution method has not yet been taught. This can be hard for some students to accept at first as many of them are used to apply given algorithms to solve tasks (Hiebert, 2003), not having to struggle with mathematical properties or concepts.

3.2 Productive struggle

Imagine for a moment that you visit a friend in a city where you have not been before. She picks you up at the station and you walk together, first to her apartment and then to a restaurant to eat dinner. All the time, as you walk through the city, you talk and have a wonderful time. Suddenly you realize that you have no idea where you are, no idea of how you got there or how you would find your way back if your friend suddenly would leave. If you took this walk with your friend every day for a week you would probably learn the way even though you don't know anything else about the city. Now, imagine that you come back a year later, the restaurant has moved and so has your friend. How will you find the way? Well with a map of course. This could be a little tricky and you'd have to notice more things (e.g., architectural markers and street names). It could be a bit of a struggle but eventually you'll get to the right place. After finding your way this time you would be more confident when visiting this particular city again.

This every-day example illustrates the importance of the need to struggle with central objects or concepts. There is however an important difference between a struggle and a *productive struggle*. The former could be achieved by having students doing mathematics in a dark room or by giving the students extremely difficult tasks. This struggle would be considered unproductive or undesirable. Tasks that impose a surmountable productive struggle with intrinsic mathematical ideas may give a more lasting impression or knowledge. This is exactly what happens when students are working with well-designed CMR-tasks (Jonsson et al., 2016). The student will have to struggle since there is no apparent way of solving the task and if the task design is good this struggle will not be overwhelming and negative but rather manageable and positive for learning.

The importance of struggle for learning has been noted by researchers in both educational science (e.g., Hiebert & Grouws, 2007; Jonsson, Norqvist, Liljekvist, & Lithner, 2014; Niss, 2007) and cognitive psychology (e.g., Bjork & Bjork, 2011; Pyc & Rawson, 2009; Wiklund-Hornqvist, Jonsson, & Nyberg, 2014). Within the realm of educational science, the importance of productive struggle for learning of mathematical concepts are discussed. Hiebert and Grouws (2007) give an overview of the significance of having to put some effort into learning something. They argue that the effort that is directed to the task at hand will be beneficial for learning. In a study where students practiced solution methods by either given algorithms or by constructing algorithms Jonsson et al. (2014) saw that the more effortful construction process was more beneficial for learning than using a given method. They discussed if this had to do with the struggle itself or if the design of the practice and test tasks could influence this (i.e., that the constructive practice tasks were similar to the test tasks while the imitative practice tasks were not). Jonsson et al. (2016) took this discussion further and examined if an effortful struggle was more influential for learning than practicing in the same way as being tested (i.e., transfer appropriate processing). The conclusion was that an effortful struggle that focused construction of the solution method was more beneficial. This resounds well with findings in cognitive psychology where there is much evidence for the benefits of struggling with important concepts and structures. Pyc and Rawson (2009) concluded that effortful retrieval from memory will be more beneficial for learning than easy retrieval. While practicing, this effortful retrieval can be achieved by repeated testing. Wiklund-Hornqvist et al. (2014) showed that repeated testing was more efficient for later retrieval than re-reading information. The repeated testing invokes more afterthought than just re-studying a concept and this is also considered as an effortful and productive struggle[2]. Bjork and Bjork (2011) summarizes the importance of what they chose to call desirable difficulties in the following statement: “Conditions of learning that make performance improve rapidly often fail to support long-term retention and transfer, whereas conditions that create challenges and slow the rate of apparent learning often optimize long-term retention and transfer”. To create some challenge or productive struggle, novel tasks (e.g., problem solving tasks) can be utilized.

3.3 Problem solving

One of the competencies that are stressed as important in many mathematics syllabi (NCTM, 2011; Skolverket, 2011b) is to become a proficient problem solver. In the Swedish syllabus this competence is formulated as the ability to “formulate, analyze, and solve mathematical problems and also evaluate strategies, methods, and results”.

A student always brings prior knowledge into every task-solving situation. It could for example be that the student has seen similar tasks before and therefore know how to embark on solving the task. Schoenfeld (1985) describes four different categories of knowledge that contribute to problem solving. The first is *resources*, the content knowledge that the student has acquired during previous schooling and that could be of importance for the particular task at hand. This could for example be knowledge about how to subtract 9 from 4 or how to differentiate the function $f(x) = 1/x$. If a student experience gaps in her resources, let’s say that she does not know how to add fractions, the learning of algebra could be impaired.

Schoenfeld’s second category is *heuristics*, the strategies and techniques needed to solve the problem. Here we talk about methods of solving tasks and in what order procedures should be done. The idea of heuristics in problem solving was first formulated by George Pólya in 1945. Pólya (1945) formulated four heuristic principles that could be applied to all problem solving: understand the problem, devise a plan, carry out the plan, and revise your work. More specific heuristics has of course also been formulated (e.g., ‘draw a figure’ or ‘try to solve a simpler task’).

Schoenfeld’s third category is *control* (or as he later renamed it *meta-cognition*), in this case control over which strategies and resources to select and use. This includes reflecting on your own thoughts and on your available knowledge to choose wisely. All teachers have seen examples of lacking control. It might be that students use the addition strategy of common denominators when multiplying fractions or that students solve non-existing equations while simplifying algebraic expressions. A student with good

[2] For a review on the testing effect, see Dunlosky, Rawson, Marsh, Nathan, and Willingham (2013).

control will make the most of her resources so that she will be able to solve novel tasks in a more efficient way.

Lastly the student's *belief systems*, his or her personal view on mathematics, is important for how the student will tackle novel tasks (Schoenfeld, 1985). This would include the thoughts you have about mathematics as a subject and your mathematical abilities. A student whose mathematical beliefs are poor will probably be more prone to give up on novel tasks and will have a harder time to control his or her resources and heuristics.

Additionally, Jackson, Garrison, Wilson, Gibbons, and Shahan (2013) argued that contextual aspects also can be important to consider during task setup, to bridge the eventual gap of information if there are contextual features that are unfamiliar to the student. They argue that key contextual features should be explicitly addressed to make task solving more effective.

Schoenfeld elaborated on Pólya's four steps by describing the problem solving procedure as containing six possible stages: reading the task, analyzing the task, exploring methods, making a plan, implementing the plan, and verifying the result. Not all of these stages have to be present during the solving of a problem. For example, the analysis of the task could be enough to generate a plan on how to proceed and then the exploration phase would be un-necessary (Schoenfeld, 1985). Schoenfeld also discovered that there was a significant difference between novice problem solvers and experts in how much time they put into analyzing a problem. Novice problem solvers typically decided quickly on an approach and pursued it even if there was clear evidence that the strategy was not bringing them closer to a solution. Experts put more time in analyzing the task, formulate and implement a plan, and verifying it to be able to go back and re-think the strategy if needed (Schoenfeld, 1992).

However, Blum and Niss (1991) indicates that teachers seems to think that even though problem solving is important it will i) need additional knowledge about other subjects and ii) take much time to implement. Time is also expressed as an issue by Boris (2003) when comparing teachers' mathematical beliefs and their practice. Another, and maybe more time efficient, way to focus the important mathematics during task solving could be to engage in tasks that promote creative mathematically founded reasoning. Creative tasks do not have to be as challenging as problem solving and can include elementary reasoning as well as more elaborated reasoning. This will however require that the task-design put emphasis on a particular mathematical hurdle that the students need to learn, and that the task does not reveal the solution method for the student. This is basically what creative mathematically founded reasoning tasks does as we will see in the following section.

3.4 A framework for mathematical reasoning

As much of this research is based on the research framework for mathematical reasoning that Lithner (2008) suggested, it seems appropriate to give a short summary of the different aspects of it. The framework provides a basis to analyze student's reasoning, primarily with respect to the distinction between using available (memorized or given) solution methods and constructing the solution. It can also be used to classify mathematical tasks with respect to the mathematical reasoning they promote and/or assess. The reasoning promoted by the task is depending on the individual's prior knowledge and the text, guidance, or examples that are available at the time of task solving. The reasoning sequence starts with the given task and continues to an answer and the reasoning that is carried out is the product of the task, the individual's thoughts, and the milieu.

3.4.1 Reasoning sequences

When solving a mathematical (or maybe any other) task you have to decide where to start. Schoenfeld (1985) observed that novice problem solvers often put less time into preparation and choosing than experts. While the experts put a lot of thought into preparation the novices were quicker to dive into an unprepared and often unsuitable problem solving process. The solving process could, as Lithner (2008) suggests, be seen as a directed graph where implementation of a solution strategy (edges) are connected by instances (vertices) which indicate both a momentary state of knowledge and of the subtask (see Figure 1). These subtasks comprise both explicitly written subtasks and implicit subtasks that the reasoner formulate during the solution process. The edges consist of solution processes that are more or less outspoken, where the reasoner is implementing the strategy of choice for the specific subtask. This implies that the task at hand can be solved or answered along different paths through the graph.

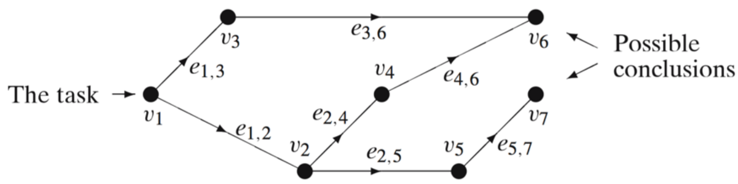

Figure 1: Reasoning sequence as retrieved from Lithner (2008).

Let me give you a simple example of this with the following task: *Jane has a salary* of *€1800 per month. How much will her salary be if she will get a 15% raise?* Depending on the prior knowledge of percent this task could be solved by at least three paths.

Path 1:

Find what 1% of €1800 is.

a. 1%: €1800/100 = €18

Find what 15% of €1800 is.

b. 15%: 15 · €18 = €270

Add €270 to €1800.

c. €1800 + €270 = €2070

<u>Answer:</u> Jane’s salary will be €2070.

Path 2:

Find what 15% of €1800 is.

d. 0,15 · €1800 = €270

Add €270 to €1800.

c. €1800 + €270 = €2070

<u>Answer:</u> Jane’s salary will be €2070.

Path 3:

Calculate the new salary by finding 115% of the old salary.

e. 1,15 · €1800 = €2070

<u>Answer:</u> Jane’s salary will be €2070.

If we should try to draw this simple example (provided that the task solver does not make any other assumptions or calculations than what is given here) the graph would look like Figure 2.

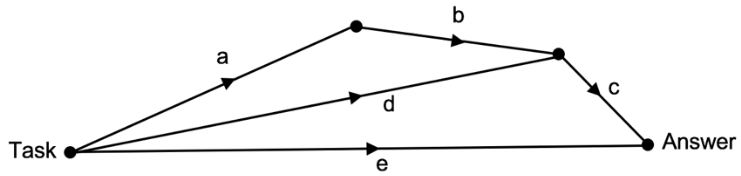

Figure 2: Reasoning graph of the task-solving example above.

3.4.2 Creative Mathematically Founded Reasoning

Lithner (2008) identifies two major reasoning types, imitative and creative reasoning, and then proceeds to divide these into sub-categories. When there is not enough information at hand to solve the task with a known solution method (i.e., by an algorithm or by recalling memorized answers) it can still be solved but another type of reasoning must be used. At least some parts of the reasoning sequence must then be constructed by the task solver and argued for by connecting it to the intrinsic mathematical properties important for the task. Reasoning that involves both novelty and mathematically founded arguments is called Creative Mathematically founded Reasoning (CMR). CMR is defined by Lithner (2008) as follows:

> *Creative mathematically founded reasoning* (CMR) fulfils all of the following criteria.
>
> 1. Novelty. A new (to the reasoner) reasoning sequence is created, or a forgotten one is re-created.
> 2. Plausibility. There are arguments supporting the strategy choice and/or strategy implementation motivating why the conclusions are true or plausible.
> 3. Mathematical foundation. The arguments are anchored in intrinsic mathematical properties of the components involved in the reasoning.

The creativity here should not necessarily be seen as something extraordinary or ingenious but rather as the construction of a, for the task solver, new reasoning sequence (Lithner, 2008). As an example, a task that asks for the area of a triangle with a given height and base could be considered a creative task (denoted CMR-task) if there is no provided formula or if the students have not done this previously. The students would have to base their reasoning on what they already know (e.g., the area of a parallelogram) and then consider the triangle to be half a parallelogram. After this we could also ask the students to formulate the rule or formula by themselves. Since mathematics is an ingredient in other school subjects, CMR could also be applied in them. For example, Johansson (2015) showed that CMR can be an important element when learning physics.

3.4.3 Algorithmic Reasoning

As a contrast to CMR, reasoning that is connected to performing a recalled procedure without connecting it to mathematical properties is called Algorithmic Reasoning (AR). Lithner (2008) defines AR as follows:

> *Algorithmic reasoning* (AR) fulfils the following two conditions.
>
> 1. The strategy choice is to recall a solution algorithm. The predictive argumentation may be of different kinds (see below for examples), but there is no need to create a new solution.
> 2. The remaining reasoning parts of the strategy implementation are trivial for the reasoner, only a careless mistake can prevent an answer from being reached.

A task would be categorized as promoting algorithmic reasoning (denoted AR-task) if it is reasonable to think that the solution method could be retrieved from memory by the solver, or if the solution method is available in the instructions or a worked example (Lithner, 2008). AR-tasks seems to be quite common in textbooks across the world and at all levels of mathematics education, from compulsory school to university (e.g., Jäder et al., 2014; Lithner, 2004; Newton & Newton, 2007). An example of an AR-task could be when students are asked to calculate the area of a triangle where the height and base are provided and the formula, $A = (b \cdot h)/2$, is written at the top of the page. The focus will be to apply the formula correctly and will likely not include considerations of mathematical properties like that the triangle is half of a parallelogram (hence, the division by 2).

Lithner (2008) differentiates between different types of algorithmic reasoning, depending on what type of AR-information the task solver make use of. Commonly, the supplied AR-information will put focus on how to solve the task and not why the task can be solved in the given way. In study 2, I discuss and test another type of AR, eXplained Algorithmic Reasoning (XAR), which concerns the reasoning that occurs when a student has access to both a solution method and an explanation on why the solution method is valid. It is important to distinguish between a description that tells *how* a solution method should be applied and an explanation on *why* the solution method is valid. The former would be categorized as AR-information since it gives explicit instruction on how to perform the calculations, without explicit connection to the mathematical properties. The latter would be more than AR information since it not only describes but also justifies the solution method. The justification included in XAR could be similar to the justification that is constructed during CMR with the difference that in XAR the justification is available from the start and in CMR it is constructed as a part of the reasoning sequence.

For example, the area of a triangle could be introduced by giving the formula (i.e., $A = (b \cdot h)/2$) and showing how to apply it. This would be considered AR-information since there is no connection to intrinsic mathematical properties. The introduction could also explain why the formula is valid by describing and showing that a triangle is half of a parallelogram, ending with the solution method that the students could apply. This introduction would be classified as XAR information since it starts off from the mathematics behind the formula and explains the validity of it from this point of view. In this way the formula can be logically founded in mathematics and not only something that you just may have to accept and believe in.

Most textbook information seems to concern the description of solution methods rather than presenting the reasons for why these solution methods work (Shield & Dole, 2013; Stacey & Vincent, 2009). However, even though XAR can be found in textbooks it is always accompanied with solution methods and/or examples that are highlighted and therefore gives the impression that they are the most important pieces of information.

Thus, task design can influence which reasoning the student will use. It is also plausible that a creative task will be more cognitively demanding than an algorithmic task since the student has to construct new solution methods and not only apply provided algorithms. If cognition is crucial for creative reasoning to occur, then cognitively less proficient students will have problems even if the tasks are well designed.

3.5 Cognitive demand

Another way to characterize mathematics tasks is to sort them regarding to the cognitive demand they impose on the task solver. Stein et al. (1996) made such a distinction between tasks when trying to find factors that contribute to the preservation of cognitive demand throughout the solution process. They defined five categories of tasks based on what was demanded from the students to be able to solve them (non-mathematical, memorization, procedure without connection to concepts, procedure with connection to concepts, and doing mathematics). In their study it became clear that the tasks with higher cognitive demand (procedure with connection to concepts and doing mathematics) often lost much of this demand during the teaching and solution process. Most of this happened when the teacher (or sometimes a peer) provides help by removing the challenging aspects of the task or when the focus shifts from the concepts to finding the correct answer. This reduces a cognitively demanding task to a task where the only aim is to apply the correct procedure.

There are some connections between the way Stein et al. (1996) categorized tasks by cognitive demand and the way Lithner (2008) has categorized tasks by looking at the reasoning they will promote. Tasks that according to Stein et al. (1996) requires memorization or procedure without connection to concepts are similar to the tasks that Lithner categorizes as imitative- or algorithmic reasoning tasks (Lithner, 2008). Here the student can rely on either a memorized solution method or a method given by the textbook or by a person close by (i.e., the teacher or a peer). The tasks that require procedures with connection to concepts or that students engage in doing mathematics are comparable to Lithner's creative reasoning tasks. Here the students need to consider the mathematical properties to solve the task, either by reflecting on why a known procedure would be appropriate to use or by constructing a new mathematically founded (and justified) solution method.

Hence, there are implications that cognitive variation is a part of the puzzle, and if an individual's cognitive abilities are important for task solving and if CMR requires a capability to handle higher cognitive demand, cognition could be decisive in how students learn from solving tasks. If students' individual cognitive variation matters it is important to examine this as well as deciding which measures to use when doing so. This will be addressed in the next chapter.

4 Memory and cognition

4.1 Individual variation in cognition

How a student will handle the requirements of doing mathematics could vary a lot depending on individual prerequisites. This could for example be how well a student will be able to concentrate in a noisy classroom or if the student feels motivated to engage in the sometimes stressful conditions of a test situation. Doing mathematics sometimes taxes the individual's cognitive abilities quite extensively. There are lots of abstract information that need to be processed and even though most of us use pen and paper to ease the cognitive strain, high processing power could be important for mathematics achievement (e.g., Floyd, Evans, & McGrew, 2003; Freund, Holling, & Preckel, 2007). Hence, individual variation of cognitive ability could be an important factor to consider when studying mathematical tasks and reasoning. In the project we have therefore chosen to include and control for some measures of cognitive capacity in our experimental studies.

4.1.1 Working memory

One cognitive construct that is often connected to mathematical thinking is working memory (WM). This is the ability to simultaneously store and process information. The multi-component model of WM was first suggested by Baddeley and Hitch in 1974. Baddeley has made some additions to this model and it now contains four parts. There is a Central Executive that coordinates incoming information to the three slave systems, 1) the Phonological Loop that process auditory information, 2) the Visuo-Spatial Sketchpad that process visual and spatial information, and 3) the Episodic Buffer that handles the temporal part of the acquired information so the stories we remember are episodically coherent (Baddeley, 2000).

The connection between WM and mathematics achievement has been extensively studied and, in the chapter on working memory in the first volume of Educational Psychology Handbook, Swanson and Alloway (2012) conclude that there is much scientific proof of a link between mathematics achievement and WM. For example, Bull, Espy, and Wiebe (2008) let primary school children do both tests of mathematics skills and of WM. The tests showed a high correlation between visuo-spatial WM and math skills in primary school children. Passolunghi, Vercelloni, and Schadee (2007) also conducted a study on primary school children with the similar results, that WM is important for mathematics achievement. However, Swanson and Alloway (2012) also note that WM is not the only important factor in mathematics learning.

4.1.2 Fluid intelligence

WM is also closely connected to a construct that cognitive psychologists refer to as General fluid intelligence or fluid reasoning (Gf). This is the part of the human cognition that is devoted to problem solving. It is not unexpected that these two are interconnected. While solving any problem we need to activate WM as we need to store and process information at the same time, so the fact that Gf and WM account for some of the same processes is not strange. However, WM and Gf are not the same construct. In a study on how time constraints influence the correlation between WM and Gf, Chuderski (2015) showed that stricter time constraints on Gf tests produce a higher correlation between Gf and WM. When time constraints are removed the correlation between the two decrease. This indicates that there are processes involved in Gf that are not directly linked to WM.

There are studies that link Gf to mathematics achievement. For example, Primi et al. (2010) argue that a higher Gf, or at least higher results on Gf-tests (e.g., Raven's Progressive Matrices), implicate a steeper mathematical learning curve (i.e., faster learning). Taub, Keith, Floyd, and McGrew (2008) conducted a study on children and youths, 5-19 year olds, where they conclude that Gf is an important factor for mathematics achievement in all age groups. However, they also mention that other factors play in when it comes to how well students perform in mathematics. As CMR contains elements of problem solving, Gf could be influential for their reasoning and in turn affect the students' results.

4.1.3 Cognitive tests

Most cognitive functions are studied by using different behavioral tests designed to test the sought after ability. In the case of WM there are a number of tests available. Typical for all these is that you are supposed to process some information whilst remembering other information. For example, reading and judging the validity of a few sentences whilst remembering the last word of each sentence (i.e., reading span) or performing simple arithmetic whilst memorizing letters (i.e., operation span) (Unsworth & Engle, 2005). Gf is most often evaluated by applying a Raven's progressive matrices test which comes in a few different levels of difficulty: colored, standard, and advanced. This test is a non-verbal problem solving test where the participant has to choose the correct tile that will complete a three-by-three matrix (see Figure 3). Raven's progressive matrices are supposed to be independent of the participant's language and culture but an increase in scores has been observed over time (e.g., Brouwers, Van de Vijver, & Van Hemert, 2009; Raven, 2000; Wongupparaj, Kumari, & Morris, 2015).

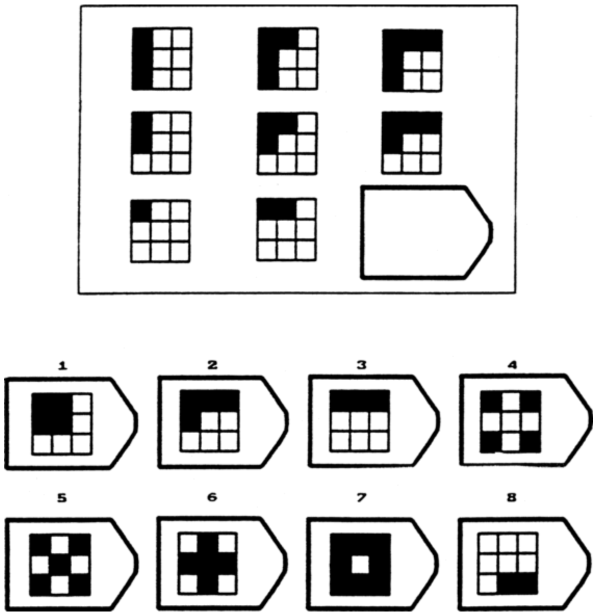

Figure 3: Example of a task similar to those used in Raven's Progressive Matrices.

4.2 Observing brain activity

Like all other organs in our bodies, the brain will develop during childhood and adolescence. Therefore, some cognitive processes that are trivial to an adult might be impossible to perform for a child. For example, the concept of time is very hard to grasp for a child while an adult finds it mostly unproblematic.

Brain imaging techniques can provide support for hypotheses about educational issues (De Smedt et al., 2010). For instance, in study 3 (in the present thesis) we showed that different mathematical practice provided, not only behavioral difference, but also long lasting neural differences in the brain (Wirebring et al., 2015a). These results supplied evidence that the hypotheses of which brain regions that became more or less active correspond to the previous behavioral data and thus strengthen the conclusion drawn about the behavioral results. By using brain-imaging techniques, it is also possible to detect cognitive processes that are not manifested in observable behavior. It has even been shown that the brain activity can be prognostic of future behavior. Wirebring et al. (2015b) showed that word pairs that were retrieved but subsequently forgotten were characterized by lower brain activity than word pairs that were retrieved and remembered. Hence, brain-imaging studies can provide information on how the brain process information that is impossible to obtain in behavioral studies. By combining brain-imaging and behavioral studies, we can more effectively evaluate different methods of learning.

Techniques have been developed to study brain activation and in later years this has been important for cognitive psychologists since the complex nature of the brain earlier only could be studied in special cases where injuries or illness did disrupt the normal brain function. Electro-encephalogram (EEG), a non-invasive technique that measures electric brain activity via electrodes attached to the scalp, has been available since the mid 20th century but this technique has a disadvantage, its lack of spatial resolution. Signals are quickly detected but it is much harder to locate their origin within the brain. To study neurological processes that correlates to cognitive activities like language or mathematics there is a need to increase this spatial resolution. Functional Magnetic Resonance Imaging (fMRI) is a comparatively new technique of registering brain activation. In contrast to classic EEG, fMRI has a high spatial resolution but a somewhat lower temporal resolution. This is due to the biological and physical processes that constitute the signal source during fMRI. This is quite complex, but in the following section I will try to explain it without going into the deeper physics of it.

4.2.1 The technique behind an fMRI image

Every voxel[3] in a typical fMRI-image is about 3x3x3 mm (i.e., about 1/500 of a teaspoon). fMRI depicts the brain (or any part of the body) in several "slices" (usually 27). The scanner makes a pass over the brain in about 2 second and this is then repeated several times with several stimuli to get a reliable picture (movie). fMRI utilizes magnetic properties of hemoglobin molecules in the blood to measure blood flow as this indicates that there is activity.

In the strong magnetic field (1,5-7T, an ordinary refrigerator magnet produces about 0,005T) within the scanner the hydrogen (nuclei) in the body will be positioned in line with the magnetic field. In the scanner there are emitters that send out radio waves in specific frequencies that will excite hydrogen atoms within water molecules. The hydrogen absorbs the radio wave energy and this will change its orientation to one angled to the magnetic field. When the radio pulse subsides the excited hydrogen return to the original orientation and release the excess energy as an energy pulse back to the detectors.

To be able to discern between the different parts of the "brain slice" the scanner is equipped with gradient coils that induce gradient magnetic fields that utilizes this sensitivity of inhomogeneous magnetic fields. This has the effect that the hydrogen atoms will orient differently depending on where in the slice they are located. This will give rise to a slightly diverse timing in the

[3] A voxel is the smallest visible element in a depicted volume. Comparable to pixels when dealing with digital photography.

return signal that is correlated to the gradient fields. And by this it is possible to determine a three-dimensional coordinate of the return signal, i.e., in which voxel the activation took place.

As the brain activates a certain area the consumption of oxygen will increase and the blood flow will need to increase accordingly. The body does this by enlarging the capillary vessels in the active area so that they will let more oxygenated blood through. The ratio of oxygenated hemoglobin will actually rise to a higher level than normal and this provides us with an opportunity to distinguish the active areas from the surrounding, non-active, tissue. As the de-oxygenated hemoglobin molecule is paramagnetic it distorts the return signal from the tissue but when the concentration of deoxyhemoglobin decreases, as the concentration of oxyhemoglobin is rising, the signal will become stronger. This contrast between in deoxy- and oxyhemoglobin gives us the possibility to detect a useful signal (often called a BOLD-signal). For a more comprehensive description of fMRI see for example Huettel (2014).

4.2.2 A downside to fMRI

A weakness of this imaging technique is that the fMRI-scanner gives us an indirect image of neural activation. This means that we can’t be absolutely sure that the increase of blood flow indicates that the main activity is depicted. The scanner can’t distinguish between neuron clusters forwarding information and clusters that does the actual processing of the information. This makes it even more important to use a good experiment design and relevant contrasts (e.g., comparisons between a) reading and finding spelling errors and b) reading and solving mathematics tasks) to make sure that what is being measured is the active brain area (Huettel, 2014).

4.2.3 fMRI and Mathematics

There has been quite extensive research in cognitive neuroscience about which brain areas that are active when mathematics is involved. However, most of the studies are done concerning quite simple mathematical concepts, e.g., addition, subtraction or one-digit multiplication. This is partly because the fMRI-technique is comparatively new but also due to the fact that the technique places restrictions on the experiment design. Time constraints are a reason that more complex mathematical procedures have not been studied in fMRI. A long exposure time would decrease the signal-to-noise ratio and give few results or results that are difficult to interpret. There is also a problem with complex tasks that would demand pen and paper or similar tools to be solvable, since physical movement or speech also generate brain activity. This would add activation patterns to the result and can sometimes be hard to distinguish from the pattern that you are really interested in (Huettel, 2014).

The studies that have been made (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003; Delazer et al., 2003; Houde, Rossi, Lubin, & Joliot, 2010; Ischebeck, Zamarian, Egger, Schocke, & Delazer, 2007; Wirebring et al., 2015a; Zamarian, Ischebeck, & Delazer, 2009) have all identified interesting neural networks that are in use while making calculations, solving novel tasks, or recalling from memory. Some of these networks are common with other subjects as well (e.g., connections to long-term memory or processing of complex tasks) and some are more specifically linked to mathematics (e.g., abstract number sense or calculations). In the following section I will describe the main neural networks and their connection to mathematics based on their function.

4.2.4 Brain functions connected to mathematics

There are some basic functions that are in use when a person is doing calculations or more complex mathematics. There are rules or principles that we have learned some time ago that need to be retrieved from memory. We have to make simple calculations using for example addition or multiplication. While we do this there is a need for our working memory to briefly store and retrieve information that is being processed. If the task at hand is novel or complex we need to figure out how to handle this new situation and how to use our prior knowledge to solve the task. There might also be visual representations to consider or to manipulate in some way. All this is handled by the brain although in different areas. The brain is a very complex and interconnected organ and most processes activate large portions of the brain.

4.2.4.1 Memory retrieval

Retrieval from long-term memory is important for all tasks, mathematical or not. Important ideas, rules, or rote learned knowledge can be found here and can be more or less easy to retrieve. It is fairly simple to recall rote-learned information (e.g., the multiplication table) which can be useful to relieve our working memory by automatization. However, since rote-learned information does not carry information on underlying ideas or basic concepts it is not always obvious which rote-learned knowledge that should be used. For example, I still recall a few German words that were supposed to control *direct object* (durch, für, gegen, ohne, um), but since I cannot remember what a *direct object* is I have little use of them. As discussed previously, knowledge that is achieved with some effort will be easier to recall (e.g., Bjork & Bjork, 2011; Jonsson et al., 2016). For example, Wirebring et al. (2015a) (study 3 in this thesis) showed that students that learned mathematics by a given solution method had to work harder to recall this method than students that constructed the solution method themselves. The former group also showed a higher activation was in the *Angular Gyrus (AG)*, an area related to for example reading, mathematics, memory retrieval and social cognition (Seghier, 2013).

Dehaene et al. (2003) argue that the angular gyrus is active during mathematics since verbal or linguistic properties are the basis of arithmetic tasks. They also propose that rote learned addition and multiplication are stored in verbal memory in the same way that grammatical rules can be remembered as a ditty. This view is shared with Ischebeck et al. (2007) as they observe that the angular gyrus and temporal lobes are activated when retrieving rote-learned mathematics, i.e., simple addition and the multiplication table. Perhaps this activation of the angular gyrus is a marker for memory retrieval (Grabner et al., 2009) or maybe it is an indication that the trained "knowledge" is manifested verbally (Delazer et al., 2003; Zamarian et al., 2009). The AG has been found to be active in non-mathematical tasks as well and this could indicate the verbal connection.

4.2.4.2 Novel and complex tasks

When a person encounters a novel or complex task, working memory will be activated to a higher degree. There is also need for memory retrieval and structuring of information and prior knowledge. Much if this work is done in the frontal lobe of the brain although there are also other parts of the brain that seem to have a part in manipulation of information in working memory (Koenigs, Barbey, Postle, & Grafman, 2009). Studies have shown that higher complexity yields more activation of the *Prefrontal Cortex*, an area connected to problem solving and working memory. This area is more active during childhood then during later years (Houde et al., 2010; Zamarian et al., 2009). This might not be unexpected since more tasks are novel as you are younger. Complex tasks often have a visual component as well, either as a sketch, diagram or graph or as information which induce the solver to make mental pictures. The visuo-spatial processing is conducted in the *Posterior Superior Parietal Lobule*, further back in the brain (Dehaene et al., 2003). This indicates that a truly complex mathematics task will be hard to study in fMRI since there are so many areas active at the same time. There is also the temporal difficulty mentioned previously, that a lasting task will decrease signal-to-noise ratio and give results that are harder to analyze.

4.2.4.3 Calculation and number sense

Most mathematics, at least in pre-university schooling, include numbers and calculations with numbers to some degree. Some of the rote-learned computations can be retrieved from long-term memory without the need for processing, but mental arithmetic is still needed to solve even the simpler tasks. Much calculation is conducted in working memory but there is also an area that seem to be activated only during mathematics, the *Horizontal Intra-Parietal Sulcus*. Dehaene et al. (2003) observed that this area was not activated by words in general but by number words, which led them to conclude that this area was the most mathematically specific of the three they studied. The horizontal intra-parietal sulcus is proposed to code the abstract meaning of numbers and is activated by mental arithmetic, e.g., subtraction.

4.3 Summary

As mathematics is activating a large portion of the brain, most studies have to be designed to pinpoint specific brain functions or type of mathematics. Additionally, increased time for reflection would add interference to the fMRI signals, and so would also aid in form of pen and paper, calculators or asking questions do. Put together, examining mathematics with fMRI is not an easy task but with considerable thought on experiment design it still is possible to distinguish brain activation connected to mathematics from other cognitive tasks (e.g., reading). One such experiment is reported in study 3 in this thesis.

The previous sections have given us an overview over the fMRI-technique and results from some previous studies connecting brain activity to the learning of mathematics. There are some benefits in using this kind of methodology since fMRI can give a deeper insight into the brain processes that govern our behavior. The fMRI-technique can also help to sort out processes that behavioral studies will have a harder time to pick out. For example, in study 3 we used fMRI to help with the explanation to the significant advantage of CMR-practice over AR-practice in study 1. In study 3 we could see that the CMR-group activated brain areas connected to memory retrieval processes in a significantly lesser degree than the AR-group did. The AR-group also had higher activation in areas connected to working memory. Together with the results from study 1 this could explain how AR-practice differs from CMR-practice. Students that practice with CMR seems to have an easier access to the practiced solution methods and easier to apply them than the AR-students will have. This was of course implicated by the behavioral results in study 1 as well, but the question about if the difference in test results in study 1 could have been due to a higher activation for CMR-students in other areas where complex tasks are processed. However, study 3 showed no other areas within the mathematics network where the CMR-students had higher activation levels than the AR-students. Therefore, the reason for the significant advantage of CMR over AR is somehow connected to the deeper encoding and ease of retrieval of CMR-practiced solution methods. This result could maybe have been found out in a series of behavioral studies as well but the use of fMRI gave us this result in a single experiment. The fMRI-experiment also gave us implications in what to pin-point in coming studies, namely, the reason for the deeper encoding of CMR-practiced solution methods.

5 Method

When the project 'Learning mathematics by Imitative and Creative Reasoning' began the project group was formed by researchers from mathematics education, cognitive psychology, and cognitive neuro-science. One of the benefits of this constellation was that there were different methodologies that met and, as all involved were interested to learn from each other, there has been both interesting discussion and education during the design processes. Within the studies that this thesis is based upon there are many different methods used and the following sections will address them and connect them to both mathematics and the object of study in each study. As a member of the project group my work has comprised task design, experiment design, data collection, data analysis, and writing. The task- and experiment design phases have taken quite a lot of time during the start of the project and later on, data collection and analysis took their time as well. In the beginning of the project extensive piloting was done, since the interventions (and partially also the analysis method) were completely new, to secure that the students were able to solve the tasks and that the tasks did promote the desired reasoning. Therefore, the articles were not submitted until later on in my doctoral studies. In the following sections the design process will be elaborated to describe the different development stages involved.

5.1 Task design

In all studies mathematics tasks are important. In three of the studies (1, 2, and 3) the tasks were designed by researchers to be as tightly connected to specific reasoning types as possible. This was important since these studies focused the outcome of specific reasoning. This will be addressed further in section 5.1.1. In the last study the tasks we study as part of the students reasoning, are designed by the teachers as part of their ordinary preparation for the lessons. These tasks were not as strictly formulated as the tasks in study 1-3 and therefore not as easily distinguished as to what type of reasoning they would promote. The teacher-designed tasks are discussed further in section 5.1.2.

5.1.1 Design by researchers

The tasks in studies 1-3 were designed by researchers to promote either AR or CMR. The goal was that both these versions of the tasks should have the same target knowledge (i.e., a solution method in the shape of a formula). The AR-tasks would need to explicitly give the solution method while the CMR-tasks required the students to be able to construct the solution method by themselves. To increase the likelihood of this we designed the CMR-tasks with three sub-tasks as elaborated below. The design would be as similar as possible to eliminate that eventual layout differences would influence the results.

The design of the tasks was an extensive process that took place during a couple of years where a lot of different tasks were tried out and adjusted to be tried again, through several pilot-studies. The main reasons for the extended design process was the need to find tasks that a) could reach the same target knowledge either via AR or CMR, b) were not familiar to the students so that they already knew the solution method in advance c) not too difficult to solve by CMR and d) not so easy that solving by CMR provided no challenge. This left us a comparatively small window of implementation where the tasks were just hard enough to be solved by CMR, without the students falling into an AR-mode of reasoning because of familiarity with the tasks.

The first pilot-study was a think-aloud study with four students, where we video-recorded their work. The analysis indicated that our hypothesis concerning the importance of CMR held but also that the initial tasks were to extensive to be used in a large-scale study. Therefore, the tasks were adjusted and new tasks were constructed for a second pilot-study. This time two classes were involved and the students solved several multiple-choice tasks individually on a computer, with an observer seated next to them. The task solving process was recorded and after all tasks were solved the observer and the student went through the recording and discussed difficulties and the choices made by the students[4]. Analysis showed that some of the tasks were not suitable. These tasks were based on fictitious mathematics which did not engage the students. The third pilot-study tested some of the old tasks and some new tasks and this time the students (two classes) both practiced and were tested with a computer where the software also recorded their answers. The data was purely quantitative and comprised answers and solution times. After this trial, smaller changes were made to the tasks and instruction and then we regarded the tasks ready for the larger data-collection.

Basically, students that are solving AR-tasks are presented with a formula and the task leaves it up to the students to decide whether or not to think about the mathematical properties behind the formula (Figure 4). It is not necessary to do so to solve the task but it is possible. During the first two pilot-trials we observed a few AR-students that applied CMR in the first couple of AR-tasks but this was not common practice. One of these students explained afterwards that he wanted to check if we tried to fool him with an erroneous formula, but when the first formula checked out he trusted the tasks and continued without controlling the properties.

[4] Some results from this pilot-study was analyzed and reported in (Liljekvist, Lithner, Norqvist, & Jonsson, 2014).

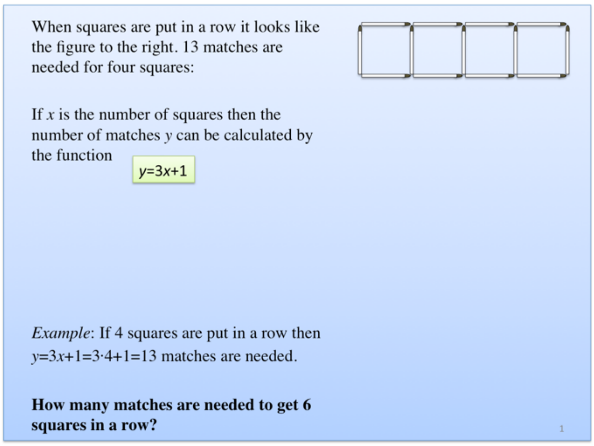

Figure 4: Example of an AR-task

In one of the studies (study 2) XAR-tasks were used. They were constructed by reuse of the AR-tasks but with an additional explanation of why the given solution method works. In some cases, an extra picture was added to clarify the explanation further (see Figure 5). The given explanations where controlled by four experienced teachers before the study to see if the explanations were reasonable in comparison to explanations in textbooks.

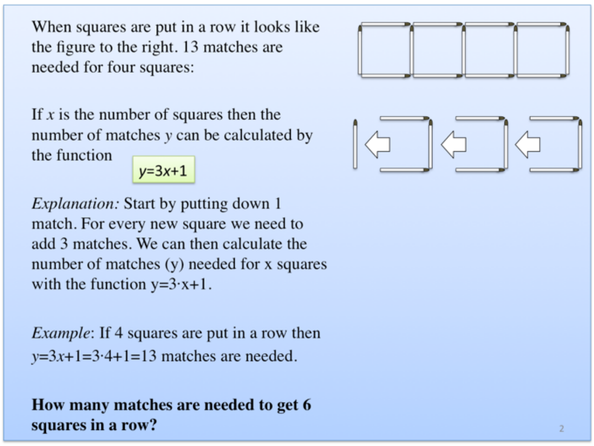

Figure 5: Example of a XAR-task

Students that solve CMR-tasks will not have a choice. They are forced to consider the mathematical properties in order to be able to solve the tasks (unless they are guessing, and strict guessing can almost never yield a correct answer). The tasks were designed to give them a first task that could be solved by observing and mentally extending the accompanying figure a few steps, a second task that forced them to consider a more generalized idea and, a third task that asked for a formalized algebraic expression (Figure 6).

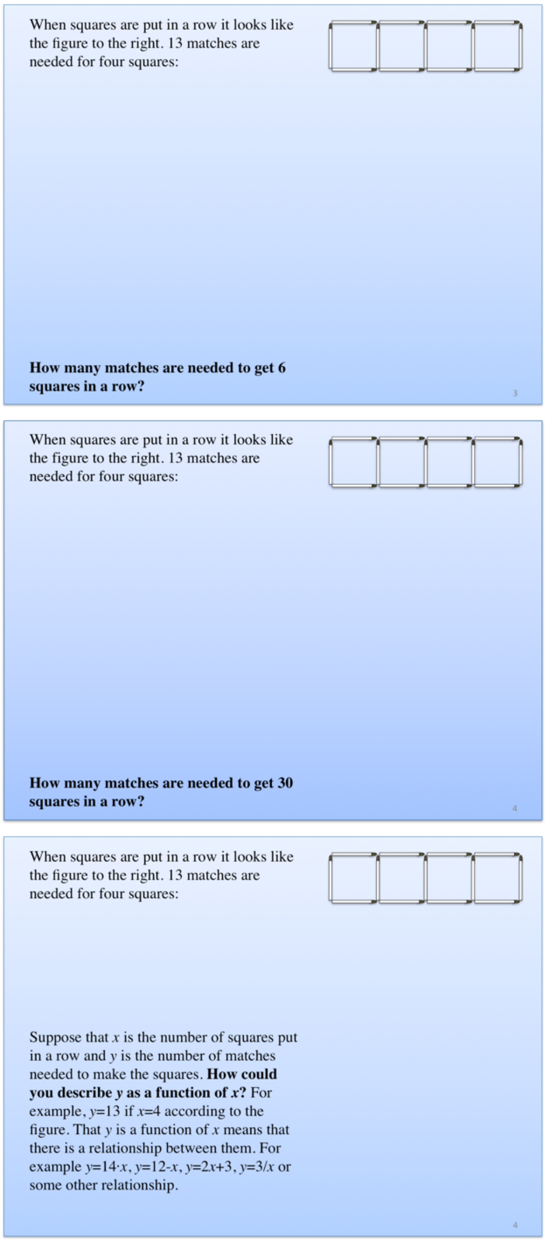

Figure 6: Example of a CMR-task (top), second CMR-task (middle), and last CMR-task (bottom).

5.1.2 *Authentic, design by teachers*

In study 4 we observed students' work with authentic teacher made tasks[5]. These tasks were not primarily designed to capture all aspects of CMR but were rather chosen as tasks with potential for CMR from a larger number of tasks that the teacher planned to use. We judged the task to have potential for CMR based on the novelty aspect in mind.

Two of the tasks were problem solving tasks with geometrical focus (i.e., optimizing volume of a cylinder and calculating area of a trapezoid). Both tasks had a clear goal and a given start, but little instruction on how to reach the goal. Both these tasks were classified as CMR-tasks. The teacher explained that the purpose of the task was for the students to practice problem solving. The other task in this study had clear instructions on how to proceed until the last subtask which had less instruction and was judged to have CMR-potential. The teachers ambition with the task was to let the students build on previous knowledge about differentiation to find a new method of differentiation connected to rational functions (i.e., the quotient rule).

5.2 Data collection

The included studies report on three experimental studies (study 1-3) and one observational study (study 4). In the following sections I will describe the methods used in these four studies to present similarities and differences between them regarding research design and sample. All participants in all studies were given and signed a written informed consent.

5.2.1 *Behavioral-experiments*

The experiment design in two of the studies (study 1 and 2) addressed differences in learning depending on the type of reasoning that the practice tasks promoted (i.e., AR, CMR, and in study 2 also XAR). In both experiments the sample comprised natural-science students from Swedish upper-secondary school (16-17 year olds). In study 1 we had 91 participants and in study 2 there were 104 participants, which we met for three sessions that took about 45 minutes each.

During the first session the students took two cognitive tests, Raven's Progressive Matrices and Operation Span, to measure fluid intelligence and working memory respectively. The students also provided background information (i.e., mathematics grade, age, and gender). These measures were then used to match the participants into two (study 1) or three (study 2) similar groups for the following session.

The second session consisted of a practice session where the students solved practice tasks via a computer program. The software saved their

[5] See study 4 for the complete tasks.

progress (i.e., solution time and answers) on a server. Within each of the practice groups students solved either AR-, CMR-, or XAR[6]-tasks. The AR- and XAR-groups solved a total of 70 sub-tasks while the CMR-group solved 42 sub-tasks, due to the additional time needed to complete the CMR-tasks. After practice the students had encountered 14 different tasks with different solution methods that we later tested for during the last session.

A week after the second session we met the students for the last time while they took the test. They were tested on all 14 solution methods with three test-tasks for each method. The first test-task explicitly asked them for the algebraic expression they used during practice. The time on this task was restricted to 30 seconds to restrict (re)construction of forgotten information. We hypothesized that the students that did not remember the formula could perhaps remember the solution idea they used during practice, therefore the second test-task asked for a numerical answer and was also restricted to 30 seconds for the same reasons as the first test-task. The third test-task asked for the same numerical answer as the second but with no time restrictions to allow for eventual (re)construction. During the test-session the software again recorded solution times and answers.

The AR-practice and the test tasks were designed to be as similar to textbook tasks and teacher made tests as possible, asking for solution methods that were practiced during the second session and for the use of these methods. This was done to try to ensure that the test did not especially benefit the CMR-students. A follow-up study have also shown that the reason that CMR seems to be more effective is not connected to any similarity between practice and test tasks (Jonsson et al., 2016).

5.2.2 fMRI-experiment

The fMRI-study (paper 3) focused on the eventual difference in brain activity during the test, after practicing via AR or CMR. In this study 40 of the 72 participants were upper-secondary students (18-20 yo) while the rest were first year university students (18-22 yo). Practice was again made at a computer that recorded practice data. The practice groups were matched into two similar groups based on mathematics grade, gender, and their score on Raven's matrices. The test session was done individually in an fMRI-scanner which put some restrictions on the experiment design.

As the fMRI-scanner registers all brain activity it would be very hard to distinguish the mathematics from eventual activity pertaining motoric functions, speech etc. Hence the test was done with multiple choice questions where the participant left his/her answers via a response-pad where each finger corresponded to an answer. Because of this the first test-task (i.e., recalling the formula) was removed and the second and third test-tasks were

[6] Only used in the study of paper 2.

adjusted to different numerical questions with the same time restriction (30 seconds to read the task + 6 seconds to choose an answer). After each test-task there was a baseline task to control for perception, attention, and reading. This task asked the participants to identify if there were any spelling errors in a text with the same layout as the mathematics tasks. This base-line task was then used to single out the mathematics specific processes from for example, reading and visual processing.

5.2.3 Observations

The observational study (paper 4) had a completely different approach. We started out by constructing an observational instrument (i.e., a document that let us structure our notes). This was piloted in an upper-secondary class followed by a rudimentary analysis process which led us to reconsider some parts of the observational instrument. We also realized the need for pre- and post-interviews with the teacher and post-interviews with the student groups that we observed. We also added a curriculum reading log in which the teacher elaborated on the materials used in preparation and what would be used during class. Another pilot study was made to try out the new tools, also in an upper-secondary class. This time we saw that the tools worked out so we went on by trying to find teachers that wanted to participate in the study.

We visited four teachers at a Swedish upper-secondary school. The teachers all agreed to participate but since only two of them had planned to use tasks with potential for CMR during their lessons we included these two teachers in the study. Within each lesson we randomly selected two student groups to observe while they solved the mentioned tasks. This was done to be able to have a closer look at how the reasoning sequences unfolded and how their progression through the task was connected to their reasoning.

We used observational notes, audio-recordings, and pictures to capture the discussion and work of the students. We also gathered information on the class and the intention for the lessons from the teachers and had a post interview with the students to entangle some critical moments that we observed. Altogether we observed four student groups (three pairs and one triplet).

5.3 Methods of analyses

The method of analysis differed a bit between the four studies since study 1-3 used quantitative data and study 4 relied on qualitative data. In the following sections I will describe the analysis methods and how they differed from each other.

5.3.1 Statistical

All results in studies 1-3 were quantitative and therefore analyzed statistically. Study 1 and 2 only used behavioral data and this was quantified by the software that we used to collect it. The same was true for study 3 but here additional fMRI-data was analyzed. The behavioral data was first scanned for outliers. After this a statistical analysis was chosen and performed depending on the nature of the data. In study 1 we chose to perform an analysis of variance with practice group (AR or CMR) as the fixed factor and the test score as the dependent variable. We also performed a linear regression analyses to evaluate the impact of the different measured factors (practice result, mathematics grade, cognitive index and gender) on the test score. Study 2 was analyzed with a multiple analysis of co-variance since there were two dependent variables (practice score and test score) and three practice groups, and since the matching of cognitive proficiency between the groups got disrupted by drop-outs it was controlled for as a co-variate. There were also linear regression analyses made that evaluated the predictors of the test score, similarly to the first study. In the third study the behavioral data was analyzed with an analysis of variance where the dependent variable was evaluated with practice group as the fixed factor. The fMRI-data was analyzed in several steps to sort out the activation depending on the practice condition. First the fMRI-images were adjusted for time and spatial differences, secondly the mathematics condition was compared to a reading condition to delimit a mathematics network and finally the activity within the mathematics network was analyzed to distinguish any differences between the two practice groups. In study 1 and 2 all statistical analyzes were conducted in SPSS[7] and in study 3 the analysis was done in SPM8[8].

5.3.2 Qualitative analysis

As the data in study 4 was of the qualitative kind (e.g., transcripts, notes, and pictures) we needed to find a different way to analyze it and this took quite some time. The audio recordings were transcribed and we started to identify moments when the reasoning sequence took a new direction. These vertices were then categorized in regards to the reason for the change in solution method. We also controlled if the next vertex indicated a progress in the task solving process to be able to distinguish if the solution method brought the students closer to the solution or not. The edges between the vertices were then analyzed according to which type of reasoning dominated each edge.

When we had the vertices, edges, and all classifications ready, we visualized the complete reasoning sequence in a graph where the x-axis represented the turns in the students' discussion and the y-axis represented progression

[7] Statistical Package for the Social Sciences, version 22 and 23.
[8] Statistical Parametric Mapping, version 8, which is an add-on to Matlab.

towards the answer. The graphs were then analyzed to see if we could find patters and co-occurrences that were connected to the task design (AR or CMR), the reasoning of the student group, and the student groups motivation and persistence. A much more elaborate description of the methodology can be found in the included manuscript (i.e., study 4).

6 Summary of the articles/Result

In this section I will give a short summary of each of the included studies. I will also present an additional result from studies 1-3 that was not reported in the articles. The additional result concerns information about the practice sessions and can serve as a comparison to the test results that are reported in the included studies.

6.1 Study 1 – Learning mathematics through algorithmic and creative reasoning

The aim of this study was to investigate the learning effects of practicing mathematical tasks through AR and CMR on task-solving performance.

We let 131 upper secondary students from the natural-science program in Sweden practice on solution methods to 14 different tasks, either by AR or CMR. The two groups were matched on cognitive prerequisites, mathematics grade and gender. The AR-group got five practice tasks where a solution method (i.e., a formula) was presented and this was repeated for all 14 solution methods. The CMR-group got three tasks with increasing difficulty for each solution method where the last task was to construct a mathematical formula that described the sought after relation. One week later we gave both groups the same test on the 14 solution methods. The test asked three questions for each solution method. The first asked for the formula and was restricted to 30 seconds. This was to ensure that they had to remember the formula and not have time to (re)construct it. The second question asked for a numerical answer and was also restricted to 30 seconds, again to ensure that no (re)construction would take place. The third question was the same as the second but now with a 300 second time limit to give enough time for eventual (re)construction of a forgotten formula.

After excluding participants from the sample due to attrition and in a few cases to prevent ceiling effects we were left with 91 students (48 AR and 43 CMR). The result showed that the CMR-group significantly outperformed the AR-group on the test, both on the composite level (see Figure 7) and on the three different tasks. A closer study of the data also revealed that high cognitive capacity was more important for the AR-group to perform well during the post-test than for the CMR-group. This is contrary to common beliefs as many regard problem solving and similar tasks as “only for the high achievers”.

In the discussion we try to find reasons to why CMR seems to be more efficient. One reason that comes up is that the CMR-group has to struggle with mathematical properties of the task to find a solution to do this. The AR-students get to see and use the correct formula repeatedly but do not have to struggle at all. The positive struggle that the CMR-group is subjected to might be one reason to why they perform better. The theory of didactical situations also suggest that you learn better by constructing the solution compared to imitating a given method. Thus, CMR will produce better opportunities to understand and learn mathematics.

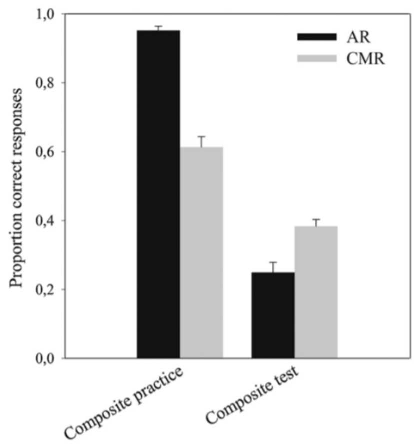

Figure 7: Practice and test scores on composite level as retrieved from Jonsson et al. (2014).

6.2 Study 2 – The affect of explanations on mathematical reasoning tasks

The aim of this study was to see if the addition of an explanation of the mathematical principle behind the formula in the AR-task would enhance the efficiency of the AR-tasks.

104 upper secondary students from the natural-science program in Sweden were recruited to participate in the study. The method was similar to study 1 and the AR and CMR-tasks were the same but another group, XAR, was added. The XAR-group got the same tasks as the AR-group but with an additional explanation that briefly but carefully described the principle behind the given formula. In other words, while the AR-tasks included information about *how* to solve the task the XAR-tasks also included mathematical arguments clarifying *why* the suggested method was correct.

All three groups were matched on cognitive capacity (i.e., Ravens matrices and operation span), mathematics grade and gender. One week after the practice session the students took the same test as in the previous study. The results indicate that the added explanation did not give any significant effect on performance compared to the AR-group. The tendency from the earlier study that the CMR-group outperformed the AR-group was still there but with fewer participants it did not show up as significant. Compared to the XAR-group though, the CMR-group performed significantly better (see Figure 8).

Cognitive capacity was still more important for the AR-groups than for the CMR-group but the added explanation seemed to lessen this effect slightly for the XAR-group.

In the discussion the non-effect from the added explanation is discussed in terms of how Brousseau's theory of Didactical Situations (1997) actually predicts this result. The importance of struggle again comes into play and the eventual lack of engagement in the explanation is also hypothesized to be an explanation.

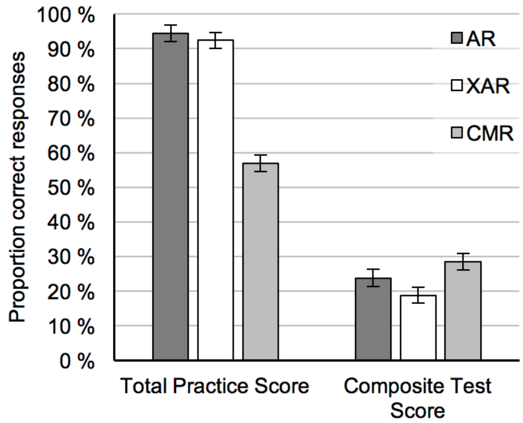

Figure 8: Practice and test scores as retrieved from Norqvist (2016).

6.3 Study 3 – Learning mathematics without a suggested method: Durable effects on performance and brain activity

The aim of this study was to replicate parts of the first study and add the perspective of brain activity. Hypothetically, the performance would be similar to the first study but since the test was made in an fMRI-scanner with multiple-choice questions we could not be sure. Another hypothesis was that the CMR-group would show less activity in the left angular gyrus since study 1 had shown that CMR-practice yielded a better recollection of the solution methods than AR-practice.

73 students, 40 from the third year at upper secondary school and 33 first year engineering students were recruited to participate. All were right-handed and had normal or corrected-to-normal vision. The participants were divided into two matched groups, AR and CMR, based on cognitive prerequisites, mathematics grade and gender. The students practiced on nine different task types with different solution methods and the practice method was the same as in study 1. Six days after practice the students took a test while in an fMRI-

scanner. This test consisted of one multiple-choice for each solution method. Between each mathematical test task there was a baseline task, checking for spelling errors, that served as a contrast in the later analysis.

The result showed that the CMR-group outperformed the AR-group on the test (see Figure 9A). It was also apparent that the AR-group had a higher activation of the left angular gyrus (see Figure 9B) and the left precentral cortex. The results also show that the right superior parietal cortex is important for mathematical performance.

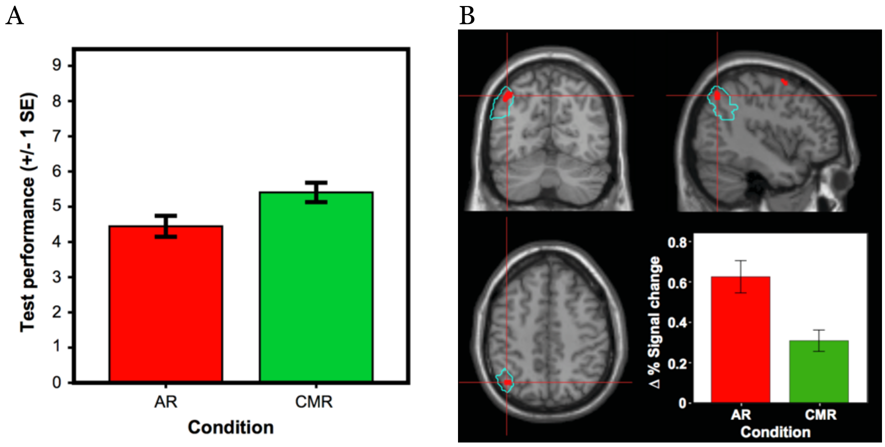

Figure 9: (A) Mathematics test scores for the two groups. (B) Difference in activation between groups in Left Angular Gyrus, as retrieved from Wirebring et al. (2015)

We concluded that the CMR-group again outperformed the AR-group on the post test. We also discussed the two indicated areas where the AR-group had a higher activation rate than the CMR-group. The higher activity in the angular gyrus indicates that the AR-students have to work harder to retrieve the formula from memory. The difference in the precentral cortex indicates that there is more stress on working memory in the AR-group.

6.4 Study 4 – Unraveling students' reasoning: analyzing small-group discussions during task solving

The aim of this study was to examine students' reasoning to see how the reasoning sequence would unfold in actual classroom situations. We were also interested in how students' reasoning would influence their progression towards a solution.

We visited two classrooms in an upper secondary school and observed two student groups in each classroom for the time it took them to complete a task, constructed and presented to them by the teacher. One of the four student groups did not complete the given task during the observed lesson. Initial analysis showed that there were two interesting dimensions to regard, group characteristics (i.e., the student-group's motivation and persistence) and task design (i.e., AR or CMR). After transcribing the audio-recordings we have segmented them into sections by utilizing Lithner's (2008) framework of mathematical reasoning. A moment when the students' reasoning took a new trajectory was called a vertex and the segment between two such vertices was called an edge (Lithner, 2008). After the classification we compared the vertices and determined how they compared to each other regarding progress towards the answer (i.e., return to previous ideas – step down, no progression – horizontal edge, or progress – rising edge or step up). The edges were then categorized according to the students' reasoning (i.e., either CMR or AR). We then visualized the students' reasoning in graphs (see Figure 10) and analyzed the patterns, the amount of rising edges and types of reasoning, as well as how the group characteristics and task design would influence reasoning and progress.

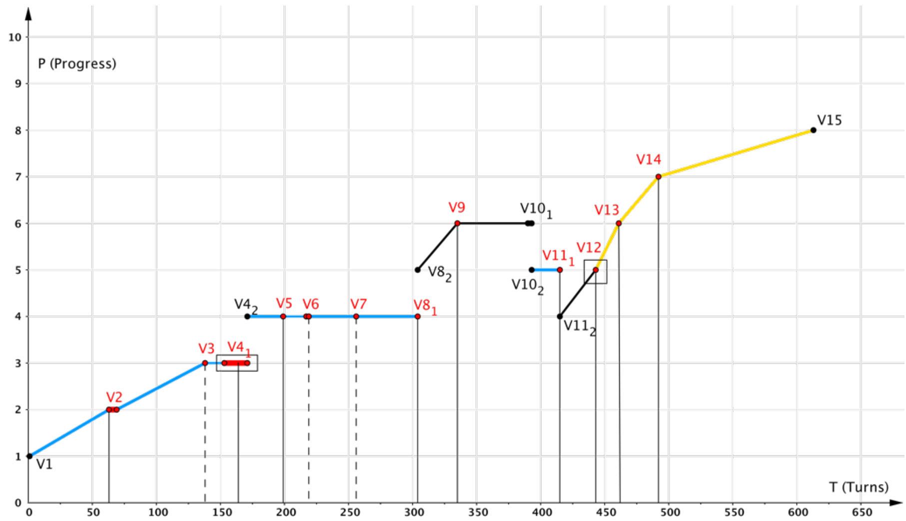

Figure 10: Example of visualization of reasoning sequence as retrieved from Van Steenbrugge & Norqvist (2016). Blue indicates AR, Yellow CMR, and Black un-characterized reasoning (all other markings are explained in the included paper (Study 4)).

The result showed that task design is important for which reasoning the students will use. Although an AR-task does not exclude CMR, it only occurs in our data if the students have difficulties and strive to handle them by themselves. We also observed that group characteristics were important for the chosen reasoning type. Student groups that were less motivated and less persistent were more prone to giving up and using AR than the more motivated and more persistent student groups.

6.5 Additional result

In study 1-3 we also gathered data on the practice session. However, we did not report or discuss this extensively in the studies. What is obvious is that CMR-practice is much more taxing on the cognitive abilities than AR-practice. This becomes evident if we observe which variables that predicts the practice result. For AR-practice none of the included variables (i.e., gender, cognitive proficiency index, mathematics grade, practice time) are predictive of the practice result while the CMR-practice result is highly dependent on cognitive capacity and mathematics grade (see Table 1). This is consistent throughout study 1 and 2 but not apparent in study 3, maybe because of the smaller sample[9]. This is almost the opposite to the result from the test where the AR-groups test result was predicted by cognitive measures and to some degree mathematics grade and gender while the CMR-groups test result was predicted by their practice result (see Table 2). In the tables below the XAR-condition from study 2 is merged with the AR-condition since they performed similarly.

Table 1

Regression Analysis Summary For Variables Predicting Practice Result

	AR			CMR		
Variables	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*
Study 1						
Cognitive index	.031	.014	.349*	.079	.047	.249
Mathematics grade	.004	.004	.151	.031	.011	.418**
Gender	.022	.023	.131	.060	.058	.135
Study 2						
Cognitive index	.008	.008	.120	.128	.033	.432***
Mathematics grade	.006	.002	.375**	.041	.008	.555***
Gender	.013	.012	.120	-.068	.053	-.144
Study 3†						
Raven's matrices	-.895	1.825	-.129	2.088	3.847	.126
Mathematics grade	-.344	.343	-.272	.986	.801	.288
Gender	-4.794	3.408	-.353	3.063	7.789	.089

**p*<.001, ***p*<.01, ****p*<.05.

† includes one of the sample groups (students).

[9] Only data from one of the sample groups were available from study 3 at the time of writing this thesis.

Table 2

Regression Analysis Summary For Variables Predicting Test Result

	AR			CMR		
Variables	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*
Study 1						
Cognitive index	.151	.047	.492**	-.003	.039	-.010
Mathematics grade	.013	.013	.155	.004	.010	.046
Practice score	-.027	.472	-.008	.897	.131	.811***
Gender	-.005	.072	-.009	-.034	.048	-.069
Study 2						
Cognitive index	.059	.018	.282**	.032	.024	.172
Mathematics grade	.030	.005	.588***	.010	.007	.219
Practice score	-.060	.286	-.019	.382	.110	.597**
Gender	.117	.028	.342***	.018	.032	.060
Study 3†						
Raven's matrices	9.046	3.806	.472*	2.071	4.074	.102
Mathematics grade	1.726	.733	.494*	.830	.872	.198
Practice score	1.087	.534	.534	.607	.253	.497*
Gender	17.363	7.502	.463*	2.825	8.159	.067

**p*<.05. ***p*<.01. ****p*<.001.
† includes one of the sample groups (students).

7 Discussion

In the included studies there are some coherence in results that could be noted. *First*, studies 1-3 show that CMR is more effective than (X)AR[10] regarding both memory retrieval and re-construction of practiced solution methods. *Second*, studies 1-3 also show that practicing by (X)AR gives rise to higher taxation on cognitive abilities during the test situation than CMR-practice does. This is confirmed by the neuro-cognitive data in study 3, which shows that AR is more neurologically demanding during the test situation. *Third*, the additional data shows that CMR-practice will however tax cognitive proficiency during practice while (X)AR-practice does not to the same extent. *Fourth*, study 4 confirms that if CMR is preferred, the task design is very important. Even though CMR is not excluded when (X)AR-information is available, it is much less likely to occur, especially if the students lack motivation and persistence. Bear in mind that the discussion is limited to the results in the four included studies and the additional result in the previous chapter, hence the word "shows" is used to mean in the context of these studies.

During the remainder of this chapter these results will be discussed in relation to the research questions (Sections 7.1-7.3) as well as in relation to limitations, generalizability, and implications for further studies and teaching practice (Section 7.4)

7.1 How will the task design influence students' solutions process, mathematical reasoning, and brain activity?

From a theoretical point of view both Brousseau (1997) and Lithner (2008) argue that tasks that contain prescriptive solution methods will most likely be solved with said methods. And according to Brousseau (1997), this will not lead to the creation of any new knowledge. Tasks should instead be designed to promote the construction of new knowledge by emphasizing important concepts or mathematical properties and at the same time not give away the entire solution method. This is basically what happens in tasks with potential for CMR. In study 4, it was clear that students preferred to use AR as far as possible and as long as the task was designed with given methods these were used. The lack of effect from the provided explanations in study 2 can also be seen as a sign of that students prefer clean AR and do not bother with redundant information. However, in a few instances in study 4 students did use CMR when solving AR-tasks. Common for these instances was that the students had difficulties in solving the task with the given method, which led them to discuss the mathematical properties that was important and decide on how to move on.

---

10 (X)AR denotes both AR and XAR

The student-group's motivation and persistence also seemed to be important factors for CMR to occur. One of the groups in study 4 was less motivated and less persistent than the others and when they met a CMR-task they never chose to reason creatively. Their reasoning was based upon formerly known methods and when they did not work the students started exploring methods by random. This could also be due to what Bjork and Bjork (2011) calls undesirable difficulties (e.g., an unbridgeable gap between prior knowledge and task requirements). The extra effort that can be so beneficial for mathematics learning (e.g., Bjork & Bjork, 2011; Hiebert & Grouws, 2007) and which is needed to perform CMR (Jonsson et al., 2016) will of course require tasks that are within reach for the students' resources and heuristics (Schoenfeld, 1985).

By design, tasks that focus the use of a given solution method will promote AR. CMR is not excluded from this type of task but is much less likely to occur than AR. Providing an explanation as was done in study 2 (i.e., an explanation to why a solution method works and is valid), does not seem to increase test scores, although the explanation is focusing the important mathematical properties. The explanation will not induce the productive struggle that CMR does by definition, and this might be why XAR-information does not improve test results compared to AR.

CMR-practice also seems, according to the results in study 3, to contribute to less effortful memory retrieval and less brain activity during the test-situation. The students that practiced by AR had a significantly higher activation in a part of the brain that is connected to verbal memory retrieval, the left Angular Gyrus (e.g., Delazer et al., 2003; Ischebeck et al., 2007; Seghier, 2013). This would indicate that the solution methods were encoded as a string of words, much like an automated multiplication table, rather than as mathematical relations or properties. As will be discussed in the next section, this lack in retention of mathematical relations or properties can affect the test situation by inducing higher strain on cognitive abilities. That our results showed no areas where brain activity was higher for the group that practiced with CMR-tasks than with the AR-tasks confirms that retrieval and eventual reconstruction during the test was less effortful for the CMR-group.

Although this thesis indicates that promoting CMR is preferable, AR can sometimes be the appropriate choice. There are some concepts or properties that are difficult to grasp or construct for students until later on in their schooling, for example that if you multiply two negative numbers the answer will be positive. There is also need for automatization of procedures to be able to focus one's cognitive abilities to the difficult parts of a problem. It would be impractical if every little calculation would need the full attention of the brain's mathematics network. The problem is therefore not that AR occurs or that rote-learning is common, but that these types of non-reflective learning is too common. When learning occurs without connection to mathematical properties or concepts it becomes fragmented, hard to remember, and as Boaler (1998) saw irrelevant.

7.2 How will students' cognitive variation affect their solution rate and does task design matter?

In studies 1-3 cognition plays a significant role in the test-situation, and to a larger extent if the students have practiced by AR or XAR. The CMR-students do not seem to be as influenced by cognitive proficiency during the test situation, but rather by the extent of how well they have performed during practice. This conclusion was also confirmed by fMRI in study 3, where AR showed to be more demanding during the test situation. However, as the additional results show, CMR is taxing high on cognition during practice while (X)AR does not. So, in a sense, test-results for CMR-students are also a dependent on cognitive abilities, although in an indirect way.

In our studies students that practiced by CMR performed better than the (X)AR-students during the test. One could argue that this difference is due to a performance boost of the capable and high achieving students that would benefit most from creative reasoning. However, in study 1 the largest difference in test result, between the two practice groups, can be found in the lowest cognitive tertile.

The struggle provided by the CMR-tasks could be one reason to why these tasks render higher test results, even though many students did not solve all practice tasks. As was suggested by Hiebert and Grouws (2007), discussed in studies 1-3, and later on supported by Jonsson et al. (2016), effortful struggle with important mathematical concepts and properties is essential for mathematical understanding. Jonsson et al. (2016) also concluded that task design is vital for what level of effort students will allocate to the solving process. Furthermore, Bjork and Bjork (2011) points out that practicing with desirable difficulties does not only increase test results but also improve long-turn retention and transfer of skills. The importance of devoting effort to a task was also argued for by Brousseau (1997) in the theory of didactical situations, where he suggests that the teacher should delegate the responsibility of the task solution to the students within the a-didactical situation. Thus, there are both theoretical and empirical arguments that CMR may be considered in the teaching practice.

7.3 How can these results influence teaching practice?

The result that (X)AR-practice puts less strain on cognition than CMR-practice, and that the opposite is true during tests is maybe not surprising but can be rather alarming. If practice in the classroom does not prepare the students to cope with the coming test situation, then what we test is not whether or not students have learned mathematics, but rather the students' cognitive abilities. This could be deceitful for both students and teachers. If the teacher observes that the students can solve most tasks they may believe that the students are learning. If the difference between the practice- and test scores from study 1-3 would be seen as indications of what could happen in a classroom, there could be a significant drop in performance from practice to test. Of course the experiment scores are somewhat an extreme case of this but, from discussions with teachers from lower secondary school to university, it does not seem uncommon that they meet students that are either underestimating or overestimating their abilities, especially when it comes to clearing or failing tests. Bear in mind that the students practiced by themselves without any help from teacher or peers. This means that all difficulties that the students met during their practice had to be either overcome by themselves or left behind. What the result would be if the students would have received help is a question left to another study, but from the results of Stein et al. (1996) we can assume that the results might be even better for the CMR-group, provided that the help was of good quality (i.e., addressed conceptual hurdles) that did not collapse the didactical situation.

Since much lesson time is devoted to solving tasks (Mullis et al., 2012; Wakefield, 2006), some of these problems could be solved by presenting the students with more well designed and mathematically challenging tasks that focus creative reasoning and mathematical properties rather than unconnected procedures. Constructing and designing new tasks can however be tedious work, but many textbook tasks have all components needed. The problem is often that there is too much information for the tasks to have potential for CMR. The following example from an upper-secondary textbook can illustrate this:

> *422. In a rectangle the sides are 8 cm and 6 cm long. The long sides will be shortened by 30% and the short sides will be increased by 25%.*
> *a) How long are the sides in the new rectangle?*
> *b) What is the area of the new rectangle?*
> *c) What is the percentual difference between the areas of the old and new rectangles?*

To construct a task with higher potential for CMR we could just remove some of the explicit sub steps and thereby leave it up to the student to consider which steps that are needed to solve the tasks (provided that the student does not know any solution methods by heart). Hence, we would end up with this task:

> *422. In a rectangle the sides are 8 cm and 6 cm long. The long sides will be shortened by 30% and the short sides will be increased by 25%. What is the percentual difference between the areas of the old and new rectangles?*

We could even remove the lengths so that the students would have to make an assumption about how long the sides are, or try to make a more general solution.

By removing the prescriptive parts of the task it becomes more effortful and creative. Now each student can have an idea on how the task could be solved and a discussion between classmates could focus on the mathematical properties rather than on implementation of known procedures. The teacher also has the opportunity to easily adjust the tasks to individual students by removing more or less of the guiding steps, and by helping students with conceptual problems. This adjustment takes much less time than having to construct the task from scratch and though the original task probably will be solved faster there will be less struggle with important mathematics. This might be one way of utilizing the CMR-idea as a way to construct tasks that are more effortful for the students but still concerns the content that the textbook prescribes.

It is also necessary to consider the result that AR-practice (as discussed under 7.2) is more cognitively demanding during test than during practice. If students fulfill the learning goals by rote learning and if these learning goals were best assessed by giving the students cognitively demanding test tasks, all would be well. However, as many studies show, this is not the case. If we want students to become proficient in mathematics, rote learning alone will not be enough (e.g., Bjork & Bjork, 2011; Hiebert, 2003; Schoenfeld, 1985). Personally I do not think that this is fair to the students. If students are supposed to solve mathematical problems and reason creatively they must get a chance to practice these competencies as well as the necessary procedures.

Studies 1-3 are made in an experimental setting and the classroom environment is of course much more complex. However, the studies of this thesis combined with many other studies in mathematics education give us reasonable evidence that task design and students reasoning are important aspects to consider when discussing students' mathematical learning.

In classroom practice the teacher could address the question about task appropriateness by choosing tasks that will fit into the curriculum and presenting them at a moment where the students can solve the task but do not know a simple procedure to do so. The teacher also has the possibility to present new content in such a way that the students has to consider mathematical properties instead of being served pre-defined procedures.

7.4 Limitations and implications for further studies

The included studies give indications on how to take the idea of creative reasoning into a classroom study. The involved teachers would have to be either educated in the different types of reasoning to be able to work as creatively as possible, or scripted as to what help and which response they would give the students. The teacher introductions and the tasks that the class would be working with have to a larger extent to be designed with CMR in mind to engage as many students as possible in this type of reasoning. It would also be good for a study if it could encompass a complete section of the subject knowledge that the class should learn that year (e.g., percentages or linear equations).

Designing tasks that would fit the rather small window comprising what the students know and what they are about to learn has also been a difficulty that we have encountered. As have been addressed earlier, CMR-tasks need to be solvable but not by any previously known solution method, and this have been a challenge for the research group during the task design process since we have little direct evidence of what the students' prior knowledge were. We chose not to test the students' prior knowledge since a test of the particular solution methods in the experiment would have influenced both the practice and post-test. Therefore, we decided to rely on more indirect sources (i.e., which tasks the students had met in the textbook and their mathematics grade) as a measure of prior knowledge. It would have been of interest for the tasks to address topics in the current curriculum and as we only visited the classrooms thrice it was hard to find tasks that fit.

It would also be interesting to meet with textbook authors and discuss the development and testing of a new type of textbook that could put more focus on creative reasoning from the beginning of each chapter. Such a textbook could be evaluated by comparing observations of students' work with it and with one of the common textbooks, as well as with pre- and post-tests of knowledge. However, a new textbook alone is insufficient unless teaching is adopted to match the intentions of the new material.

There are also many adjustments that could be made to the existing studies to examine how, for example, repeated practice, delayed tests, or different student groups would influence the results. In fact, one such study has already been made by one of our pre-service teachers, where the object was to see if a more heterogeneous and slightly younger group of students would yield the same result as study 1. The result showed that the same pattern occurs during practice and also the importance of the need to think through the tasks for this group of students (Wikman, 2015).

Since studies 1-3 were pseudo-experiments, we tried to reduce complexity as much as possible to be able to study the influence of the tasks rather that of something else. This meant that information to the students were scripted and quite vague when it came to the the purpose of the different sessions. The fact that the students did not know that they were going to be tested could also have been important for their test performance. Students could have concentrated harder to remember the solution methods and maybe even tried to rehearse them before the test. Then again, it was not a high-stake test so the students might not have bothered to engage in rehearsing since they have plenty of other school-related and higher-stake-tasks to attend to.

Letting students talk to each other while solving the tasks could also be interesting to study. From a result perspective this could increase the CMR-students practice performance and according to the results in study 1-3, the test scores. On the other hand, the help the students give each other could be of an AR-type which might lessen the focus on important mathematical properties and hence decrease the test-scores. It would also be interesting to observe the student discussions, the questions and help they give each other, both in the CMR-group and the AR-group, to notice if the task design will be important for the quality of these discussions.

Another interesting but maybe impractical study would be to make an fMRI-study of the practice situation. It would be very interesting to see how the CMR-practice would compare to the AR-practice on a brain activation level. Alas, at the moment I have no inkling on how such a study would be designed or which parts of the practice situation that could be or should be observed. Since the fMRI-environment puts restraints on the experiment design a lengthy creative reasoning sequence would be hard to observe, with all the noise that would infiltrate the data. Maybe some parts of the reasoning process could be focused or maybe some sort of multiple choice questions could lessen the time needed to construct a solution method. Maybe the productive struggle could be isolated and observed but as I said, at this moment I am just speculating. This kind of study could however give us deeper insight into why CMR-practice seems to be so effective compared to AR-practice and which brain processes that are active in sustaining this difference.

7.5 Conclusion

The studies included in this thesis confirm that task design is important for what type of reasoning the students will engage in, and that the reasoning is subsequently important for what is learned, just as Brousseau (1997) and Lithner (2008) argued. The included studies also suggest that creative reasoning is not only beneficial for the high achievers but equally (if not more) fruitful for cognitively less proficient students. As a complement, the additional result indicate that CMR-practice is highly taxing on cognition while during the test practice performance becomes more important. Conversely, practicing by AR seems to be quite effortless while the test seems to place a higher demand on cognition. The results of study 4 also informed us that CMR is important for progression through demanding tasks and that the reasoning also is dependent on student characteristics, such as motivation and perseverance along with subject knowledge.

These findings are of importance for both teacher practice and textbook design. It is essential that students are well prepared for coming tests as well as for future life and none of these will benefit from knowing a lot of superficial procedures. Even if creative mathematical reasoning always can be used to understand concepts and mathematical properties, task solvers may rather use an available algorithm to solve a task as rapidly as possible. If students are taught mathematics mostly based upon becoming proficient in using different procedures, they may be even less likely to engage in creative reasoning if it is not required from the task (Schoenfeld, 1985). Therefore, including more creative mathematically founded reasoning tasks, and teaching that supports students' work with such tasks, is important if we want students to grow up to become mathematically literate.

8 References

Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417-423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. (1974). Working Memory. *Psychology of Learning and Motivation, 8*, 47-89. doi:10.1016/S0079-7421(08)60452-1

Bergqvist, T., & Lithner, J. (2012). Mathematical Reasoning in Teachers' Presentations. *Journal of Mathematical Behavior, 31*(2), 252-269.

Bergqvist, T., Lithner, J., & Sumpter, L. (2008). Upper Secondary Students' Task Reasoning. *International Journal of Mathematical Education in Science and Technology, 39*(1), 1-12.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society*. (pp. 56-64). New York, NY, US: Worth Publishers.

Blum, W., & Niss, M. (1991). Applied Mathematical Problem Solving, Modelling, Applications, and Links to Other Subjects--State, Trends and Issues in Mathematics Instruction. *Educational Studies in Mathematics, 22*(1), 37-68.

Boaler, J. (1998). Open and Closed Mathematics: Student Experiences and Understandings. *Journal for Research in Mathematics Education, 29*(1), 41-62.

Boesen, J., Helenius, O., Bergqvist, E., Bergqvist, T., Lithner, J., Palm, T., & Palmberg, B. (2013). Developing mathematical competence: From the intended to the enacted curriculum. *Journal of Mathematical Behavior, 33*, 72-87. doi:10.1016/j.jmathb.2013.10.001

Boris, H. (2003). Teachers' Mathematical Beliefs: A Review. *Mathematics Educator, 13*(2), 47-57.

Brousseau, G. (1997). *Theory of didactical situations in mathematics*. Dordrecht ; Boston: Kluwer Academic Publishers.

Brouwers, S. A., Van de Vijver, F. J. R., & Van Hemert, D. A. (2009). Variation in Raven's Progressive Matrices scores across time and place. *Learning and Individual Differences, 19*(3), 330-338. doi:10.1016/j.lindif.2008.10.006

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*(3), 205-228. doi:10.1080/87565640801982312

Chuderski, A. (2015). The broad factor of working memory is virtually isomorphic to fluid intelligence tested under time pressure. *Personality and Individual Differences, 85*, 98-104. doi:10.1016/j.paid.2015.04.046

De Smedt, B., Ansari, D., Grabner, R. H., Hannula, M. M., Schneider, M., & Verschaffel, L. (2010). Cognitive Neuroscience Meets Mathematics Education. *Educational Research Review, 5*(1), 97-105.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cogn Neuropsychol, 20*(3), 487-506. doi:10.1080/02643290244000239

Delazer, M., Domahs, F., Bartha, L., Brenneis, C., Lochy, A., Trieb, T., & Benke, T. (2003). Learning complex arithmetic - an fMRI study. *Cognitive Brain Research, 18*(1), 76-88. doi:10.1016/J.Cogbraines.2003.09.005

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest, 14*(1), 4-58. doi:10.1177/1529100612453266

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between Measures of Cattell-Horn-Carroll (CHC) Cognitive Abilities and Mathematics Achievement across the School-Age Years. *Psychology in the Schools, 40*(2), 155-171.

Freund, P. A., Holling, H., & Preckel, F. (2007). A Multivariate, Multilevel Analysis of the Relationship Between Cognitive Abilities and Scholastic Achievement. *Journal of Individual Differences, 28*(4), 188-197. doi:10.1027/1614-0001.28.4.188

Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., & Neuper, C. (2009). To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving. *Neuropsychologia, 47*(2), 604-608. doi:10.1016/J.Neuropsychologia.2008.10.013

Henningsen, M., & Stein, M. K. (1997). Mathematical Tasks and Student Cognition: Classroom-Based Factors That Support and Inhibit High-Level Mathematical Thinking and Reasoning. *Journal for Research in Mathematics Education, 28*(5), 524-549.

Hiebert, J. (2003). What research says about the NCTM standards. In J. Kilpatrick, W. G. Martin, D. Schifter, & National Council of Teachers of Mathematics. (Eds.), *A research companion to Principles and standards for school mathematics* (pp. x, 413 p.). Reston, VA: National Council of Teachers of Mathematics.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 371–404). Greenwich, CT: Information Age.

Houde, O., Rossi, S., Lubin, A., & Joliot, M. (2010). Mapping numerical processing, reading, and executive functions in the developing brain: an fMRI meta-analysis of 52 studies including 842 children. *Developmental Science, 13*(6), 876-885. doi:10.1111/J.1467-7687.2009.00938.X

Huettel, S. A. (2014). *Functional magnetic resonance imaging* (3. ed. ed.). Sunderland, Mass.: Sunderland, Mass. : Sinauer Associates.

Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., & Delazer, M. (2007). Imaging early practice effects in arithmetic. *Neuroimage, 36*(3), 993-1003. doi:10.1016/J.Neuroimage.2007.03.051

Jackson, K., Garrison, A., Wilson, J., Gibbons, L., & Shahan, E. (2013). Exploring relationships between setting up complex tasks and opportunities to learn in concluding whole-class discussions in middle-grades mathematics instruction. *Journal for Research in Mathematics Education, 44*(4), 646-682.

Johansson, H. (2015). *Mathematical Reasoning-In physics and real-life context.* (Doctoral Thesis), University of Gothenburg, Gothenburg.

Jonsson, B., Kulaksiz, Y. C., & Lithner, J. (2016). Creative and algorithmic mathematical reasoning: effects of transfer-appropriate processing and effortful struggle. *International Journal of Mathematical Education in Science and Technology*, 1-20. doi:10.1080/0020739X.2016.1192232

Jonsson, B., Norqvist, M., Liljekvist, Y., & Lithner, J. (2014). Learning mathematics through algorithmic and creative reasoning. *The Journal of Mathematical Behavior, 36*, 20-32.

Jäder, J., Lithner, J., & Sidenvall, J. (2014). Reasoning requirements in school mathematics textbooks: an analysis of books from 12 countries. *Submitted.*

Koenigs, M., Barbey, A. K., Postle, B. R., & Grafman, J. (2009). Superior Parietal Cortex Is Critical for the Manipulation of Information in Working Memory. *Journal of Neuroscience, 29*(47), 14980-14986. doi:10.1523/Jneurosci.3706-09.2009

Liljekvist, Y., Lithner, J., Norqvist, M., & Jonsson, B. (2014). Reasoning in alignment with the task design? Students actual handling of algorithmic and creative mathematical reasoning.

Lithner, J. (2000). Mathematical Reasoning in Task Solving. *Educational Studies in Mathematics, 41*(2), 165-190.

Lithner, J. (2003). Students' Mathematical Reasoning in University Textbook Exercises. *Educational Studies in Mathematics, 52*(1), 29-55.

Lithner, J. (2004). Mathematical reasoning in calculus textbook exercises. *The Journal of Mathematical Behavior, 23*(4), 405-427.

Lithner, J. (2008). A Research Framework for Creative and Imitative Reasoning. *Educational Studies in Mathematics, 67*(3), 255-276.

Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. International Association for the Evaluation of Educational Achievement. Herengracht 487, Amsterdam, 1017 BT, The Netherlands.

NCTM. (1989). *Curriculum and evaluation standards for school mathematics*: National Council of Teachers of Mathematics. Commission on Standards for School Mathematics.

NCTM. (2011). *Common core state standards for mathematics*: National Council of Teachers of Mathematics. Commission on Standards for School Mathematics.

Newton, D., & Newton, L. (2007). Could Elementary Mathematics Textbooks Help Give Attention to Reasons in the Classroom? *Educational Studies in Mathematics, 64*(1), 69-84. doi:10.1007/s10649-005-9015-z

Niss, M. (2007). Reactions on the state and trends in research on mathematics teaching and learning. From here to Utopia. . In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning : a project of the National Council of Teachers of Mathematics* (Vol. 2, pp. 1293-1312). Charlotte, NC: Information Age Pub.

Niss, M., & Jensen, T. H. (2002). *Kompetencer og matematiklæring: Idéer og inspiration til udvikling af matematikundervisning i Danmark* (Vol. 18): Undervisningsministeriet.

Norqvist, M. (2016). The affect of explanations on mathematical reasoning tasks. Manuscript in preparation.

OECD. (2015). *Education at a Glance 2015*: OECD Publishing.

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The Precursors of Mathematics Learning: Working Memory, Phonological Ability and Numerical Competence. *Cognitive Development, 22*(2), 165-184.

Pólya, G. (1945). *How to solve it; a new aspect of mathematical method*. Princeton, N.J.,: Princeton University Press.

Primi, R., Ferrão, M. E., & Almeida, L. S. (2010). Fluid intelligence as a predictor of learning: A longitudinal multilevel approach applied to math. *Learning and Individual Differences, 20*(5), 446-451. doi:10.1016/j.lindif.2010.05.001

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*(4), 437-447. doi:10.1016/j.jml.2009.01.004

Raven, J. (2000). The Raven's Progressive Matrices: Change and Stability over Culture and Time. *Cognitive Psychology, 41*(1), 1-48.

Schoenfeld, A. H. (1985). *Mathematical problem solving*. Orlando, Fla.: Academic Press.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. *Handbook of research on mathematics teaching and learning*, 334-370.

Seghier, M. L. (2013). The Angular Gyrus: Multiple Functions and Multiple Subdivisions. *Neuroscientist, 19*(1), 43-61. doi:10.1177/1073858412440596

Shield, M., & Dole, S. (2013). Assessing the potential of mathematics textbooks to promote deep learning. *Educational Studies in Mathematics, 82*(2), 183-199. doi:10.1007/s10649-012-9415-9

Sidenvall, J., Lithner, J., & Jäder, J. (2015). Students' Reasoning in Mathematics Textbook Task-Solving. *International Journal of Mathematical Education in Science and Technology, 46*(4), 533-552.

Skolverket. (2000). *Grundskolan : kursplaner och betygskriterier* (1. uppl. ed.). Stockholm: Statens skolverk : Fritzes offentliga publikationer.

Skolverket. (2011a). *Läroplan för grundskolan, förskoleklassen och fritidshemmet 2011*. Stockholm: Skolverket : Fritze distributör.

Skolverket. (2011b). *Läroplan, examensmål och gymnasiegemensamma ämnen för gymnasieskola 2011*. Stockholm: Skolverket : Fritze distributör.

Skolöverstyrelsen. (1970). *Läroplan för gymnasiet* (6. tr. ed.). Stockholm.

Skolöverstyrelsen. (1980). *Läroplan för grundskolan. Allmän del : mål och riktlinjer, kursplaner, timplaner*. Stockholm: LiberLäromedel/Utbildningsförlag.

Stacey, K., & Vincent, J. (2009). Modes of Reasoning in Explanations in Australian Eighth-Grade Mathematics Textbooks. *Educational Studies in Mathematics, 72*(3), 271-288. doi:10.1007/s10649-009-9193-1

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building Student Capacity for Mathematical Thinking and Reasoning: An Analysis of Mathematical Tasks Used in Reform Classrooms. *American Educational Research Journal, 33*(2), 455-488.

Swanson, H. L., & Alloway, T. P. (2012). Working memory, learning, and academic achievement. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol 1: Theories, constructs, and critical issues* (pp. 327-366). Washington, DC, US: American Psychological Association.

Taub, G. E., Keith, T. Z., Floyd, R. G., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly, 23*(2), 187-198. doi:10.1037/1045-3830.23.2.187

Unsworth, N., & Engle, R. W. (2005). Working Memory Capacity and Fluid Abilities: Examining the Correlation between Operation Span and Raven. *Intelligence, 33*(1), 67-81.

Utbildningsdepartementet. (1994). *Läroplaner för det obligatoriska skolväsendet och de frivilliga skolformerna : Lpo 94 : Lpf 94*. Stockholm: Utbildningsdep. : Fritze.

Van Steenbrugge, H., & Norqvist, M. (2016). Unraveling students' reasoning: analyzing small-group discussions during task solving. Manuscript in preparation.

Wakefield, J. F. (2006). *Textbook usage in the United States: The case of U.S. History*. Paper presented at the International Seminar on Textbooks, Santiago, Chile.

Wiklund-Hornqvist, C., Jonsson, B., & Nyberg, L. (2014). Strengthening concept learning by repeated testing. *Scand J Psychol, 55*(1), 10-16. doi:10.1111/sjop.12093

Wikman, A. (2015). *Att tänka kreativt i matematik: Om hur olika resonemang uppmuntrar lärande* [To think creatively in mathematics: How different reasoning encourages learning]. Examensarbete, Umeå universitet, Institutionen för matematik och matematisk statistik. Från http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-113981

Wirebring, L. K., Lithner, J., Jonsson, B., Liljekvist, Y., Norqvist, M., & Nyberg, L. (2015a). Learning mathematics without a suggested solution method: Durable effects on performance and brain activity. *Trends in Neuroscience and Education, 4*(1-2), 6-14. doi:10.1010/j.tine.2015.03.002

Wirebring, L. K., Wiklund-Hornqvist, C., Eriksson, J., Andersson, M., Jonsson, B., & Nyberg, L. (2015b). Lesser Neural Pattern Similarity across Repeated Tests Is Associated with Better Long-Term Memory Retention. *Journal of Neuroscience, 35*(26), 9595-9602. doi:10.1523/Jneurosci.3550-14.2015

Wongupparaj, P., Kumari, V., & Morris, R. G. (2015). A Cross-Temporal Meta-Analysis of Raven's Progressive Matrices: Age groups and developing versus developed countries. *Intelligence, 49*, 1-9. doi:10.1016/j.intell.2014.11.008

Zamarian, L., Ischebeck, A., & Delazer, M. (2009). Neuroscience of learning arithmetic-Evidence from brain imaging studies. *Neuroscience and Biobehavioral Reviews, 33*(6), 909-925. doi:10.1016/J.Neubiorev.2009.03.005