PH.D.-AFHANDLING

PERNILLE BØDTKER SUNDE

Strategies in Single-Digit Addition: Patterns and Perspectives

AARHUS UNIVERSITET
DPU - DANMARKS INSTITUT FOR PÆDAGOGIK OG UDDANNELSE

Strategies in Single-Digit Addition: Patterns and Perspectives

Dansk titel:
Strategier i addition: Mønstre og perspektiver

Pernille Bødtker Sunde
PhD Dissertation
June 2019

Danish School of Education, Department of Educational Psychology, Aarhus University

Supervisor: Associate professor Lisser Rye Ejersbo, Aarhus University

Co-supervisor: Professor Uffe Jankvist, Aarhus University

The PhD scholarship was financed by The National PhD Council in Educational Research, and organised in collaboration between DPU, Aarhus University and VIA University College.

AARHUS UNIVERSITY

VIA University College

ISBN: 978-87-7507-460-0
DOI: 10.7146/aul.349

List of Papers

I. Sunde, P. B., Sunde, P., & Sayers, J. (2019). Sex Differences in Mental Strategies for Single-Digit Addition in the First Years of School. *Educational Psychology*. https://doi.org/10.1080/01443410.2019.1622652

II. Sunde, P. B., Sunde, P. (2019). Development and variance components in single-digit addition strategies in year one. Accepted for publication in *Proceedings of the Eleventh Congress of the European Society for Research in Mathematics Education.*

III. Sunde, P. B., Sunde, P. (Manuscript). Single-Digit Strategies in Year One Predicts Achievement in Year Four.

IV. Sunde, P. B. Sayers, J. (2017). Danish teachers' expectations of year one pupils' additive competence. Paper presented at NORMA 17. https://www.mnd.su.se/om-oss/evenemang/norma-17/programme/paper-session-2/sunde-and-sayers

V. Sunde, P. B. (manuscript). Teaching addition: Teachers' perspectives on teaching and learning number and addition in year one.

Paper I and II has been reproduced with permission from the publishers.

Summary

This dissertation addresses students' use and development of mental strategies in single-digit addition in the first years of school in Denmark. The overarching aim was to investigate longitudinal and cross-sectional patterns in Danish students' development and use of strategies for single-digit addition in relation to 1) different teaching practices, 2) later mathematical achievement and 3) teachers' perspectives on teaching and learning of number and arithmetic.

The work builds on data from two independent studies: A) a longitudinal study of the development of strategy use from year one to year four (147 students) and B) a study of six year one classes (83 students and six teachers) based on two assessments of students' strategy development in year one (October/November and April/May), along with teacher interviews on the teaching of number and arithmetic and classroom observations of teaching. The latter study was subsequently supplemented by data on later (year four) achievement in arithmetic, knowledge of fractions and word problem solving.

Students' strategy use was scored in one-to-one assessment interviews by sequentially presenting the student with 36 flash cards, each of which showed a single-digit addition problem using the addends 2 to 9. The solving strategy for each addition problem was scored in one of the seven categories: 1) Gives up, 2) Miscalculates, 3) Counting all, 4) Counting on, 5) Direct retrieval, 6) Derived facts +, and 7) Derived facts −. To ensure a sufficient number of observations in each strategy use category in some of the statistical analyses, the initial seven categories were reduced to the following four: 1) Error: the student gives up or repeatedly miscalculates, 2) Counting: all varieties of counting procedures, including finger counting, verbal, or self-report of mental counting, 3) Direct retrieval: the sum is automatized, and 4) Derived facts: addends are decomposed and automatized sums are used to calculate the answer (e.g. $4 + 5 = 4 + 4 + 1$ or $5 + 5 - 1$).

The dissertation comprises a synthesis and five individual papers. Paper I describes the proportional use of four distinct strategies ('error', 'counting', 'derived fact strategy' and 'direct retrieval') as a function of school age (year one to year four) and sex, based on data from Study A and verified by data from Study B. Paper II is a student-level statistical analysis of how strategy use changes over a half school year in year one (October/November–April/May) and the extent to which this development varies between classes in Study B. Paper III is an analysis of the extent to which strategy use in year one predicts mathematical achievement in year four (number and arithmetic, knowledge of fractions and word problem solving). Paper IV includes a qualitative analysis of teachers' expectations for year one students' additive competence in Study B. Finally, Paper V reports the same teachers' perspectives on teaching and learning number and addition in year one.

The key findings can be summarised as follows. From year one to year four, Danish students' use of counting strategies decreased, and their use of fact-based strategies increased (Paper I). However, over a half learning year in year one (November to April), this longitudinal developmental variation in strategy use (reflecting increasing number knowledge with age) was overridden by substantial individual variation, much of which

could be related to sex (Papers I–II). Hence, in year one, girls used counting three times more often on average than boys, a difference equivalent to at least two years' development (Paper I). Strategy use patterns varied little between classes (Papers I–II) and did not develop differently across classes in year one. It follows that differences in teaching practice did not result in any traceable differences in the pace of development of strategy use during year one (Paper II). The variation in strategy use across students was more or less constant throughout year one, with no indication that students who used counting most were reducing their use of this unsophisticated strategy more rapidly over time than students who used this method less often from the outset (Paper II). Measures of strategy use in year one—especially frequency of use of counting all—correlated significantly with mathematical achievement in year four. For two of three measurements of mathematical achievement (knowledge of fractions and word problem solving), strategy use patterns in year one accounted for variation unexplained by measures from a standard achievement test (Paper III). In other words, information about *how* a student performed single-digit addition in year one was a better predictor of later achievement than *how well* the student performed addition in year one (Paper III). Correlations between strategy use in year one and mathematical achievement in year four were generally stronger for boys than girls (Paper III). It can be hypothesised that this difference may reflect girls' greater inclination to use counting even after mastering retrieval strategies. If so, strategy use patterns in year one is a poorer predictor of later achievements in girls than in boys because in girls, a high frequency of use of 'counting all' to a higher extent reflects old habits and to a lower extent inability to use more advanced strategies than is the case for boys.

Taken together, these results indicate that habits of strategy use are deeply rooted within the individual child and seem to have been established either before or at the outset of formal schooling, changing only slowly and gradually over time. These results highlight the relevance of students' early understanding of number and arithmetic (i.e. strategy use) both as indicators of later achievement and as a focus for intervention and targeted teaching to support student development in all cases. From a teaching perspective, the results may prove useful in two ways. First, strategy use 'profile' from the assessment interviews could provide information on the student's risk of developing mathematical difficulties or becoming low achiever. In that respect, excessive use of counting strategies (and especially counting all) in the early years of school seems a valuable screening parameter. With early detection of students in need of special attention, targeted teaching can be initiated before difficulties emerge. Secondly, if the strong correlation between strategy use patterns in early year one and achievement in year four reflects a causal relation (which remains to be investigated), a stronger teacher focus on developing number sense and flexible strategy use for arithmetic in the earliest school years may potentially enhance students' mathematical achievements in later years and for life.

Keywords: mental strategies; arithmetic; single-digit addition; years one to four; mathematical achievement; sex differences; teaching practices

Dansk resumé

Denne afhandling omhandler danske folkeskoleelevers brug og udvikling af hovedregningsstrategier til etcifret addition i de første fire skoleår. Det overordnede mål var at beskrive variation og udvikling i danske elevers brug af hovedregningsstrategier til etcifret addition i de tidligste klassetrin, samt at undersøge sammenhænge mellem elevernes strategibrug i 1. klasse og 1) forskelle i undervisningspraksis, 2) senere matematiske viden og færdigheder og 3) lærernes perspektiver på undervisning og læring inden for tal og regning.

Afhandlingen bygger på analyser af data fra to uafhængige studier. Studie A er en undersøgelse fra samme skole af 147 elevers (hver elev testet mellem 1 og 3 gange) strategibrug fra 1. til 4. klassetrin. Studie B er en undersøgelse fra tre andre skoler af seks 1. klasser (83 elever og seks lærere), hvor elevernes strategibrug blev undersøgt to gange i henholdsvis oktober/november og april/maj. Derudover blev lærerne interviewet om deres perspektiver på undervisning og læring inden for tal og regning. Studie B er siden blevet suppleret med data fra en senere undersøgelse af elevernes viden og færdigheder inden for tal og regning, brøker og tekstopgaver i 4. klasse.

Strategibrug blev undersøgt med en standardiseret interviewprocedure, hvor eleverne blev præsenteret for 36 etcifrede additionsstykker, som de fortalte hvordan de løste. Elevens metode (strategi) til at løse hvert enkelt regnestykke blev inddelt i syv kategorier: giver op, regner fejl, tæller alt, tæller videre, automatisering, regruppering med addition (f.eks. $4 + 5 = 4 + 4 + 1$) og regruppering med subtraktion (f.eks. $4 + 5 = 5 + 5 - 1$). I nogle af analyserne, blev kategorierne slået sammen til i alt fire kategorier: fejl (giver op og regner fejl), tælling (tælle alt og tælle videre), automatisering og regruppering (regruppering med addition og subtraktion).

Afhandlingen består af en kappe og fem individuelle artikler. Artikel I beskriver udviklingen i den proportionale brug af de fire strategikategorier ’fejl’, tælling’, ’regruppering’ og ’automatisering’ som funktion af klassetrin (1. til 4. klasse) og køn. Analysen er baseret på data fra studie A og er verificeret på data fra studie B. Artikel II er en statistisk analyse af ændringerne i strategibrug på elevniveau over et halvt skoleår (oktober/november og april/maj) i 1. klasse og i hvor høj grad ændringer i strategibrug varierer mellem klasser i studie B. Artikel III er en analyse af sammenhængen mellem strategibrug i 1. klasse og matematiske færdigheder i 4. klasse (tal og regning, brøker og tekstopgaver). Artikel IV er en kvalitativ analyse af lærers forventninger til, hvad deres elever skal kunne inden for tal og regning efter 1. klasse. Artikel V analyserer de samme læreres perspektiver på undervisning og læring inden for tal og regning.

Afhandlingens hovedkonklusioner kan opsummeres som følger: Strategibrugen til etcifret addition ændres generelt fra 1. til 4. klasse med gradvist mindre brug af tællestrategier og en øget brug af regrupperingsstrategier og automatisering (artikel I). Set over et halvt skoleår (fra november til april) i 1. klasse, blev denne variation i udviklingen af strategibrug oveskygget af en betydelig individuel variation, som i høj grad kunne forklares ved køn (artikel I og II). I 1. klasse brugte piger således tælling tre gange så ofte i gennemsnit som

drenge, hvilket svarer til en forskel på mindst to læringsår (artikel I). Der var lille variation mellem klasser i mønstre af strategibrug (artikel I og II), og ændringer i strategibrugen i 1. klasse var ikke forskellig for de forskellige klasser. Heraf følger, at forskelle i undervisningspraksis ikke havde nogen målbar effekt på hvordan og med hvilken hastighed elevernes strategibrug ændredes i løbet af et halvt år i 1. klasse (artikel II). Variationen i strategibrug eleverne imellem var mere eller mindre konstant i 1. klasse uden tegn på, at de elever, som brugte tælling mest i oktober/november, reducerede deres brug af denne strategi frem mod i april/maj uforholdsmæssigt mere end de elever, som brugte tælling mindst (artikel II). Strategibrug i 1. klasse – især frekvensen af brug af 'tælle alt' – var signifikant korreleret med testscore for matematiske færdigheder og viden i 4. klasse. For to ud af tre testtyper i 4. klasse (brøker og tekstopgaver) forklarede strategibrug i 1. klasse variation som ikke blev forklaret af en standardiseret matematiktest i 1. klasse (artikel III). Med andre ord, kunne man ud fra testdata af hvordan eleven regnede i 1. klasse opnå et bedre mål for hvor godt eleven klarede sig i brøk- og tekstopgaver i 4. klasse end man kunne fra tilgængelige testdata på hvor rigtigt eleven regnede i 1. klasse (artikel III). Korrelationerne mellem strategibrug i 1. klasse og testscorer i 4. klasse var generelt stærkere for drenge end for piger (artikel III). Dette kan muligvis skyldes pigers tilbøjelighed til at bruge tælling som foretrukken regnestrategi i 1. klasse, selvom de fuldt ud mestrer regrupperingsstrategier. Hvis det er tilfældet, er pigernes strategibrug ikke i lige så høj grad som hos drengene et udtryk for deres talforståelse og regnefærdigheder.

Resultaterne tyder på, at strategibrugsmønstre er dybt forankret i det enkelte barn, at dette er etableret før eller ved den tidlige skolestart, samt at strategibrugen kun ændres langsomt og gradvist over tid. Dette viser betydningen af elevernes tidlige forståelse af tal og regning (strategier), både som indikator for senere viden og færdigheder, men også som fokus for intervention og særligt tilrettelagt undervisning med det mål at støtte alle elevers udvikling. Fra et didaktisk perspektiv er resultaterne interessante på to måder. Først og fremmest tyder resultaterne på, at elevens tidlige 'strategiprofil' giver værdifuld information om elevens risiko for at opleve vanskeligheder i matematik. I den henseende er især brugen af tællestrategier, og specielt 'tælle alt', i de tidlige skoleår efter alt at dømme en god screenings-parameter. Hvis elever i potentielle vanskeligheder kan opdages tidligt, kan man igangsætte tiltag før eventuelle vanskeligheder etableres. Hvis den stærke korrelation mellem elevers strategibrug i 1. klasse og deres evne til at løse mere avancerede matematikopgaver i 4. klasse er udtryk for en kausal sammenhæng (hvilket stadig mangler at blive eftervist), vil et øget fokus på at udvikle talforståelse og fleksibel strategibrug i de yngste år potentielt kunne fremme udviklingen af elevers matematiske viden og færdigheder i de senere skoleår.

Emneord: hovedregningsstrategier, aritmetik, etcifret addition, 1.-4. klasse, matematiske færdigheder og viden, kønsforskelle, undervisningspraksis

Preface and acknowledgements

When I left my job as a schoolteacher to become a PhD student, a former colleague said to me ‘don’t forget where you come from’. These words have stayed with me ever since as a reminder of why I embarked on this project in the first place: to create useful knowledge that will ultimately enhance mathematics learning in schools. When I worked as a teacher and consultant in primary schools, I had a special interest in students who struggled with mathematics, and I regularly observed that many students experienced difficulties in mathematics around years three and four—difficulties that appeared to relate to a failure to grasp the basic principles behind mathematical problems. To better identify these struggling children at an earlier stage, my friend Pernille Pind and I began to develop an easily used screening tool that assesses students’ strategies for single-digit addition. When developing this tool and subsequently applying it in practice, I was surprised by the diversity of students’ strategies for solving different types of simple digit addition problem. It gradually became clear that there were some interesting links between students’ strategy use and their achievement in mathematics.

My interest in this topic was further stimulated when I was asked to assess year seven and eight students who were struggling with mathematics. For even the simplest addition problems, they all seemed to count a lot—exactly as I had observed in the case of much younger students. At case level, it also struck me that students who used derived facts strategies on at least some single-digit problems in year one seemed to excel in mathematics more generally. In contrast, students who relied on counting strategies seemed to develop at a slower pace or eventually struggled with mathematics. These observations and experiences motivated me to adopt a more analytical and stringent approach to investigating these patterns, the immediate result of which is presented in this thesis. It is my humble hope that the present findings may provide teachers with new insights and tools to help students who struggle with mathematics.

As a PhD student, I have also come to value collaboration, both to achieve better academic outcomes and to improve myself in ways that would be impossible if working entirely alone. My co-authors and colleagues have challenged and enhanced my theoretical worldview and my analytical methodology, as well as my writing and the rigour of my logic and reasoning. Many people have made this thesis possible, most importantly the participating students and teachers; without you, this thesis would never have come to fruition, and I am grateful for your valuable insights into the teaching and learning of arithmetic in the early years of school.

Pernille Pind, you are perhaps the most important influence of all. In summer 2013, you persuaded me to write a contribution for the NORSMA7 conference about our joint work on developing the screening test that became the starting point for my PhD. Lisser Rye Ejersbo and Uffe Thomas Jankvist, I thank you for agreeing to supervise my work. Each in your own way, you provided valuable feedback and advice throughout the process. Judy Sayers, meeting you was a turning point in my life as a PhD-student. You convinced me not to give up and invited me to stay at Stockholm University for almost four months, where I have met many interesting scholars and other doctoral students. You convinced me that my

project was interesting and that the data contained important stories to be told. You taught me a lot about how to conduct research and write papers, for which I am truly grateful. I hope I have managed to deliver a few of those stories here. Paul Andrews, I thank you for the many conversations at Stockholm University about life as a researcher, the ups and downs of being a PhD student and the role of theory in educational research. Finally, I wish to thank all my colleagues at VIA University College, especially my office colleagues Kaj, Adrian, and Henning. You have listened patiently to my excitement when a new result emerged, engaging in discussion and supporting me whenever you could. Pernille Ladegaard Pedersen, I thank you for many fruitful discussions and coffee breaks, as well as the fun we had when participating in conferences and courses. Lóa Björk Jóelsdóttir, you deserve special thanks for taking care of my job in the last busy months of writing the thesis and making sure I met deadlines.

Through courses and conferences, I have met a number of people who took the time to engage, and their interest has also been important for my own and the project's development. Lieven Verschaffel, I thank you for your encouraging comments. Pirjo Aunio, I thank you for encouraging me to keep up my work and for your very concise and relevant comments, both on the topic in general and in particular on how to wrap up a thesis.

Finally, I would like to thank my family for their great support and understanding. To my two children, Lise and Kristian, I thank you for your immense patience when forced to listen to intense discussions over dinner about an interesting new solution strategy for 5 + 7 (there are surprisingly many!) and for managing on your own at times when I could not be there for you. Lise, thank you for the 'PhD emergency kit'—a box filled with chocolates for any possible crisis and the PhD comic book for whenever I needed a laugh (strongly recommended to other students)—which helped me through the last months. Kristian, thanks for your always valid comments—like this one, three days before deadline: 'Mom, as you always say to me, finishing in time is not a problem. You just have to start well in advance'. Peter, my husband and statistical consultant, you made it all possible. Without your support and engaged discussion, I would never have managed. When I was giving up, you made me continue. It has been a great pleasure to write papers with you, both those included here and others. I am grateful for all your valuable feedback and for enduring my shifting moods.

Thank you all.

Contents

1. Introduction

Numerical competences in the early years of school have been shown to influence later success, both in school and in life after school (Duncan et al., 2007; Parsons & Bynner, 2005). It therefore seems likely that leaving school with inadequate numerical competences has negative consequences, both for the individual and for society as a whole (OECD, 2010). It follows that even a modest improvement in the proportion of those who leave school with adequate numerical competences may yield substantial societal benefits.

In a Danish context, data from the final examination after ten years of compulsory education indicate that a significant proportion of Danish school leavers still find arithmetic challenging. Over the last two decades, 20–30% of students fail to answer whole-number arithmetic problems in the final examination. According to Danish Ministry of Education data from 2011, 25% of examinees were unable to correctly solve the problem '1018 – 619 = ?', and 46% could not correctly solve the problem '8032:8 = ?' (UVM, 2011). As these arithmetic problems are designed to elicit number-based strategies and are more difficult to solve using a standard algorithmic approach, this seems to indicate that a significant proportion of Danish students leave school without adequate numerical competences.

While many students experience difficulties with the four arithmetic operations throughout school, this pattern is likely to be established in the early school years, when there is a focus on number and arithmetic. In my former career as a primary teacher and mathematics consultant (for more than ten years), I observed that many of these difficulties appear around years three and four, when students are introduced to more complex multiplication and division, i.e. operations for which additive reasoning and counting strategies are no longer sufficient.

So, why do some students have such difficulty answering simple arithmetic questions after 10 years of compulsory school? This project addresses those aspects of early schooling that seem important for the development of arithmetic proficiency. To that end, I investigated young students' development of number and arithmetic competence, along with their teachers' perspectives on teaching and learning of number and arithmetic. As the issue has not to my knowledge been investigated in depth in a Danish context, it is hoped that this knowledge will be of use to both educators and policy makers.

The project is entirely problem-driven, both in the sense that the research problem guided my methodological choices (Arcavi, 2000; Ellis & Levy, 2008; Sternberg, 2007) and because it addresses a 'real world' problem (D. E. Gray, 2014; Robson, 2011). Applied research seeks to provide answers and possible solutions for stakeholders (D. E. Gray, 2014; Niiniluoto, 1993; Robson, 2011)—in this case, educators and policy makers. Based on the existing literature, the conceptual framework (Figure 1.1) shows how the project and its various components are interrelated and embedded in a process of problem-driven applied research.

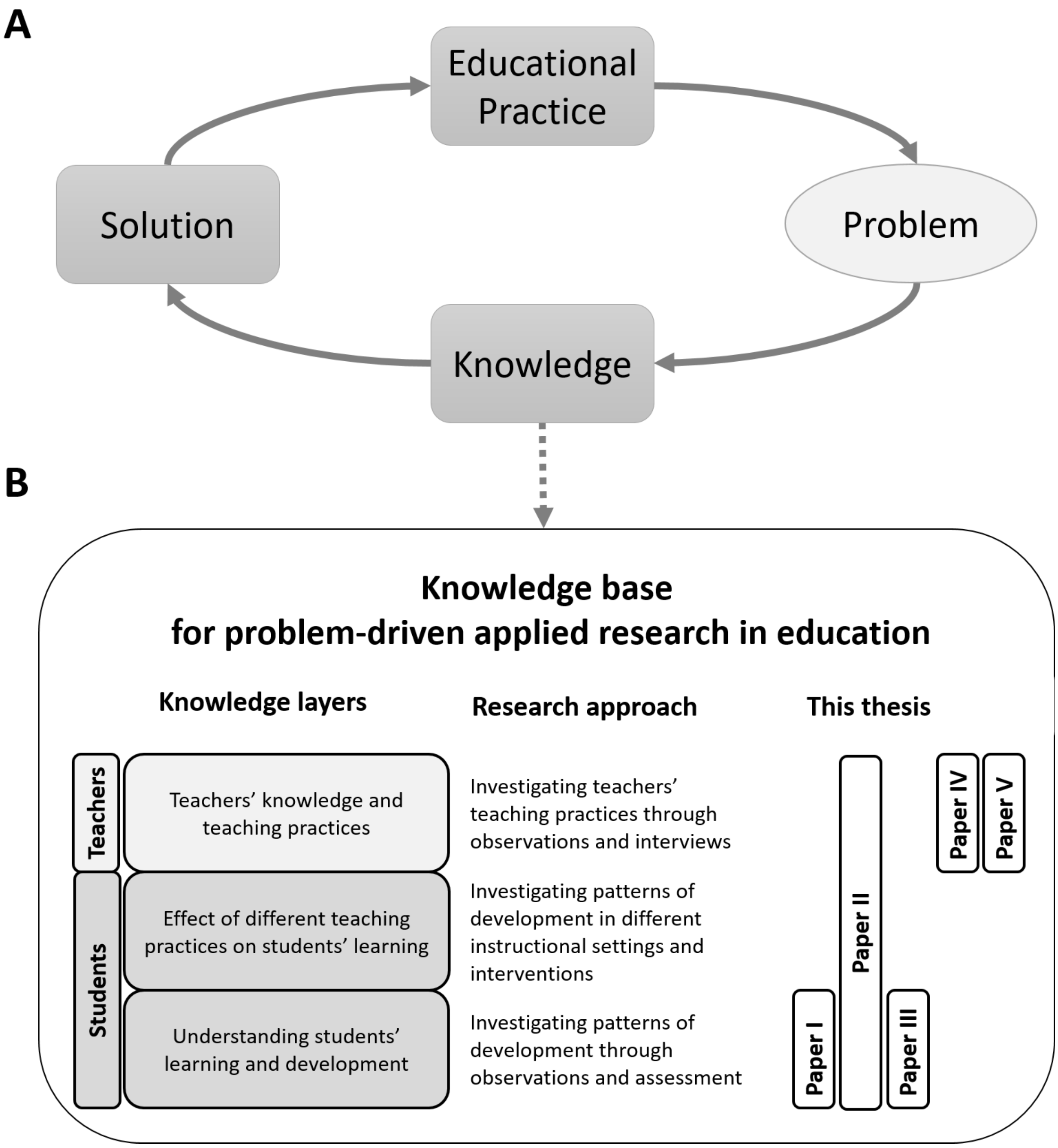


Figure 1.1 A tentative conceptual model of (A) the steps in problem-driven applied research and (B) the knowledge base needed to provide answers and solutions to 'real world' problems.
A: Problem-driven applied research can be understood as a cyclic process, typically starting with an apparent problem that must be identified and explicated. The next step is to gather the relevant knowledge; this involves combining existing and new research-based knowledge. The third step is to provide an answer or solution to the initial problem and to implement the findings in practice. In the process, solving the fundamental problem may draw attention to other previously ignored problems, or new problems may emerge as a consequence of the changed practice. For that reason, applied research can often be characterised as a cyclic process, in which new knowledge and answers lead to new problems and questions.
B: The different layers of the knowledge base needed to provide answers and relevant solutions for practitioners. The research focus is to identify patterns and mechanisms. For each knowledge layer, a research approach is proposed, and the contribution of each individual paper is indicated.

In the knowledge base underpinning solutions for practice (Figure 1.1, B), there are three layers or elements of knowledge. As these are interrelated and of equal importance, these are not to be understood as hierarchical levels but as different layers of knowledge. The bottom layer (students' learning and development) is what needs to be improved; along the interplay between learning and teaching practices, this is the 'students' knowledge block. The 'teachers' layer comprises knowledge of teachers and their teaching practices. These three overlapping and interacting layers of knowledge inform the investigation of the relevant patterns and mechanisms.

After outlining the arguments that define the project's focus, a brief overview of the study and data is followed by an overview of the dissertation.

1.1 Number, arithmetic and mathematics achievement

Number knowledge and arithmetic are important for the development of mathematical competences (e.g. Cowan et al., 2011; Gersten, Jordan, & Flojo, 2005), and the teaching and learning of arithmetic is a key topic in educational research (Nunes, Dorneles, Lin, & Rathgeb-Schnierer, 2016). The foundations for arithmetic skill and competence are laid in the early years of school; in particular, mental arithmetic strategies are generally considered important for developing arithmetic competences (e.g. Cowan, 2003; Nunes et al., 2016).The present study investigates the teaching and learning of strategies for mental addition during the early school years in a Danish context.

Internationally, the study of how arithmetic strategies are developed and used—especially for addition—has a long history (e.g. Ashcraft, 1995; Baroody & Dowker, 2003; Carpenter, Hiebert, & Moser, 1981; Siegler, 1996; Thornton, 1978). In a Nordic context, one major contributor is Snorre Ostad, whose longitudinal study of arithmetic strategy development compared students with and without mathematical difficulties (Ostad, 1997a). In light of the extensive existing research on this subject, it is reasonable to ask why there is a need to investigate students' development of strategies for mental single-digit addition. I argue here that there are at least two reasons.

1) Addition is the first arithmetic operation that students encounter in school mathematics and lays the foundation for further development. Simple arithmetic abilities are related to number sense (e.g. Desoete & Grégoire, 2006; Fuson & Burghardt, 2003; Geary, 2013), and the student's early use and development of these strategies indicates the quality and structure of their knowledge of number and their ability to reason with numbers (Fuson, 1992). The first such strategy students learn is *counting,* in which it suffices to understand each number as an entity (Fuson, 1992; Verschaffel, Greer, & De Corte, 2007) or to see numbers as objects on a number line. More advanced strategies, such as *derived fact strategies* require students to 'see' numbers as objects that can be structured in many ways (Ambrose, Baek, & Carpenter, 2003; Verschaffel et al., 2007). Students' early number and arithmetic competences predict later achievement (e.g. Desoete & Grégoire, 2006); the qualitative use of mental arithmetic strategies early in school is known to be a valid predictor of later mathematical achievement and difficulties (Gersten et al., 2005; Ostad, 1997a; Price, Mazzocco, & Ansari, 2013). If this link is causal—that is, if mental strategy use or underlying knowledge of number drives development of mathematical achievement—a faster and more efficient transition in early arithmetic learning to the use of more sophisticated mental strategies (possibly through development of number knowledge) may support a deeper

understanding of number and arithmetic and the ability to comprehend more complex mathematics.

2) Previous studies have shown that students' development of strategies in arithmetic differ substantially, both at individual level (e.g. Dowker, 2014; Siegler, 1996; Torbeyns, Verschaffel, & Ghesquiere, 2004a) and between countries and cultures (e.g. Shen, Vasilyeva, & Laski, 2016). This has been attributed to differences in didactic approaches in early arithmetic education (Geary, Bow-Thomas, Liu, & Siegler, 1996; Shen et al., 2016), as well as to culture-specific extracurricular factors (Campbell & Xue, 2001). As only a few studies to date, such as Gaidoschik (2012) in Austria, have investigated the relationship between classroom teaching practices and students' arithmetic strategy development, it is useful to examine teachers' knowledge and practices in this regard in a Danish context.

1.2 Study overview

This dissertation investigates students' use and development of mental strategies for single-digit addition in the first years of school in Denmark. The initial aims were to explore and describe 1) students' development of strategy use, 2) teachers' perspectives on teaching and learning number and arithmetic; and 3) teaching practices in the year one classroom, and how these are related. For reasons explained below, two additional issues emerged in the course of the study: 4) the extent to which strategy use patterns in year one were associated with or predicted mathematical achievement in year four; and 5) the existence of systematic sex differences in strategy use.

The present work builds on data from two independent studies: A) a longitudinal study of 147 students' development of strategy use from year one to year four and B) a study of students' strategy development in six year one classes (83 students and six teachers) based on assessment data (twice in year one: October/November and April/May), as well as interviews with teachers (on teaching number and arithmetic) and observations of classroom teaching. Students' use and development of strategies was investigated in one-to-one assessment interviews of each student on 36 single-digit addition tasks, and their general proficiency with numbers and arithmetic was measured using a standardized test.

Three years later, a subsequent study measured the ability of study B students (now in year four) to solve more complex mathematical problems involving multiplication, division, fractions and word problems. These data provided an opportunity to analyse the extent to which strategies used to solve simple arithmetic questions in year one predicted students' later ability to solve more complex mathematical problems. These data on mathematical achievement in year four formed part of a larger research project. Although data is not complete and could not be fully analysed, I have chosen to include them here, as they provide important insights into the importance of students' early use and development of mental strategies in arithmetic.

Finally, it became apparent in both studies that boys and girls differ in their use and development of mental strategies. For that reason, gender was included as a variable in all the analyses of strategy use. Both psychological and educational research have established that boys and girls may differ in terms of general mathematics performance (e.g. Penner & Paret, 2008; Voyer & Voyer, 2014) and cognitive skills such as mental rotation (e.g. Levine, Huttenlocher, Taylor, & Langrock, 1999; Moè, 2018). In a Danish context, girls score lower than boys in the final examination in mathematics (Figure 1.2), and TIMSS and PISA results confirm that Danish boys generally outperform girls (Mullis, Martin, Foy, & Hooper, 2016; OECD, 2016). As sex differences in behaviour and achievement can be considered politically controversial (Frank, 2019), it is important to emphasize that, from a learning perspective, such differences are not in themselves particularly interesting. In learning situations, each student should be viewed as an individual with their own unique prerequisites for learning, and teaching practice should vary accordingly. Based on the overall framing of the problem in Figure 1.1, any sex differences are therefore interesting primarily as a 'framing' variable encapsulating more informative causal predictor variables that remain to be deconstructed in subsequent analyses.

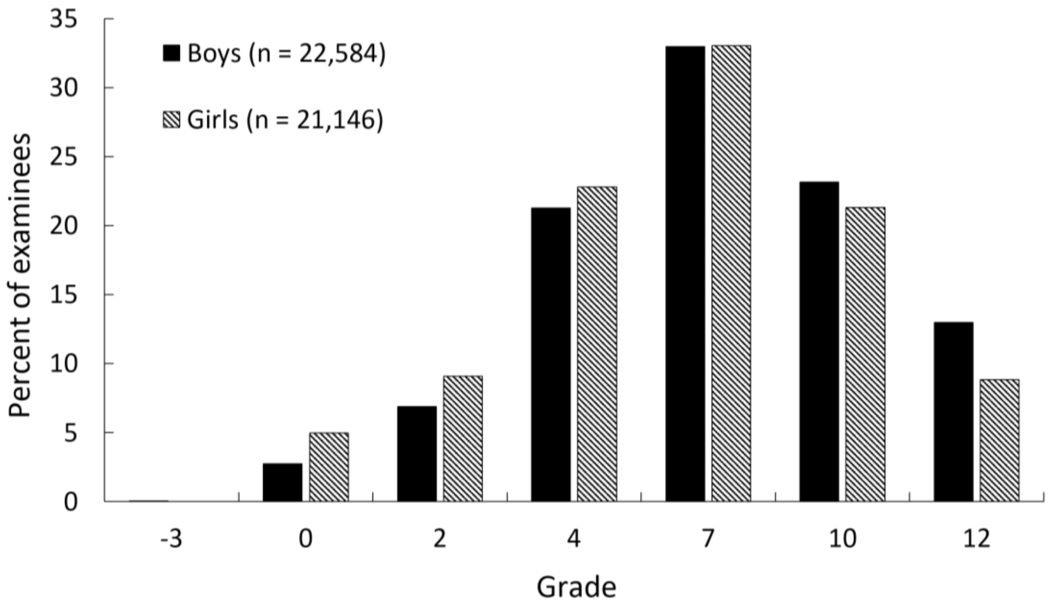


Figure 1.2 Distribution of grades in Danish National elementary school examination of general mathematical skills, May 2017 (by sex). The scale defines -3 (no performance) and 0 (improper performance) as failing grades, and 2 (just passed) to 12 (very good) as passing grades (roughly comparable to grades E,D,C,B,A on the ECTS scale. The mean score is 6.6 for girls and 7.2 for boys. These results are similar to those for examination years 2014, 2015 and 2016.
(Based on data available at https://uddannelsesstatistik.dk/Pages/Reports/1058.aspx)

1.3 Overview of the dissertation

The rest of the dissertation is organized as follows. Chapter 2 reviews the literature on arithmetic strategy development and gender differences. Chapter 3 specifies the study's objectives and research questions. Chapter 4 describes methods and methodological choices, including ethical considerations. Chapter 5 presents summaries of the five papers. Finally, in Chapter 6, I discuss the results from the individual papers and consider the implications for teaching within the overall problem-oriented context of teaching and learning number and arithmetic in Danish primary schools.

Specific results are detailed in the five papers attached to this thesis.

Paper I is based on a statistical analysis that describes the proportional use of four different strategies ('error', 'counting', 'derived fact strategy', and 'direct retrieval') as a function of school age (year one to four) and gender.

Paper II describes a statistical analysis of how strategy use changes over a half school year (October/November–April/May) in year one and the extent to which strategy use and development varies between classes.

Paper III reports preliminary findings on the association between students' strategy use in year one and mathematics achievement in year four.

Paper IV describes a qualitative analysis of teachers' expectations for their year one students' additive competence.

Paper V reports teachers' perspectives on teaching and learning number and addition in year one.

2. Strategy use in single-digit addition

This chapter on mental strategies for single-digit addition begins with a historical overview of relevant strategies and models of development. I go on to review existing knowledge about factors that influence the use and development of strategies, including individual and group differences and socio-cultural context. Then, I review existing knowledge concerning the relationship between mental strategies and general achievement in mathematics. Finally, I review sex differences with respect to the aforementioned topics.

2.1 Introduction

Mental strategies in arithmetic provide the foundation for more complex mathematics (e.g. Cowan et al., 2011) and have been found to predict later mathematics achievement (Gersten et al., 2005; Ostad, 1997a; Price et al., 2013). These strategies are based on number knowledge and understanding (Ambrose et al., 2003; Fuson, 1992; Threlfall, 2002; Verschaffel et al., 2007) and provide insights into the quality and structure of a student's number knowledge and ability to reason with numbers.

Students' task performance is a result of their 1) repertoire of strategies, 2) the strategy distribution, i.e. frequency of using different strategies, 3) strategy efficiency, i.e. the speed and accuracy of executing a strategy, and 4) strategy selection, i.e. the flexibility of strategy choice (Lemaire & Siegler, 1995). Thus, the choice of a strategy to solve a specific problem is influenced by the individual student's repertoire of strategies and the efficiency of execution—that is, their speed and accuracy in executing a specific strategy (e.g. Siegler, 1996; Siegler & Lemaire, 1997), where speed and accuracy are positively correlated with frequency of use (e.g. LeFevre, Sadesky, & Bisanz, 1996; Siegler & Lemaire, 1997). The present study focuses on patterns of students' strategy repertoire for single-digit addition and frequency of use.

In general, there is no unequivocal definition of *strategy* (see for example Ostad, 1997b). Siegler and Jenkins (1989) distinguished between *procedure* and *strategy*, arguing that while a procedure represents the only way to achieve a goal, strategies are procedures chosen from among others. They define a strategy as 'any procedure that is nonobligatory and goal directed' (p. 11). Ostad (1997b) distinguished between task-specific and general strategies, defining the former as '*organized, domain-specific, nonobligatory patterns of decisions activated when confronted with mathematical (arithmetical) problems, and goal directed to attain the solution of the problem'* (p. 12). For present purposes, I adopt Siegler and Jenkins' (1989) usage of strategy as an *optional* series of actions directed towards solving a specific problem as distinct from a *procedure* as a *fixed* series of actions. The term *mental strategies* is used here to refer to calculations that do not involve written algorithms.

2.2 Strategies in single-digit addition

Addition is the first arithmetic operation that students are introduced to in primary school. However, children's first experiences of arithmetic occur long before formal schooling (e.g. Baroody & Wilkins, 1999), and in reality, most students already have some knowledge of numbers and basic arithmetic skills when beginning primary school (e.g. Clements & Sarama, 2007). Exposure to practical everyday problems and play activities (e.g. in kindergarten) mean that, before receiving any formal teaching in arithmetic, most students are perfectly capable of solving problems such as 'You have two cars and I have four. How many do we have all together?' Because these early experiences involve the manipulation of concrete objects, children's first addition strategy is based on counting. The development of counting strategies has been thoroughly described elsewhere (e.g. Baroody, 1987, 1989; Fuson, 1992) but can be briefly summarised as follows. The first counting strategy that a child develops or invents is counting *all*. Both addends are first counted independently (1-2 cars and 1-2-3-4 cars) and then all together (1-2-3-4-5-6 cars). The next step is to skip individual counting of the addends, instead counting all (1-2-3-4-5-6 cars). As they become capable of recognising small quantities without counting or when performing symbolic addition, the child gradually begins to use counting on, initially from the first addend (3-4-5-6 cars). In turn, the *min* strategy (minimum counting) is developed, counting on from the larger addend (5-6 cars).

As the child practises addition by counting, an increasing number of sums are automatized, leading to the development of so-called *fact-based* strategies. *Direct retrieval* (also called 'fact retrieval' or simply 'retrieval') is direct recall from memory of the solution to an addition question without any intervening computation. Known facts of this kind are an important prerequisite of *derived fact* strategies, which have been explored in many recent studies (e.g. Dowker, 2014; Gaidoschik, 2012; Laski, Ermakova, & Vasilyeva, 2014). Derived fact strategies are also known as *decomposition* or *partitioning* strategies. These build on known facts and involve transforming the addition question into more simple questions. Derived fact strategies fall into two distinct groups based on 1) known facts (often doubles) or 2) adding up to 10 (e.g. LeFevre et al., 1996). An example of the former is solving the unknown addition of 7+8 by transforming it into the known sum of 7+7 and then adding 1. An example of adding up to 10 is solving the same problem (7+8) by partitioning its addends—for example, dividing 7 into 2 and 5, adding 8 and 2 to get 10 and then adding the remaining 5 to get the final sum (15).

Derived fact strategies are generally considered more advanced, as they build on a more complex understanding of numbers. Fact-based strategies not only require some facts to be automatized (E. M. Gray, 1991; Threlfall, 2002) but also depend on an understanding of relations between operations and a concept of numbers as objects that can be partitioned in many ways (Ambrose et al., 2003; E. M. Gray & Tall, 1994; Verschaffel et al., 2007). Understanding of part-whole concepts has been found to relate to primary students' use of derived fact strategies (Canobi, 2004). On the other hand, counting builds on an understanding of numbers as standalone entities—for example, that five is *before* six but is not *part* of six (Fuson, 1992). Unlike retrieval and fact-based strategies, counting does not require an understanding of part-whole relations; to count, a number line- or number sequence-based understanding of numbers will suffice.

The above strategies can be performed either by using manipulatives or by purely mental means. In a series of studies (Carr & Alexeev, 2011; Carr & Davis, 2001; Carr, Steiner, Kyser, & Biddlecomb, 2008), Carr and colleagues draw a distinction between *manipulative* and *cognitive* strategies. Manipulative strategies refer to all kinds of counting supported by manipulatives of various kinds, such as fingers or marks on paper. In contrast, cognitive strategies comprise all counting strategies that do not involve manipulatives (e.g. silent mental counting), and derived fact and direct retrieval. This leads to a slightly different categorisation of strategies as compared to those described above.

These different addition strategies involve a development from counting to direct retrieval and derived fact strategies (Carpenter & Moser, 1984; Torbeyns, Verschaffel, & Ghesquiere, 2004a). The next section describes how this understanding (or developmental model) has developed over time, based on different measuring methods and an increased focus on individual differences.

2.2.1 Measuring strategy use and models of development

Our understanding of how and when strategies in mental arithmetic develop has changed substantially over time, in part because of changes in how strategy use is measured. Early models of strategy use (primarily from cognitive psychology) focused on the sequence of acquisition of new strategies and frequency of use and were based on measurements of reaction time (Ashcraft, 1992). In general, problems involving large addends take longer to solve. This so-called ‘problem size effect’ has been explained in terms of how stored arithmetic facts are represented in memory and by speed of counting: the larger the addends, the longer it takes to count, and reaction time for direct retrieval is shorter than for counting procedures. The interpretation of differences in reaction time for tasks of different size or difficulty resulted in the development of models that focused on two main strategies: counting—Groen and Parkman’s (1972) min model of minimum counting from the larger addend—and direct retrieval—Ashcraft’s (1982) network or structural model.

Both of these models are based on the notion that addition is solved *either* by counting *or* by direct retrieval. In that sense, these explanatory models centred on task characteristics and the idea that direct retrieval replaces counting with time and practice. Siegler (1996) used the metaphor of a *staircase* to capture the Piagetian view of development that underpinned this linear or sequential approach, in which one strategy is gradually abandoned when a new and more efficient strategy is acquired. This is especially pronounced in research on counting, which is one of the most thoroughly investigated groups of strategies (e.g. Ashcraft, 1982; Baroody, 1987; Carpenter & Moser, 1984; Groen & Parkman, 1972).

These models conflicted with reports that students use a variety of strategies. The growing focus on self-reported strategy use revealed that students use strategies other than counting and direct retrieval. With the qualitative approach in educational research, interviewing students and asking them to describe their actual problem-solving processes, other strategies like derived fact or mixed strategies gradually influenced models of strategy use (Carpenter et al., 1981; Carpenter & Moser, 1984). Combining this qualitative approach (student self-reports) with quantitative measurements (chronometric measures of reaction time), Siegler (1987, 1988) not only confirmed that children use multiple strategies but also demonstrated that these derived fact strategies could account for some of the deficits in earlier models based solely on counting and direct retrieval (Ashcraft, 1992). On analysing these strategies independently, Siegler (1987) found that the min model was a good predictor of solution time when students reported counting from the larger addend, but not when students reported using other strategies.

The shortcomings of the min model were addressed in the *distribution of association* model (Siegler & Shrager, 1984) and later in the *strategy choice* model (Siegler & Jenkins, 1989), building on the assumption that a problem is associated with both correct and incorrect answers, as well as various procedures. The more often a problem is practised and associated with a correct answer, the higher the 'peak' for the association of problem and correct answer and the procedure that leads to it. Siegler and colleagues (Siegler & Jenkins, 1989; Siegler & Shrager, 1984) argued that as non-retrieval strategies (e.g. counting) more often lead to wrong answers for large problems, the distribution of association for these larger problems will be less peaked, so explaining the problem size effect.

Studies of reaction times for doubles or ties further supported these findings. In general, doubles exhibit substantially faster reaction time patterns and smaller problem size effects than non-ties (Groen & Parkman, 1972; Thevenot, Barrouillet, Castel, & Uittenhove, 2016). In a study comparing adults' self-reported strategy use with reaction times for single-digit addition problems, LeFevre et al. (1996) found that ties were solved primarily by direct retrieval while non-ties where solved using a variety of strategies. They also found that strategy choice had a major impact on reaction time, and this could account for the problem size effect (LeFevre et al., 1996); for example, the problem size effect was modest for problems solved by direct retrieval.

Adding change in strategy use to the models, Siegler (1996) demonstrated that this change develops gradually and involves a shift in frequency of use rather than the replacement of one strategy by another. Siegler's (1996) *changing wave* or *overlapping waves* theory of strategy change proposed that multiple strategies are available and in use at the same time.

Increasing reports of variability and developmental differences in strategy use at both individual and group levels (e.g. E. M. Gray, 1991; E. M. Gray & Tall, 1994; Siegler, 1996) led to a shift of focus from *which* strategies are used at *what* developmental stage—that is, *the sequence of development*—to *how* change in strategy use occurs and under *what conditions* different strategies are used—that is, *patterns of development*.

The general understanding is that strategies in single-digit addition develop from counting to direct retrieval and derived fact strategies, and that the change happens gradually (Carpenter & Moser, 1984; Siegler, 1996; Torbeyns, Verschaffel, & Ghesquiere, 2004a). While several studies have confirmed this pattern (e.g. Torbeyns, Verschaffel, & Ghesquière, 2002; Wylie, Jordan, & Mulhern, 2012), this developmental pathway has been called into question in recent years; instead, it has been suggested that direct retrieval may just involve very fast unconscious counting (Thevenot et al., 2016). Gaidoschik (2012) found that counting on is not a prerequisite for automatization of addition facts, and he suggested instead that derived facts facilitate this development.

2.2.2 Adaptive strategy choice

The factors that influence students' strategy choices relate not only to specific task characteristics but also to characteristics of the individual student. These may include cognitive variables (e.g. Geary, Hoard, & Nugent, 2012), as well as the individual student's repertoire and frequency, efficiency, adaptability and flexibility of strategy use (e.g. Bailey, Littlefield, & Geary, 2012; Rechtsteiner-Merz & Rathgeb-Schnierer, 2015; Torbeyns et al., 2002; Torbeyns, Verschaffel, & Ghesquière, 2005, 2004). Although definitions of these concepts are somewhat inconsistent and overlapping (see for example Nunes et al., 2016), they contribute to the understanding of why some students develop adaptive expertise while others do not. Hatano (2003) defined adaptive expertise as ‘the ability to apply meaningfully learned procedures flexibly and creatively’ (pp. xi). In other words, adaptive strategy choice means selecting a strategy while taking account of the efficiency of applying a specific strategy to a specific problem and choosing one that will most rapidly and accurately produce an answer (Siegler & Lemaire, 1997; Torbeyns et al., 2005).

The choice/no-choice method described by Siegler and Lemaire (1997) facilitates comparison of the student’s actual strategy preferences with the most appropriate strategy from a task analysis perspective, so providing an unbiased assessment of the student’s adaptive strategy use. The student is tested in two situations: 1) a free choice situation (where the student is free to choose their preferred strategy) and 2) a no-choice situation (where the student must apply a specific strategy). This method of investigating strategy use and development takes account of the student’s repertoire of strategies and their speed and accuracy in applying those strategies on tasks with specific characteristics. For example, from a task characteristic perspective, the addition task 7+8 would be solved most efficiently by a double +/- 1—that is, either 7+7+1 or 8+8-1. From a student perspective, however, this might not be the best choice of strategy if, for instance, the individual is not efficient with doubles (in terms of correctness and speed). Depending on the student’s repertoire and efficiency with different strategies, a better strategy might be to use friends of ten, solving the task by splitting 8 into 3 and 5 and then adding 7+3+5, or alternatively by counting on from 8. In this way, the choice/no-choice method (Siegler & Lemaire, 1997; Torbeyns et al., 2005) makes it possible to analyse strategy use based on the individual student’s skills—that is, their speed and accuracy when using a specific strategy for a specific task. In so doing, this approach extends our understanding of the relationship between efficiency and adaptability in strategy use.

The factors outlined above relate either to task characteristics or to individual student characteristics. However, arguing for the importance of socio-cultural context, Verschaffel, Luwel, Torbeyns and Van Dooren (2009) defined adaptive choice of strategy as ‘… *the conscious or unconscious selection and use of the most appropriate solution strategy on a given mathematical item or problem, for a given individual, in a given sociocultural context*’ (p. 343). For example, Bjorklund and Rosenblum (2002) found that 6- and 7-year-old children used more sophisticated strategies to solve addition problems when playing Chutes and Ladders than when asked to solve the same problems in an academic context. The children also made more errors in the academic context than when playing the game.

To sum up, the factors that influence students’ use of specific strategies can be assigned to three main categories: (1) factors related to *task characteristics*; 2) factors related to *individual characteristics* (e.g. cognitive factors, individual repertoire, speed and accuracy); and 3) factors related to the *socio-cultural context* in which the student acquires knowledge and executes the strategies (e.g. school environment, teaching practices, parental influence).

Figure 2.2 depicts the relationship of task characteristics, individual student characteristics and socio-cultural context to the student’s use of strategies, and the relationship between strategy use and mathematics achievement. These three categories should not be considered absolute and discrete—that is, socio-cultural context may influence individual characteristics and mathematics achievement or vice versa. However, these distinctions serve as a useful guideline for research purposes as a way of decomposing the different factors that contribute to variations in student strategy use.

The next section unpacks and illustrates how task characteristics, student characteristics and socio-cultural context relate to strategy use and reviews the literature on the relationship between students' strategy use and later achievement in mathematics. Finally, the relevance of sex in relation to these themes will be exemplified.

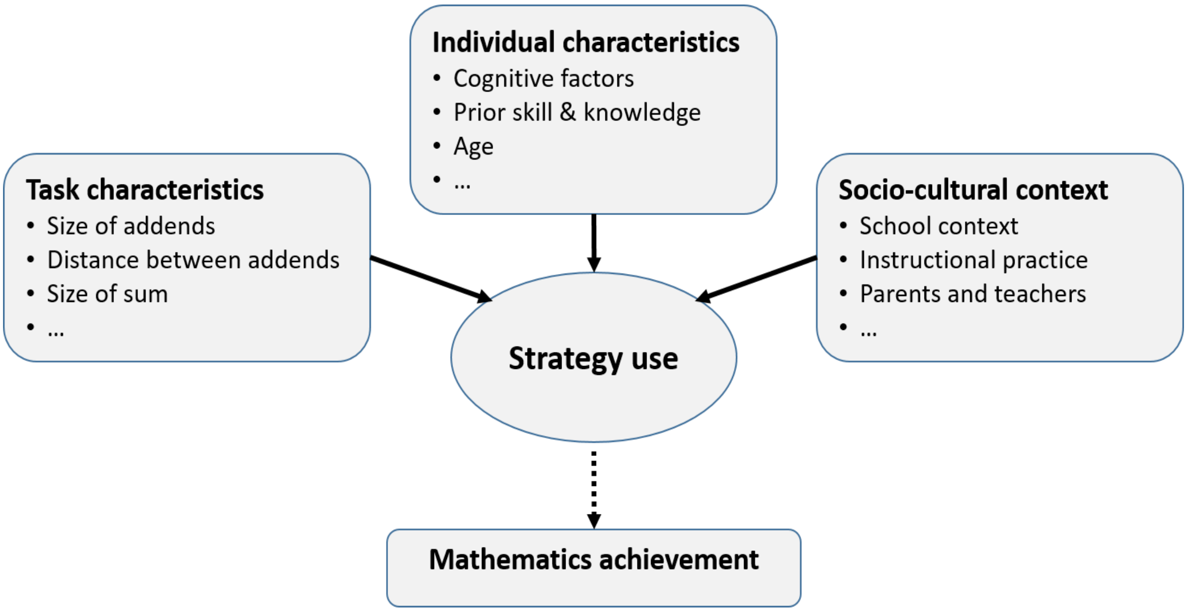


Figure 2.2 Overview of factors influencing student strategy use and development and the relationship between strategy use and mathematics achievement. Students' adaptive choice of strategy is influenced by task characteristics, individual student characteristics and socio-cultural context.

2.3 Task characteristics

For any given student, the optimal strategy for solving a given problem will depend in part on the nature and complexity of the problem or task itself. For example, a particular student may choose to solve 3+3 by direct retrieval, 7+2 by counting on from 7, and 8+7 by using a derived fact strategy (e.g. 8+8-1), simply because, for this student, these are the easiest and fastest strategies for the respective tasks. When analysing differences in strategy use, it is important to be aware that (optimal) strategy use differs—not only between students for the same problem but also within students for different types of problem.

Differences in strategy choice for tasks with different characteristics have been demonstrated for kindergarteners (e.g. Laski et al., 2014) and young students (e.g. Blöte, Klein, & Beishuizen, 2000). For adults, too, LeFevre et al. (1996) identified a clear relationship between task characteristics, chosen solution strategy and reaction time. For tasks with sums greater than 10, derived fact and direct retrieval strategies were used with almost equal frequency, even though direct retrieval was faster. For sums up to 10, fast and accurate direct retrieval was most frequent, and derived facts were rarely used. Counting was used most often when adding 1, 2 or 3, and reaction times were predicted by the required number of counts.

The influence of task complexity on strategy choice has also been demonstrated for kindergarteners. For example Laski et al. (2014) found that kindergarteners were more likely to use decomposition for complex multi-digit problems or for problems involving bridging ten than for simple problems. Similarly, Blöte et al. (2000) found that, during a teaching intervention, Dutch year two students developed a good understanding of the relationship between number characteristics and calculation procedures for addition and subtraction of numbers up to 100. However, they did not fully use this knowledge when choosing a solution strategy; factors other than number characteristics also influenced strategy choice, and the students used more flexible strategies for context problems than for symbolic problems (Blöte et al., 2000).

The choice/no-choice method outlined in 2.2.2 makes it possible to pinpoint the influence of task characteristics on individual strategy choice, although this method has some limitations. As noted by Torbeyns and colleagues (Torbeyns, Arnaud, Lemaire, & Verschaffel, 2004; Torbeyns, Verschaffel, & Ghesquiere, 2004b) the number of no-choice situations should reflect all the strategies applied by students in the choice situation. This might entail multiple no-choice situations, with practical implications for implementation of the experimental setup and participants' performance (Torbeyns et al., 2004).

From the researcher's perspective, if different task characteristics influence the individual's choice of strategy, it is important to bear this in mind when specifying tasks and/or analysing and interpreting results. When investigating strategy use in single-digit arithmetic, it is appropriate to use a set of combinations of addends that are representative of all combinations of addends; leaving out doubles or certain combinations could influence the results. In statistical analyses of patterns of strategy use, the inclusion of task characteristics in statistical models may account for apparent background 'noise' that would otherwise blur patterns of variation between students.

2.4 Individual student characteristics

Individual student characteristics include cognitive factors and prior knowledge and skill. Cognitive factors (or domain-general factors) relate to how the student learns and to how they acquire and execute skills and competences. Prior knowledge and skills (or domain-specific skills) reflect previous learning and are known predictors of later achievement. In many cases, the effect of prior knowledge explains much of the students' development as compared to other factors (e.g. Hailikari, Nevgi, & Komulainen, 2008). Prior skills and knowledge may relate to the same mathematical topic—for example, prior skills in retrieving facts predicts later skills (e.g. Bailey et al., 2012)—or to other prerequisites for the specific skill or competence in question, such as specific number skills that influence development of strategy use (e.g. Gaidoschik, 2012; Laski et al., 2014).

2.4.1 Cognitive factors

Cognitive factors such as mental rotation skill (e.g. Laski et al., 2013; Moè, 2018) and working memory (e.g. Geary et al., 2012; Paul & Reeve, 2016) have been linked to aspects of number and arithmetic in general, as well as to the use and development of specific mental arithmetic strategies. For instance, Geary and colleagues established that working memory (central executive component) were related to strategy use and developmental changes (Geary, 2011; Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Geary et al., 2012; Stoet & Geary, 2013). Imbo and Vandierendonck (2007) investigated the importance of working memory for arithmetic performance in grades four to six. They assessed students' strategy use when solving large sum single-digit addition problems in a series of choice/no-choice conditions, with and without a dual task requirement, which loaded the executive component of working memory. In three no-choice conditions, the students used either counting, direct retrieval or decomposition to solve the tasks. Imbo and Vandierendonck's (2007) results suggest that central executive resources are involved in arithmetic performance, and they reported that working memory played a greater role in derived fact and counting than in direct retrieval. They also found that the negative influence of executive working memory load decreased with age, indicating that less working memory resources are needed with increased learning and more frequent use of direct retrieval.

Foley, Vasilyeva and Laski (2017) reported that limited short-term visuospatial memory influences the use of decomposition, so reducing accuracy. In a study of first grade girls, Laski et al. (2013) found a positive relation between use of fact-based strategies and mental rotation skills, and a negative relation between use of counting and visuospatial ability. Casey, Lombardi, Pollock, Fineman and Pezaris (2017) identified a direct pathway from the spatial skills of first grade girls to their fifth grade mathematics reasoning.

Processing speed is another cognitive factor that predicts arithmetic fluency. Cowan and Powell (2014) found that it accounts for variation in basic calculation fluency among year three students, and Fuchs et al. (2006) found that processing speed in conjunction with attentive behaviour and phonological decoding was the most important predictor of addition and subtraction fact fluency.

As well as arithmetic fluency, cognitive variables such as working memory, visuospatial skills and processing speed seem to influence (or are at least correlated with) the use of specific strategies, and this may have interesting implications for learning. According to Cheng and Mix (2014), enhancing these capabilities may improve students' arithmetic ability; they found that training mental rotation skills improved students' arithmetic performance in multi-digit addition and subtraction and in missing-term problems (e.g. 4 +__ = 12) but not in fact retrieval.

2.4.2 Prior skill and knowledge in strategy use

Prior knowledge of different strategies and skill in using them influences students' strategy choice and development. Using the choice/no-choice method (Lemaire & Siegler, 1995; Siegler & Lemaire, 1997), Torbeyns, Verschaffel and Ghesquière (2002, 2005; 2004a, 2004b; 2004) have demonstrated that strategy repertoire, as well as speed and accuracy of application, influences strategy choice for specific tasks. The student's level of adaptive expertise (Hatano, 2003; Hatano & Oura, 2003) also influences further development. Bailey, Littlefield, and Geary (2012) showed that even small differences at the onset of development can result in large differences in strategy use. They found that preference for a specific strategy and skill in using it formed a feedback loop; early preference predicts later skill, which predicts later preference, and so on.

For whatever reason, some students initially prefer counting as a strategy for solving basic arithmetic problems and use this at the expense of other strategies for working more flexibly with numbers. This means that they also get less practice in working flexibly with numbers and therefore have fewer opportunities to develop the more complex understanding needed for advanced arithmetic (Gersten et al., 2005). Laski et al. (2013) suggested that girls' poorer accuracy in derived fact and retrieval strategies may be a result of their persistent counting, which affords fewer opportunities to practice fact-based strategies. The longer a student continues to use inadequate strategies and get the right answer, the greater the effort needed to change strategy (e.g. E. M. Gray, 1991). Students who prefer counting are more likely to use this strategy when introduced to new arithmetic problems such as multiplication. From a didactical perspective, then, it is important to recognize and support each individual student's development of strategies.

2.4.3 Number sense

Threlfall (2002) emphasized that mental calculation requires both knowledge of basic facts and fact families and a deep understanding of number and relations between arithmetic operations. It follows that number sense is an important prerequisite for developing arithmetical competence (Desoete & Grégoire, 2006; Fuson & Burghardt, 2003; Jordan, Kaplan, Ramineni, & Locuniak, 2009). However, the concept of number sense is poorly defined and depends to some extent on the research discipline (Andrews & Sayers, 2015; Rezat & Ejersbo, 2018). In cognitive psychology and neuroscience, number sense is associated with innate competences such as approximate representation of numerical magnitude and subitizing—the precise representation of small numbers of objects (e.g. Dehaene, 2011; Feigenson, Dehaene, & Spelke, 2004). In mathematics education, number sense is related to aspects of acquired number sense—for example, recognising and naming numbers (e.g. Gersten et al., 2005) and counting skills (e.g. Jordan, Kaplan, Nabors Oláh, & Locuniak, 2006). It is important to note that it is not clear whether number understanding is innate or requires teaching; for example, precision in numerical discrimination and approximate number sense change with age (Halberda & Feigenson, 2008). For present purposes, I generally refer to number sense in the understanding of skills that requires some form of teaching.

Different aspects of number sense are known to be important for arithmetic development and specific strategy use. For example, Gaidoschik (2012) found that conceptual subitizing (Clements, 1999) at the beginning of first grade had a significant effect on strategy use by the end of first grade. Specifically, better conceptual subitizing skills were associated with more frequent use of derived fact strategy. Knowledge of number symbols and number names and the ability to map symbols to quantities have also been found important for students' development (Chu, VanMarle, & Geary, 2015; Geary, 2013).

Number comparison skills, both symbolic and non-symbolic, are associated with year four students' performance in multi-digit subtraction (Linsen, Verschaffel, Reynvoet, & De Smedt, 2015) while students with mathematics learning difficulties exhibit weaker skills in processing symbolic numerical magnitude (De Smedt, Noël, Gilmore, & Ansari, 2013)—skills that are likely to be related to strategy use (Vanbinst, Ghesquière, & De Smedt, 2012, 2014). Vanbinst et al. (2012) found that students who performed better in symbolic number comparison retrieved more single-digit addition and subtraction facts and were faster in executing direct retrieval and procedural strategies (i.e. counting and derived fact strategies).

Counting skills and understanding of quantities and relations (i.e. relational skills) are known predictors of kindergarteners' arithmetical skills and mathematical performance in year one (Aunio & Niemivirta, 2010). Although strategy development generally proceeds from counting all and counting on to fact-based strategies (e.g. Carpenter & Moser, 1984), there is some evidence that skill in counting on is not a prerequisite for developing more advanced derived fact strategies (Gaidoschik, 2012; Steinberg, 1985).

There is fairly consistent evidence regarding the relationship between estimation skills and arithmetic performance (e.g. Booth & Siegler, 2008; Gilmore, McCarthy, & Spelke, 2007). For example, Cowan and Powell (2014) found that, among third graders, estimation accuracy on the number line related to basic calculation skills (fast retrieval of single-digit addition and subtractions from numbers less than 20). They also found a relationship with number system knowledge, aligning with Laski et al. (2014), who reported that kindergarteners' use of derived fact strategies based on base-10 decomposition depends on their knowledge of base-10 number structures. While the individual components of number sense are known to be important for further mathematics development (e.g. Geary, Hoard, Nugent, & Bailey, 2013; M. Mazzocco, Murphy, Brown, Rinne, & Herold, 2013), the links between these components are crucial (Gersten et al., 2005), and their combined contribution to arithmetic skills is complex and changes with age (e.g. Lyons, Price, Vaessen, Blomert, & Ansari, 2014).

In summary, individual student characteristics such as prior knowledge and preference for and skill at using specific strategies impacts the development of arithmetic strategies. To guide interventions and teaching practices, it is important to know how, when, and to what extent these influence the development and use of specific strategies and mathematical achievement in general. As there is little existing research on this issue in a Danish context, the present thesis addresses this knowledge gap in a shorter-term study of year one students' strategy development for single-digit addition, and in a longer-term follow-up of these students' arithmetic ability in year four. The thesis also investigates the views of six year one teachers in relation to the teaching and learning of number and arithmetic.

2.5 Socio-cultural context: Instruction and culture

Socio-cultural context influences students' strategy use at two levels: 1) their general *repertoire* of strategies and 2) their *choice* of strategy in specific situations (e.g. Bjorklund & Rosenblum, 2002; Carraher, Carraher, & Schliemann, 1985).

2.5.1 Socio-cultural context and development of strategy repertoire

Different instructional practices and general cultural perspectives can influence students' development of strategies and consequently their strategy *repertoire* (e.g. Geary et al., 1996; Shen et al., 2016). There is evidence that instructional practice can influence the use and development of strategies in arithmetic (Hickendorff, Torbeyns, & Verschaffel, 2018). It has also been shown that boys and girls may respond differently to the same instructional practice (Hopkins, McGillicuddy-De Lisi, & Lisi, 1997; Timmermans, Van Lieshout, & Verhoeven, 2007). Differential response to instruction has also been found for high and low achievers (Torbeyns et al., 2005; Tournaki, 2003). Gaidoschik (2012) reported that instructional practice may enhance low achievers' development of derived fact strategies, but he found that high achievers used these strategies even when not taught explicitly. With instruction, even relatively low achieving students are able to use derived fact strategies (Torbeyns et al., 2005).

Instruction also seem to play a role in the development of number sense skills in young preschool children (aged 3.8–6.1 years). In a study of students from Finland, Hong Kong and Singapore, Aunio, Ee, Lim, Hautamäki, and Van Luit (2004) found that Asian children outperformed their Finish peers on counting skills and number sense tasks such as the ability to organise and compare quantities. They argued that this could be explained by the early focus on learning mathematics in Asian countries, where children as young as three are introduced to simple mathematics concepts. In contrast, this starts at the beginning of formal schooling in Finland (i.e. at age seven).

The evidence that students' development of arithmetic strategies differs substantially, both between individuals and between countries and cultures (Campbell & Xue, 2001; Geary et al., 1996; Hickendorff et al., 2018; Shen et al., 2016), may be attributed at least in part to differing didactic approaches in early arithmetic education. In a study on the relation between Dutch and Flemish students' strategy use for multi-digit subtraction and different instructional practices as expressed through teacher questionnaires and textbook analysis, Torbeyns, Hickendorff and Verschaffel (2017) did not find a clear relation. They suggested to complement the teacher questionnaires with classroom observation to provide more valid information on the relation between classroom instruction and students' strategy use. To my knowledge, only a few studies to date have investigated instructional practices in the classroom as a factor in students' strategy development for single-digit addition.

2.5.2 Socio-cultural context and strategy choice

Another aspect of socio-cultural context's influence on strategy use relates to the student's *choice* of strategy in a given situation. For example, students may make different choices in a test situation as compared to classroom, play or everyday settings (e.g. Bjorklund & Rosenblum, 2002; Carraher et al., 1985). Bjorklund and Rosenblum (2002) found that 6- and 7-year-olds used more sophisticated addition strategies when playing a game with their parents than when solving an addition task in a school setting. Carraher et al. (1985) studied four Brazilian children (aged 9–15) with very different levels of schooling (1 to 8 yrs.) who worked with their parents as street vendors. When presented with test items based on sales transactions familiar to them from their working life, they used mental arithmetic successfully. However, when presented with the same test items as symbolic arithmetic (i.e. without the sales transaction), they used paper and pencil algorithms and returned higher error rates.

Students' strategy choice could also be influenced by classroom practices (Torbeyns et al., 2005). They suggested that adaptive strategy choice—especially for high achievers—could be influenced by 'the explicitly taught classroom rule for strategy selection' (p. 16). If textbooks and teachers recommend that near-tie problems are to be solved using a tie strategy, the student may choose to use that strategy even though they might have been more inclined to use decomposition to 10. In this way, choice of strategy in classroom settings is influenced by how students interpret the teacher's legitimization and valuation of different strategies—that is, the classroom's socio-mathematical norms (Yackel & Cobb, 1996).

In summary, as sociocultural context plays an important role in students' development and use of arithmetic strategies, it is relevant to investigate strategy use in different cultural and educational settings. By exploring students' development and use of strategies in a Danish educational context and investigating differences between classrooms, the present study encompasses the effects of classroom culture and instructional practices.

2.6 Strategies and mathematics achievement

While students are generally known to use the whole repertoire of strategies from counting to fact-based (derived fact and direct retrieval strategies) regardless of their general mathematical ability (Torbeyns, Verschaffel, & Ghesquière, 2004), there are substantial differences between both individuals and groups in their use of different strategies. While high achievers (those with strong mathematical abilities) use fact-based strategies more often than counting, the opposite is true of low achievers (Dowker, 2014; Torbeyns, Verschaffel, & Ghesquière, 2004). These findings align with longitudinal studies that show that students with mathematical difficulties commonly use more counting strategies and fewer fact-based strategies while higher achievers shift to fact-based strategies with increasing school age (e.g. Ostad, 1997a; Vanbinst et al., 2014).

The types of strategy used to solve mental addition and subtraction problems influence the development of mathematical competence (e.g. Carr et al., 2008; Cowan et al., 2011; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Ostad, 1997a). According to Carr and Alexeev (2011), the use of cognitive strategies (i.e. calculations performed without manipulatives, including silent mental counting, derived fact, and direct retrieval) in second grade was associated with the development of mathematics competency in fourth grade, whereas excessive use of counting strategies using manipulatives resulted in poor outcomes. In a longitudinal study, Geary (2011) found that early (year one) use of decomposition and more advanced counting (frequent and accurate use of the min procedure) predicted year five mathematics achievement (measured as number discrimination, rote counting, simple and complex arithmetic and solving rational number problems).

Although less is known about the relationship between specific strategy use and specific mathematical knowledge and competence, one study (Chesney et al., 2014) examined the relationship between students use of decomposition strategies and their understanding of mathematical equivalence. Students who used decomposition to solve single-digit addition problems and were able to provide more equations for a target value (e.g. equations that equals 9 or 12) showed a better understanding of mathematical equivalence.

It seems likely that the relationship between strategy use and mathematics achievement is driven by some of the same domain-general and domain-specific skills involved in strategy use. For example, Vanbinst et al. (2014) found that students with persistent mathematical learning difficulties had lower numerical (symbolic) magnitude processing skills, which seem to be related to strategy use.

As outlined above, there is a well documented link between strategy use and later mathematics achievement. However, less is known about the relationship between strategy use and specific aspects of mathematics knowledge and competence. The present study analyses the relationship between year one students' strategy use in single-digit addition and their level of arithmetic proficiency in year four, including proficiency with fractions and word problems.

2.7 Sex differences

In principle, sex can be considered a ‘framing’ variable, in the sense that it embeds a multitude of distinct characteristics as underlying factors (including both biological and socialized differences related to gender identity; see Halpern et al. (2007) for a comprehensive review). Any or all of these may contribute to sex differences in mathematics achievement and strategy use, either individually or as interactions between factors. As biological differences can be assumed to be the same across cultures, it seems likely that differences in boys’ and girls’ mathematical performance are at least in part culturally driven.

From a didactic perspective, sex differences are primarily of interest as indicators of the underlying causes of this variation. As teaching must always relate to the individual child, a better understanding of the variations among students is needed to meet individual needs. The aim here is not to distinguish girls and boys as different but to understand and acknowledge the variation that may exist between the sexes.

2.7.1 Sex differences in mathematics achievement

Sex differences in mathematics achievement have been the subject of extensive research in psychology and education (e.g. Carr et al., 2008; Fennema et al., 1998; Lachance & Mazzocco, 2006; Shen et al., 2016). However, findings regarding the nature of sex differences and their possible causes and implications are inconsistent. Indeed, there is some doubt about whether sex differences exist at all (Hyde, 2005), and about the effect and magnitude of any such differences on mathematics achievement, both in general and in relation to other variables. While there is some evidence of general sex differences, with boys outperforming girls on average, the relevance of these findings is questionable because of the relatively small effect sizes compared to other components of variance (Hyde, 2005; Hyde, Fennema, & Lamon, 1990; Penner & Paret, 2008; Shen et al., 2016) and the evidence of male overrepresentation in both high-achieving and low-achieving groups (Penner & Paret, 2008). Additionally, national and cultural differences have been demonstrated in both large-scale and small-scale studies. For example, in a study of TIMMS and PISA results, Else-Quest, Hyde and Linn (2010) found very low effect sizes for sex differences in general but very high variability in effect sizes between countries.

In their analysis of PISA results over 10 years, Stoet and Geary (2013) found a persistent inverse relationship between sex differences and national average performance in mathematics and reading in individual countries: the larger the sex difference in mathematics, the smaller the sex difference in reading. This pattern was identified both between and within countries. Sex differences in mathematics were generally small and almost non-existent at lower achievement levels. In a more recent study of PISA data, Stoet and Geary (2015) found that in 70% of countries, girls on average outperformed boys on general achievement scores (mathematics, reading and science), but top-performing boys scored higher than top-performing girls. At the individual level, girls scored higher in reading than in mathematics, and individual boys scored higher in mathematics than in reading. These patterns were unrelated to measures of gender equality. Of special interest in this context is the finding that the Nordic countries score very high on gender equality measures but also tend to exhibit wider gender gaps than other OECD countries (with the exception of Denmark and Sweden).

2.7.2 Sex differences in strategies in arithmetic

There are numerous documented cases of sex differences in strategy use for mental single-digit addition (e.g. Bailey et al., 2012; Carr & Alexeev, 2011; Carr & Davis, 2001; Carr et al., 2008; Gaidoschik, 2012; Geary et al., 2012; Imbo & Vandierendonck, 2007; Shen et al., 2016), with only a few studies reporting no differences between boys' and girls' strategy use (e.g. Dowker, 1998, 2009). However, any observed sex differences appear consistent across studies; specifically, girls seem to count more than boys while boys use direct retrieval and derived fact strategies more often than girls (Bailey et al., 2012; Carr & Alexeev, 2011; Carr & Davis, 2001; Gaidoschik, 2012; Imbo & Vandierendonck, 2007; Shen et al., 2016).

Carr et al. (2008) found that boys are more likely to use cognitive strategies for multi-digit arithmetic while girls are more likely to use manipulative strategies, with significant implications for mathematical competence. These differences are persistent in lower grades (Fennema et al., 1998); Carr and Alexeev (2011) found that sex and accuracy predicted initial strategy use in second grade but not the rate of growth in strategy use as predicted by fluency.

2.7.3 Sex differences in cognitive factors

Sex differences in mathematical ability related to number and arithmetic can be linked to underlying differences in cognitive abilities. For example, there are unambiguous sex differences in the relationship between visuospatial skills and mathematics performance among younger primary children. For matching figures tasks, however, there are no obvious sex differences among first and second graders in the relationship between performance and multi-digit number processing (Krinzinger, Wood, & Willmes, 2012), and Lachance and Mazzocco (2006) found that the relationship between cognitive factors and different measures of mathematical and arithmetical ability was only slightly stronger for primary school age boys.

One visuospatial skill that reveals consistent sex differences is mental rotation. Male 2-D mental rotational skill is better among children as young as 4½ years (Levine et al., 1999), and this remains the case among college students. For example, among high-ability (but not low-ability) college students, Casey, Nuttall, Pezaris, and Benbow (1995) found that males outperformed females on both 3-D mental rotation tasks and the Scholastic Aptitude Test–Math (SAT-M), but sex differences in SAT-M were eliminated when the results were adjusted for mental rotation ability.

On the other hand, sex differences in cognitive factors are not necessarily linked to sex differences in calculation performance and number competences. Although Moè (2018) found that eight- and ten-year-old boys scored higher than girls on mental rotation (2-D and 3-D), and that mental rotation was related to mathematics ability in ten-year-olds (but not eight-year-olds), there were no sex differences in students' numeric and arithmetic ability. Although there are some reports of sex differences in these cognitive abilities, these do not necessarily entail sex differences in arithmetic competence.

Sex differences have also been reported in approaches to wayfinding and orientation. For example, when finding their way in a labyrinth (Pintzka, Evensmoen, Lehn, & Håberg, 2016) or on a computer game map (Coutrot et al., 2018), males tend to orientate themselves according to ‘landmarks’ while women tend to follow a ‘route’ or series of directions (e.g. ‘left, then right’). Pintzka et al. (2016) found that when treated with small doses of the male hormone testosterone, healthy women’s approach to orientation became more like men’s. This indicates that sex is a ‘marker’ or ‘framing variable’ of some biological component—in this case, testosterone—that influences their behaviour. This does not mean that the specific behaviour is innate or predisposed and cannot be changed or developed through training, as several factors are involved. For example, Coutrot et al. (2018) found that the higher a country’s gender equality, the better women performed. However, the overall finding that women primarily use a ‘serial strategy’ is interesting in the present context, as girls also seem to prefer ‘serial strategies’ (e.g. counting) when performing single-digit addition.

2.7.4 Sex differences in number sense

As different aspects of number sense are important for the development of arithmetic skills (Geary et al., 2013; Lyons et al., 2014; M. Mazzocco et al., 2013), it seems likely that sex differences in these components may influence differences in arithmetic proficiency. In a study of 1,391 children in years 1–6, Hutchison, Lyons, and Ansari (2019) analysed students’ answers to a wide range of numeric and arithmetic tasks. They found convincing evidence of boys’ superiority in number-line estimation (1–100 and 1–1000), and of girls’ superiority in quick and accurate counting of dots (1-9). As these differences decreased with age, Hutchison et al. (2019) concluded that ‘for the most part, gender does not influence basic numerical processing’ (p. e75). However, one could argue that the findings of better male performance on number-line estimation tasks aligns with evidence of better male visuospatial skills (e.g. Moè, 2018). The finding that girls outperform boys in quick and accurate counting of small numbers of dots (Hutchison et al., 2019) may relate to frequently observed sex differences in single-digit addition strategies, where girls use counting more often than boys (e.g. Bailey et al., 2012; Carr & Alexeev, 2011; Carr & Davis, 2001; Shen et al., 2016).

In relation to cognitive factors, there is divergent evidence of sex differences in number sense. Among younger children, it seems there are no such differences (Aunio et al., 2004), but it is not uncommon to observe sex differences in later years. While these may be a result of social interactions, Halpern et al. (2007) noted that the causes may just as well be biological in light of the developmental timeline.

2.7.5 Sex differences as a product of socio-cultural context

Reported sex-related differences in responses to instructional practices are somewhat inconsistent. There are some reports that low-performing girls improved following guided instruction in an intervention to explore different strategies (Timmermans et al., 2007). However, the opposite effect has also been observed—for example, Hopkins et al. (1997) reported that girls benefit from direct instruction in the use of rules and algorithms. Several other studies report that differences in interactions with adults (teachers or parents) can have differential effects on boys’ and girls’ performance. For example, Chang, Sandhofer and Brown (2011) found that mothers engaged in more number-related talk with 22-month-old boys than with girls of the same age. Carr, Jessup and Fuller (1999) found that boys’ perceptions of parents and teachers’ beliefs influenced strategy use, and that teachers were more likely to teach retrieval strategies to boys.

Cultural differences in sex effects have also been reported in studies of different ethnic (cultural) and socio-economic groups (e.g. Penner & Paret, 2008). While Stoet and Geary (2015) found that sex differences in general PISA achievement were unrelated to national measures of gender equality, there is evidence that sex differences are to some extent dependent on sociocultural factors (Hyde & Mertz, 2009; Penner & Paret, 2008) and educational practice. For example, in a study of three very different cultural and educational contexts, Shen at al. (2016) found sex differences among American and Russian first graders' addition and subtraction strategies but not among Taiwanese students.

To sum up, sex differences in mathematical achievement and arithmetic strategies are widespread and well documented, but the underlying causes (biological, cultural or interactive) remain largely unknown. This incomplete understanding should in itself stimulate scientific research to enhance our understanding of how students think and learn. From a pragmatic and analytical viewpoint, the fact that sex differences often appear when investigated suggests that educational studies should always take account of this issue.

2.8 Summary

Mental addition strategies can be categorised in terms of various forms of counting, direct retrieval and derived fact. Strategy use and development is influenced by task characteristics, individual student characteristics and the socio-cultural context in which strategies are acquired or executed. Arithmetic strategies are important for students' further development of arithmetic competence and more general mathematical competence.

In the present context, socio-cultural and individual factors are of particular interest, as is the relationship between students' early strategy use and later mathematical achievement. Many factors influence the development of strategies in arithmetic; these relate to working memory, number sense or knowledge and spatial ability. While some can be attributed to biological causes and others to cultural and social influence, the 'nature versus nurture' issue is beyond the scope of this thesis. What matters here is that these factors influence the development of strategies to differing degrees at different points in time. The role of sex as one such factor indicates a biological component. From a didactical viewpoint, the point is not whether the student is a boy or a girl but that the focus must be on the individual and how teaching can facilitate learning for that student.

3. Aim and research questions

The overarching aim of this study is to investigate patterns in Danish students' use and development of strategies for single-digit addition in relation to 1) different teaching practices, 2) later mathematical achievement and 3) teachers' perspectives on teaching and learning of number and arithmetic.

To that end, these complex issues are operationalised as six research questions, which are addressed in the five individual papers (I-V).

1. How does strategy use in single-digit addition develop with age, and is there any systematic between-subject or between-class variation in patterns of strategy use among boys and girls? (Paper I)

2. To what extent does the individual student's strategy use develop during a half-year of formal mathematics teaching from their baseline level at the beginning of year one? (Paper II)

3. How much do the different components of variance (initial strategy use, time, teaching practice) contribute to the overall variation in year one students' development of strategies? (Paper II)

4. How does strategy use among year one students relate to later mathematics achievement (assessed in year four), and does information about strategy use in year one provide additional predictive information beyond a standardised test of mathematical achievement? (Paper III)

5. What do teachers expect their students to know about number and addition by the end of year one? (Paper IV)

6. To what extent do teachers' perspectives on the teaching and learning of number and addition in year one align with established knowledge concerning the foundations of numerical and arithmetic competence? (Paper V)

Because this project is problem-driven (as outlined in Chapter 1), the research questions guided the methodological choices. For that reason, the first four research questions (which relate to students) and the last two (which relate to teachers) require different quantitative and qualitative data and analyses and a mixed methods approach. As shown in figure 3.1, the research draws on three different data strands: assessment of students, teacher interviews and classroom observations. Further details of design and methods are provided in Chapters 4 and 5.

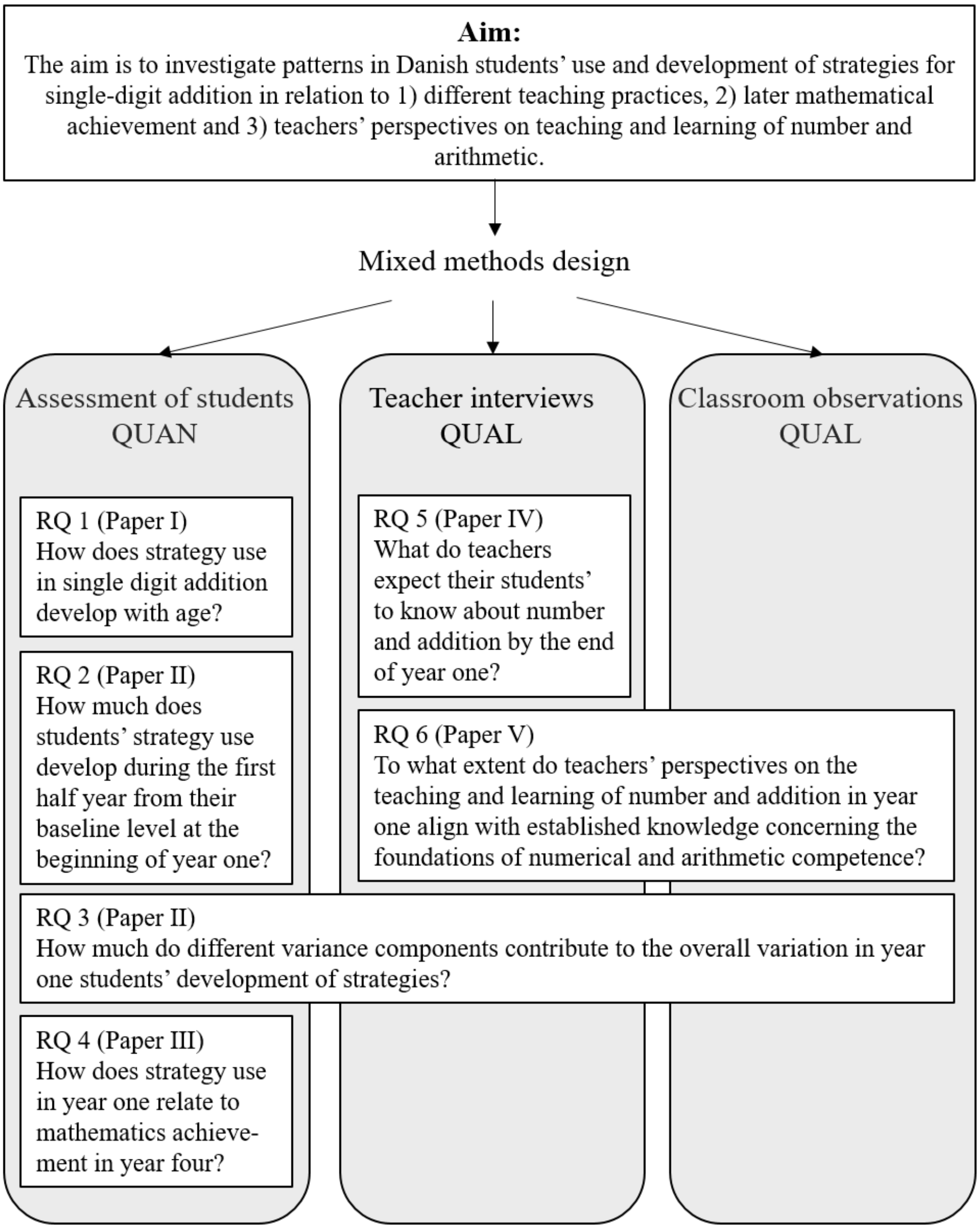


Figure 3.1 Overview of how three data strands in the mixed methods approach inform the research questions.

4. Methods

In this chapter, I present my methodological choices. First, I review the overall structure of the project and how quantitative and qualitative approaches were applied. Next, I describe the design of the project's two studies. Each study's sampling methods are presented, and the overall principles for how they were analysed are described. Finally, I present some ethical considerations that emerged during the project.

4.1. Structure of the project

The project consisted of data from two independent studies, denoted Study A and Study B in the following.

Study A consisted of a quantitative (statistical) investigation of the frequency with which different strategies are used to solve single-digit addition problems in years one to four as a function of age and sex when controlling for variation between classes and individuals. These data were collected in the years 2012 to 2016 in a rural school where I worked as a teacher. This initial general study of strategy use in single-digit addition formed the basis for my subsequent works executed in Study B.

Study B represents a more narrowly focussed, in-depth investigation aimed at elucidating how students' single-digit addition strategies developed through year one, what the teachers' educational focus and practices were, and whether these were interrelated. Study B initially ran for a single school year (2015–16), when the aforementioned topics were investigated. In the course of time, however, an extension to this study appeared in the form of test results for mathematical achievement in year four (school year 2018–19) that could be related to the individual student's strategy use profiles in year one, three years earlier.

The overall analytical design was a parallel mixed methods design where quantitative and qualitative data are collected concurrently to inform different aspects of the overall research question. Qualitative data provide in-depth understanding and "a narrative understanding to quantitative research findings" (Hesse-Biber, 2010 p. 6). Furthermore, "quantitative data can be useful for establishing generalizability of qualitative results" (Hesse-Biber, 2010, p. 6). Therefore, quantitative and qualitative data were analysed, and the results were reported and discussed independently because they addressed different parts of the overall research question. This was conducted using the concurrent embedded approach described by Creswell (2009). The independent findings were then compared and contrasted in order to provide an overall composite assessment of the problem and gain a broader perspective on the development of strategies for single-digit addition.

4.1.1 Design

As described above, this thesis consists of data from two individual studies. An overview of the data is provided in Figure 4.1. Details for each study are presented below.

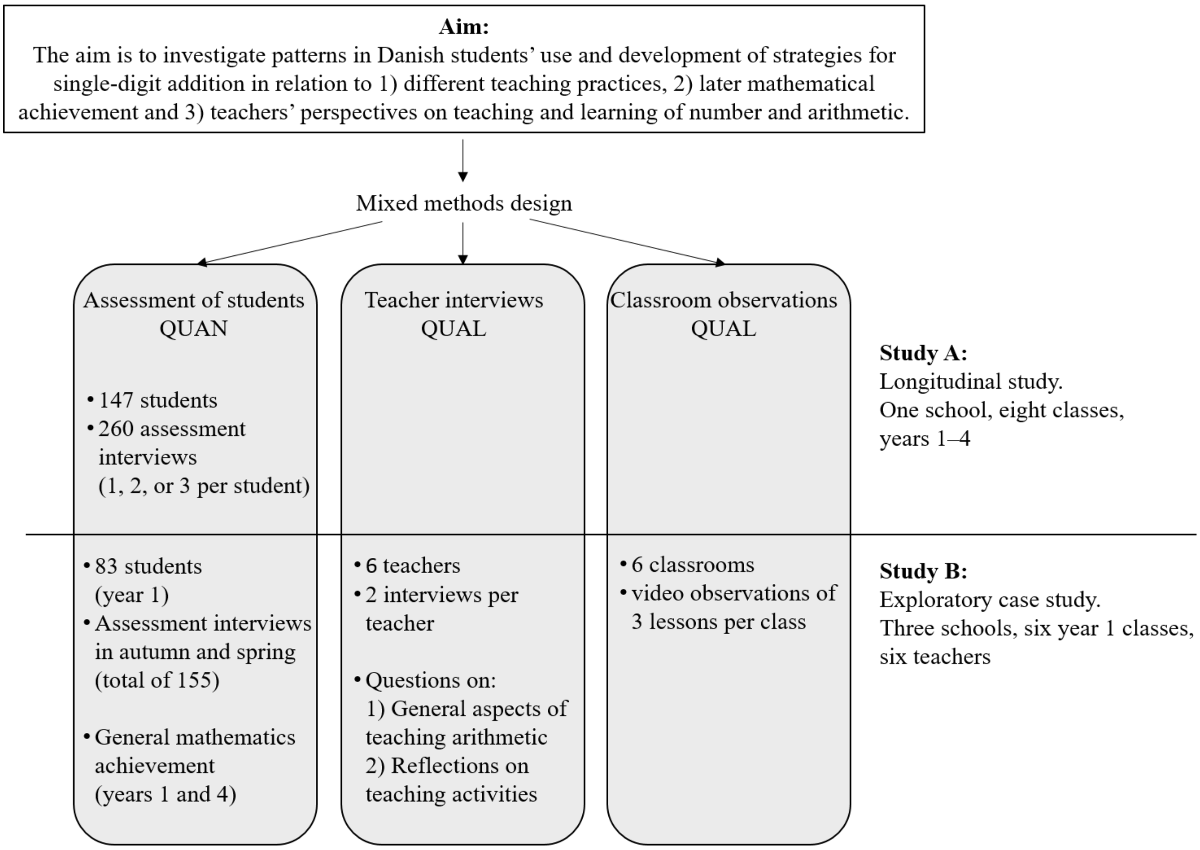


Figure 4.1. Overview of the data from Studies A and B.

Study A

In Study A, I followed the development of mental strategies in addition from year one to year four with 147 students (77 girls, 70 boys) from eight classes in a Danish primary school (340 students) in a rural district. The majority of students in each class were interviewed (Table 4.1). Three classes were tested thrice in subsequent grades, two classes twice, and three classes once (in total, 260 assessment interviews with 147 students). Since the majority of the students were tested more than once, the data include both cross-sectional information (information on variation between individuals and classes when the age factor is held equal) as well as longitudinal data (development with time or school age within the same individual).

Longitudinal data in this context means the data (strategy use) for a cohort of students were collected over time (Cohen, Manion, & Morrison, 2011; Ruspini, 2002). Longitudinal data makes it possible to analyse and describe changes over time as well as explain the changes in terms of other characteristics such as gender (Ruspini, 2002). The cross-sectional data resemble the longitudinal data in that the nature and rate of development of a number of variables were assessed. However, the data were drawn from samples of students at different ages and/or time points (Cohen et al., 2011). Longitudinal data can therefore provide information about development and change at an individual level, whereas cross-sectional data can provide information at a population level (Ruspini, 2002).

Table 4.1 Study A: Number of students assessed per class (Girls/Boys) by year, and mean dates of assessment interviews (Paper I).

	Number of assessment interviews conducted per year										Students in total	
	Year 1		Year 2		Year 3		Year 4		Total			
Class	G/B	Mean date	G/B	Mean date	G/B	Mean date	G/B	Mean date	G/B	All	G/B	All
A1					11/6	2-Nov-13	13/6	30-May-15	24/12	36	13/6	19
A2			9/4	31-Aug-12	12/4	28-Oct-13	14/5	22-Apr-15	35/13	48	14/5	19
A3	0/1	1-Nov-12	11/6	29-Sep-13	14/8	21-Apr-15			25/15	40	14/8	22
A4	13/7	21-Feb-14	8/7	10-Apr-15	8/6	2-Jun-16			29/20	49	8/7	15
A5			7/10	1-Apr-15					7/10	17	7/10	17
A6	6/10	26-Apr-15	6/9	6-Jun-16					12/19	31	6/10	16
A7	6/9	21-Apr-15							6/9	15	6/9	15
A8	9/15	6-Oct-15							9/15	24	9/15	24
Sum	34/42		41/36		45/24		27/11		147/113	260	77/70	147

Study B

Study B is a short longitudinal study where I followed six year one classes, including their teachers, and observed lessons from October to April. This provided me with in-depth information about the teaching and learning of number and arithmetic in year one. This is characteristic of an exploratory case study (Yin, 2013), which allows for an “in-depth exploration from multiple perspectives of the complexity and uniqueness of a particular . . . system in a ‘real-life’ context” (Simons, 2009, p. 21). For this research, the case study approach provided the ability to investigate the close relationships between teachers, students, and learning in a natural setting.

In advance of the study, I sent invitations to participate to school administrators and teacher consultants. After school administration gave permission for the project, I contacted the relevant teachers and asked if they wanted to participate in the project with their year one classes. The final sample comprised six classes; three from one city school (580 students), two from one town school (370 students), and one from a rural school (390 students) (Table 4.2). The schools were compensated economically for the extra time the teachers spent on the project, for example, their time for the interviews and meetings.

Table 4.2 Study B: Participating schools and teachers.

School	Teacher (pseudonym)	Gender	Age	Math-specific teacher education	Teacher years of experience (first year teaching)
R: Rural school	Allan	Male	37	No	2 years (2014)
T: Town school	Bettina	Female	49	Yes*	24 years (1991)
	Carl	Male	40	Yes	15 years (2000)
C: City School	Dan	Male	45	Yes*	20 years (1995)
	Else	Female	37	Yes	9 years (2006)
	Frida	Female	30	Yes	5 years (2010)

**The teacher education at the time only permitted teachers to teach mathematics at lower levels.*

All six teachers (three males, three females) held certificates for the Danish four-year teacher education, but the certificates varied in their specificity to math education. Their professional experience ranged from 2 to 24 years (Table 4.2). The six teachers, for whom I have used pseudonyms throughout, were interviewed in October 2015 and January 2016 about the teaching and learning of addition in year one. Classroom teaching was observed on videorecordings three times, in October 2015, December 2015, and January 2016. Excerpts from the videos were used for a stimulated recall interview in the last interview session (Figure 4.2).

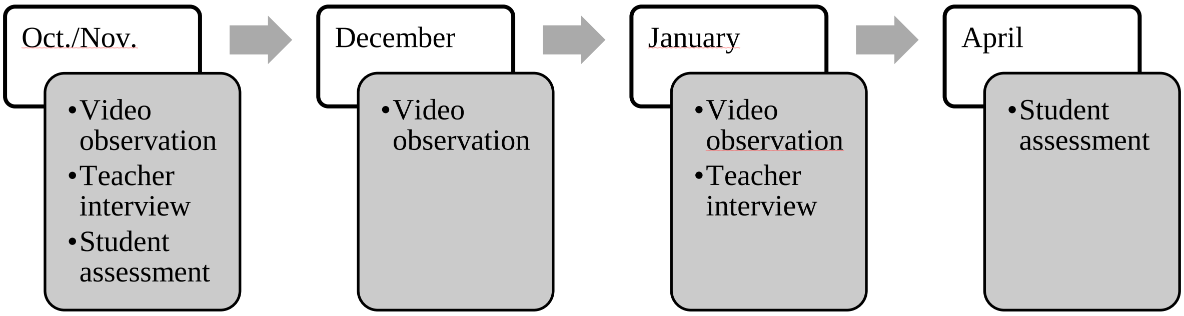


Figure 4.2. Timeline for data collection in 2015, Study B.

The students' strategy use for mental addition and general mathematics achievements were assessed in November 2015 and April 2016 (Figure 4.2; Table 4.3). This set-up provided repeated measurements separated by a time span of five months. Because the first assessment took place less than three months after the start of year one (the Danish school year runs from mid-August), the five-month interval between the first and second assessments represented approximately half of students' time spent being formally educated in mathematics in year one. For practical reasons (each student had to be interviewed separately), it was not possible to assess strategy use in all students. Therefore, to obtain a random sample for the assessment interviews on strategy use, between one-half and two-thirds of the students in each class were randomly selected, in roughly equal numbers by sex. The resulting total sample was 83 students, divided into 46 girls and 37 boys (Table 4.3). These data were used to both validate findings in Study A and to investigate short-term longitudinal development within a single school year.

Concurrent with the strategy use assessment interviews, all students were tested for general mathematical skills (see 4.2.2 for further details). As this test was conducted class-wise, a sub-sampling of students was not necessary, and all of the students who were present on the test days were included.

Table 4.3 Study B: Number and mean dates of assessment interviews with students (Girls/Boys) (Paper I).

	Number and date of assessment interviews:						Students in total	
	Round A		Round B		Total			
Class	G/B	Mean date	G/B	Mean date	G/B	All	G/B	All
B1	9/6	5-Nov-15	8/5	14-Apr-16	17/11	28	9/6	15
B2	5/5	3-Nov-15	7/5	20-Apr-16	12/10	22	6/5	11
B3	7/5	2-Nov-15	7/5	14-Apr-16	14/10	24	7/5	12
B4	6/5	10-Nov-15	7/6	15-Mar-16	13/11	24	8/6	14
B5	7/6	5-Nov-15	6/6	4-Apr-16	13/12	25	7/6	13
B6	9/9	26-Oct-15	8/6	29-Mar-16	17/15	32	9/9	18
Total:	43/36	2-Nov-15	43/33	5-Apr-16	86/69	155	46/37	83

4.1.2. Information about consent

For all studies, written parental consent was obtained in advance for each student. The declaration of consent to be signed was attached to an informational letter that informed the parents about the general aim of the study, the planned tests and observation activities, the types of data that would be collected for each child, and how the data would be used and stored. In Study A, the parents were instructed to only give consent for their children's participation in the strategy use assessment and were informed that the data would be used in an anonymized form for scientific studies. In Study B, in addition to everything above, parents were instructed to additionally give specific consent to each of the following individually: 1) that their child could participate in the assessments, 2) that their child could participate in the videorecorded lessons, and 3) that the data for their child could be used for research purposes. The parents of one student did not give consent, so in collaboration with the teacher and parents, this student did not participate in the assessments or the lessons that were videorecorded.

Prior to the assessment interviews, the students were orally informed about the aim of the activities (that I was investigating how young students solved addition questions), that participation was voluntary, and that anyone could withdraw from the testing session at any time if they changed their minds. In those cases where the assessment interviews were audio recorded, students were informed that the reason for recording the interviews was to be able to recall the conversation without uncertainty. The students were explicitly asked whether they were comfortable with the interview being recorded or if they would prefer the audio recorder be turned off. In a few cases, the student preferred to not have the assessment interview audio recorded, and as a result, those interviews were not taped.

In Study B, prior to giving their consent for participation, the teachers were first informed about the project in writing and were subsequently informed again at an introductory meeting where they could also ask their own questions about the project in general and their intended role. At the end of the project, all of the teachers were invited to read and approve the transcripts and analyses of the interviews. None of the teachers took advantage of this offer.

4.2. Data collection

4.2.1 Assessment of students' strategy use in single-digit addition

The addition tasks

Strategy use was scored in one-to-one assessment interviews by sequentially presenting the student with 36 flash cards, each of which showed a single-digit addition problem using the addends 2 to 9 (Table 4.4). The 36 tasks comprised all of the possible combinations of the addends 2 to 9, including tie sums, but only one of each pair of commutative tasks. The tasks were selected to ensure equal numbers of tasks with the larger addend first, and vice versa. The 36 tasks used in the assessment interview were selected from the standard set of single-digit arithmetic tasks (LeFevre et al., 1996). Tie sums were included as they are important in derived fact strategies. Often, tie sums are excluded (Vanbinst et al., 2014). However, excluding tie sums constrains the ability to perform detailed analyses on the development of task-specific strategy use.

To prevent situations where less confident students would give up or, despite other strategies being available, overly rely on counting if presented with tasks that were too difficult at the beginning of the interview, the students would then first be presented with the easiest addition tasks that had sums less than 10. Subsequently, they were presented with tasks with increasing sums and difficulty in the order presented in Table 4.4. The students were not provided with any manipulatives or paper and pencil, but were allowed to use fingers.

Table 4.4 Single-digit addition tasks for mental strategy assessment.Single-digit addition tasks used in assessing student's mental strategy use. The tasks were presented one at a time on flash cards in the indicated order.

Flash card no.	Addition task	Flash card no.	Addition task	Flash card no.	Addition task
1	2+2	**13**	5+5	**25**	8+7
2	3+4	**14**	3+8	**26**	4+9
3	2+3	**15**	7+5	**27**	7+7
4	4+5	**16**	4+6	**28**	8+5
5	6+2	**17**	9+3	**29**	9+6
6	4+2	**18**	6+6	**30**	6+7
7	3+6	**19**	7+3	**31**	9+9
8	4+4	**20**	5+6	**32**	9+7
9	2+5	**21**	8+2	**33**	6+8
10	3+3	**22**	7+4	**34**	8+8
11	5+3	**23**	2+9	**35**	5+9
12	2+7	**24**	8+4	**36**	9+8

One-to-one assessment interviews

The interviews with students were performed in a quiet room at the school. Each interview lasted between 10 and 30 minutes depending on the grade and the individual student. The students knew me as either a teacher at the school (Study A) or through the classroom observations (Study B). On a few occasions in Study B, the student's mathematics teacher would be present: Dan, Else, and Frida attended some student assessment interviews in November 2015, and Allan, Bettina, and Carl attended some student assessment interviews in April 2016. The students appeared to behave exactly the same in the interviews whether I interviewed them alone or was accompanied by their teacher.

Before I presented the students with the flash cards, I said, "I am investigating how young students like you do addition. I will show you some single-digit addition tasks. First, I would like you to work out the answer to the task, and then we will talk about how you found the answer. There are many ways to find the answer to an addition task. Sometimes, you might know the answer or count, or perhaps you use other tasks to find the answer. I am interested in knowing how you find the answer". When I presented a student with a flash card, e.g. 4 + 5, I first asked, "What is the answer to four plus five?" Often, the student automatically explained to me how she or he found the answer. When that was not the case, I then asked, "How did you find the answer?" and if further prompting was needed, "Did you count, did you just know the answer, or did you use some other tasks you know to find the answer?" If the student miscalculated, I asked the student to show me once more how she or he found the answer. In most cases, the student then corrected the answer. Scoring of the strategy category was based on the assessment scheme described by Geary et al. (1996) and others (Vasilyeva, Laski, & Shen, 2015). This means that the student's verbal explanation or observed procedure was scored immediately after each addition task. In case of disagreement between the student's self-report and an observed overt behaviour, the observed strategy category was chosen.

Student self-reports have been used widely in studies of strategy use. Several studies have demonstrated the validity of students' self-reported strategy use by comparing with behavioural observations and/or reaction time (Canobi, Reeve, & Pattison, 1998; Rittle-Johnson & Siegler, 1999; Siegler, 1987, 1989). This has also been demonstrated for adults (e.g. LeFevre et al., 1996), and recent neuroscientific evidence supports the validity of self-report as a method of assessing strategy use (Grabner & De Smedt, 2011). In a study on students' spelling strategies, Rittle-Johnson and Siegler (1999) found that basing strategy assessment on observed behaviour and verbal reports yields a more accurate assessment as compared to relying on only overt behaviour or verbal report. Furthermore, achievement level does not seem to influence students' awareness of their own strategy use or their ability to report it (Wu et al., 2008).

Danish students are familiar and confident with having open discussions in the classroom. Therefore, it never seemed difficult or uncomfortable for the students to explain their procedures, and accordingly, scoring on the basis of student self-report was usually straightforward.

Strategy use was assessed in a free-choice situation—that is, the student was free to choose any strategy to solve the specific task. An alternative approach often used to assess a student's actual repertoire is the choice/no-choice method, which can provide detailed information about adaptive strategy choices (Siegler & Lemaire, 1997; Torbeyns, Verschaffel, & Ghesquière, 2005). Using this approach, students would be tested in both free-choice and no-choice situations (where the student must apply a specific strategy); the number of no-choice situations equals the number of strategies investigated. In the present study, I employed four different categories of strategy (counting all, counting on, direct retrieval and derived fact), which would mean a total of five test situations for each assessment time point. As one test situation in the present study took between 10 and 30 minutes, it would not be possible to complete five such assessments for each child within a reasonable timeframe. These limitations of the choice/no-choice method have been discussed by Torbeyns and colleagues (Torbeyns, Arnaud, Lemaire, & Verschaffel, 2004; Torbeyns, Verschaffel, & Ghesquiere, 2004).

With regard to the student interviews, I conducted an informal assessment of interrater reliability beforehand. In developing this assessment procedure, I observed another person interviewing and scoring the children. Subsequent comparison of category scores revealed a maximum of 1 inconsistency per 36 items for the majority of the test sessions (a rate of less than 3%). In most cases, the majority of inconsistencies related either to the two different counting strategies or to fact-based strategies. This leads me to believe that the observed inconsistencies are unlikely to have influenced the overall pattern. As most of the student assessment interviews in Study B were recorded on tape, these data are available for future analysis in this regard.

Strategy categories

Many studies have categorized strategies in counting, derived fact, and direct retrieval; however, the level and number of subcategories can differ. Vasilyeva et al. (2015) used *counting, decomposition, retrieval,* and *other,* whereas Geary et al. (1996) further distinguished between three subcategories of counting (counting fingers, fingers, and verbal) and three subcategories of decomposition (10-based, tie-based, and other). In this study, the solving strategy for each addition problem was scored in one of the seven categories: 1) Gives up, 2) Miscalculates, 3) Counting all, 4) Counting on, 5) Direct retrieval, 6) Derived facts +, and 7) Derived facts −. Description of the strategy categories and examples of students' answers are given in Table 4.5. If the students miscalculated and subsequently corrected the answer, the strategy used for the correct answer was scored. Attempted strategy use for incorrect answers were not scored.

Table 4.5 Definition of the seven categories of strategy use with examples of students' explanations of how they found the answer to a specific single-digit addition task and the strategy category the answer was scored in. 'She' is used for convenience but represents both boys' and girls' answers. (The table continues on the following page).

Strategy category	Description	Examples of student explanations
1. Gives up	The student has no strategy and gives up.	The student cannot provide an answer and responds, "I don't know". The interviewer asks if the student could find the answer by counting or maybe by using a task she already knows the answer to, but the student says, "I don't know how to do it".
2. Miscalculates	The student miscalculates without noticing and does not correct the answer when asked to try again.	The student answers with an incorrect sum. The interviewer asks if the student can explain how she found the answer, and the student responds by saying, "I just know the answer". Upon realizing that it was wrong, she provides another wrong answer.
		The student miscounts, e.g. with 3 + 4, the student counts 3-4-5-6. The interviewer asks how she found the answer and asks the student to try again. The student still miscalculates.
		The task is 8 + 7, and the student says, "8 + 8 is 18 and then minus 1 is 17". The interviewer asks the student to explain how she got the answer and how she finds the sum of 8 + 8. The student answers, "I just know" and repeats the explanation.
3. Counting all	The student counts both addends and then all, or represents both addends with fingers, with or without counting.	The task is 3 + 4, and the student counts "1-2-3" on the fingers of one hand, "1-2-3-4" on the finger of the other hand, and finally, "1-2-3-4-5-6-7".
		The student's lips are moving as if counting 1-2-3, 1-2-3-4, 1-2-3-4-5-6-7, and when the student is asked to explain how she counted, the student's explanation is in agreement with the observation.
		The student solves the task 8 + 4 using three fingers to represent 8, says "8" and counts "1-2" on the same hand and then "3-4" on the other hand, then looks at both hands and says "12".
		The task is 3 + 4, and the student shows three fingers on one hand and four fingers on the other hand, and without visibly or audibly counting, says "7".

Strategy category	Description	Examples of student explanations
4. Counting on	The student counts on from one of the addends (there is no distinction between counting from the first or a larger addend).	The student solves 7 + 3 by saying “7” and counting on three fingers “8-9-10”.
		The student solves the task 7 + 3 by saying “8-9-10” or “7” (pause) “8-9-10”.
		When presented with the task 4 + 9, the student says “13”. When asked how she found the answer, the student explains “I counted on from 9 in my head. I said 9, and then 10-11-12-13”.
5. Direct retrieval	The student knows the answer.	For the task 7 + 3, the student immediately responds “10—I just know”.
		The student hesitates with the answer, but when asked, she responds, “I know the answer; it just took some time to find it in my brain”.
6. Derived facts +	The student decomposes the addends and subsequently uses addition.	When answering 5 + 6, the student says “5 + 5 is 10 and then one more is 11”.
		For the task 8 + 5, the student says “ I just take 2 to the 8 and then it is 10+3, that is 13”.
7. Derived facts −	The student decomposes the addends and subsequently uses subtraction.	When answering 8 + 7, the student explains “It is just 1 less than 16 because 8 + 8 is 16, so it is 15”.

To ensure a sufficient number of observations in each strategy use category in the statistical analyses for Papers I and II, the initial seven strategies were reduced to the following four (Table 4.6): 1) *Error*: the student gives up or repeatedly miscalculates, 2) *Counting*: all varieties of counting procedures, including finger counting, verbal, or self-report of mental counting, 3) *Direct retrieval*: the sum is automatized, and 4) *Derived facts*: addends are decomposed and automatized sums are used to calculate the answer (e.g. 4 + 5 = 4 + 4 + 1 or 5 + 5 − 1).

Table 4.6 Categories for scoring students’ strategy use and pooled categories for Paper I and Paper II analyses.

Strategy	Description: The student . . .	Pooled categories
1. Gives up	has no strategy and gives up	**Error**
2. Miscalculates	miscalculates without noticing	
3. Counting all	counts both addends and then all, or represents both addends with fingers	**Counting**
4. Counting on	counts on from one addend	
5. Direct retrieval	knows the answer	**Direct retrieval**
6. Derived facts +	decomposes the addends and uses addition (e.g. 4 + 5 = 4 + 4 +1)	**Derived facts**
7. Derived facts −	decomposes the addends and uses subtraction (e.g. 4 + 5 = 5 + 5 − 1)	

4.2.2 Assessment of students’ mathematics achievement

Mathematical achievement, year one

General mathematics skills were tested with MAT1 (Jensen & Jørgensen, 2007). MATs are standardized norm-referenced tests commonly used for diagnostic testing in Danish primary schools. MAT tests a wide range of mathematical skills within the areas of numbers and algebra, geometry, and applied mathematics. The test is conducted as a paper and pencil test for the whole class, and it is adapted to the curriculum for each school year (separate tests for each year). The test for year one consists of simple calculation tasks (addition and subtraction), tasks on number knowledge and the number system, problems involving geometric shapes, measurements of length and area, and statistics problems. The number of items within each topic varies from 1 to more than 10. The test is not time restricted, so time is not monitored. For a further description of MAT tests, see Sunde and Pind (2016).

Mathematical achievement, year four

Mathematical achievement in year four was assessed in a computerized test consisting of three subtests, two of which were time restricted at 10 minutes for each test (number and arithmetic, proficiency with fractions, and algebraic understanding), and one that was not time restricted (word problem solving). The teacher administered the test in the classroom, and the students used either tablets or computers. All students were familiar with using computers or tablets for testing as well as in everyday lessons. The students were tested at the beginning of year four. The subtest number and arithmetic (NUMAR) consisted of 40 arithmetic tasks in the four operations and number knowledge, whole number, and decimal number. The second subtest, proficiency with fractions (FRAC), consisted of 36 items comparing fractions and with addition and subtraction using fractions. The non-timed subtest, word problems (WP), consisted of 20 word problems in the four operations and fractions. With this test, the students were allowed to use a calculator. For all three subtests, the test score was calculated as the total number of correct answers.

4.2.3 Teacher interviews

The teacher interview consisted of two parts (see Appendix A for the full interview guide in Danish). The first part, lasting 5 to 15 minutes, was a timeline interview (Adriansen, 2012) with the purpose of capturing the teacher participants' professional perspectives (background, progression, and standpoint), and this part also contributed to establishing a relaxed and trusting atmosphere. The second part was a semi-structured interview (Brinkmann & Kvale, 2015) using an interview guide inspired by Ball (1988) and Ma (2010). Here, the focus was to gain insight into teachers' perspectives on knowledge of teaching and learning arithmetic in this phase of formal schooling. The questions referred to both the observed and videotaped lessons as well as case examples of learning situations and covered topics on teaching of arithmetic in year one, the teachers' choices of tasks and activities, how they assess students, and the prerequisites for learning arithmetic in year one (see Papers IV and V for details on interview questions). In the second interview, in January 2016, the primary focus was on the observed and videotaped lessons. Excerpts from videorecorded lessons were used for stimulated recall interviews. Video-stimulated recall has been shown to be "an effective method for revealing teachers' tacit knowledge about their pedagogy" (Powell, 2005, p. 407). In connection to the first interview in October 2015, the teachers answered a questionnaire regarding their use of textbooks and other teaching materials, manipulatives, and the amount of time spent on teacher-oriented teaching, group work, and individual work.

Each interview lasted approximately an hour, and all of the interviews were audio recorded and transcribed by student assistants, based on written instructions. I compared the transcripts with the audio recordings in NVivo to ensure the quality and to correct errors in transcription. All excerpts presented in the papers were translated into English.

4.2.4 Video observations of classroom teaching

In the period from November 2015 to April 2016, three different mathematics lessons were videorecorded in each of the six year one classes in Study B (in Classroom E, only one lesson was recorded because the teacher was on leave). The teachers had been asked in advance to plan the lessons within the topics of number and arithmetic, but otherwise perform the lessons as they would normally. I operated a single camera in the back of the classroom, and the recordings focussed on the teacher and the activities. Excerpts from the videos were chosen for the stimulated recall interview based on the criteria that it contained a teacher-initiated activity and the teacher was actively instructing and/or engaged in dialogue with the entire class. The video observations were used for video-stimulated recall in the teacher interviews and were otherwise not further analysed.

4.3 Data analysis

4.3.1 Quantitative analysis for Paper I

The focus of Paper I was to investigate how the use of different strategies in single-digit addition develop with age (within-subject variation), and how strategy use varied between boys and girls and between different classes (between-subject). The data consisted of longitudinal and cross-sectional data from Studies A and B on students' strategy use in years one to four. Data from Study A were used to establish a statistical model that parameterized the effects of the aforementioned factors. Data from Study B were used to test whether this established relationship could be reproduced with an independent data set.

Each of the four strategies (error, counting, direct retrieval, and derived facts) were analysed separately. The statistical observation unit was the answer to each of the 36 flash card questions, and the response variable would be binary with either a positive (the strategy category in question was used) or negative (the strategy category in question was *not* used) outcome. The appropriate statistical solution for modelling that type of response data is binary logistic regression.

The data structure contained multiple levels of dependence, as the students were embedded in classes with a shared learning environment, most students were tested more than once, and each test consisted of 36 tasks. When working with complex systems where context plays a major role and processes or characteristics on a higher level influence processes or characteristics on lower levels, multilevel analysis is an important analytical tool (Luke, 2004). Luke (2004) defines the multilevel model as "a statistical model applied to data collected to more than one level in order to elucidate relationships at more than one level" (pp. 7–8). The levels of data in this study were student, class, and school. Furthermore, many single-level statistical models assume observations to be independent of each other, but this is often not the case with students' cognitive development. Student learning is influenced by the context: the classroom, the teacher, the school, etc., and therefore, students' achievements in a class are not independent of each other. Multilevel models have also been shown to be useful for modelling longitudinal data and for handling missing observations (Luke, 2004).

The quantitative data on student development and use of mental strategies in single-digit addition were analysed using multilevel modelling, also known as generalized linear mixed models (GLMM) with the GLIMMIX procedure in SAS 9.4 software (Littell, Milliken, Stroup, Wolfinger, & Schabenberber, 2006)). A GLMM is an extended form of a generalized linear model (in this case a binary logistic regression model) that can account for both 'fixed' effects and 'random' effects. The fixed effects are the universal effects of main interest for the study (e.g. gender and age). Random effects are background variation attributable to replaceable factors such as student identity (Bob, Andrew, Sarah, etc.) or problem type identity ('5+5', '5+6', '5+7', etc.) that are known or suspected to contribute with variation in the response variable that are not of global interest (unlike fixed effects such as sex or school age) but have to be controlled for in order to avoid pseudo-replication. (Because Andrew is Andrew, he must be expected to some extent to solve different math problems in his own personal way. Similarly, because the math problem 5 + 5 is uniquely different from the math problem 5 + 6, on average, the two problems may, all other factors being equal, be solved slightly differently by students in general.) The variety of strategy use specific to addition task items for boys and girls separately is illustrated in appendix B.

In practice, GLMMs have two types of model statement, one for the fixed effects, which are the universal effects of primary interest for the study (e.g. differences between girls and boys with respect to the frequency with which a given strategy is used on the logit scale) and one for the random nuisance effects that first and last just have to be controlled for when estimating the fixed effects as well as predicted means for combinations of fixed effects (e.g. the frequency with which girls in the middle of first grade are using a given strategy when back-transformed from the logit scale on which the model operates). The random effects are calculated as a covariance parameter of normal distributed dispersion between the random factors (e.g. students) when all other effects have been accounted for. For a further description of the statistical modelling procedure, see Paper I.

4.3.2 Quantitative analysis for Paper II

The aim of Paper II was twofold. First, I wanted to quantify how much of an individual student's strategy use develops during the first half year of formal mathematics education relative to the student's starting point at the beginning of year one. Second, I wanted to estimate how much of the total variance in proportional strategy use in the two test rounds could be attributed to the time between the two test rounds (five months), student identity, class, and the interaction between class and time. The interaction term between class and time was of interest because any difference in student progression due to different teaching practices would be included in the term. The data for this paper were strategies used in the autumn and spring for year one students from Study B.

Analysis of change in proportional strategy use

Each of the four strategies (error, counting, direct retrieval, and derived facts) were analysed separately. In this analysis, the observation unit consisted of the summed outcome of a test round (36 flash cards). For each of the four strategies, the response variable was defined as the proportion of the 36 items that were solved with that particular strategy. This resulted in a state space ranging from 0 (the strategy in question was used to solve 0 out of 36 items) to 1 (all 36 items were solved with the strategy in question). To avoid ceiling effects of the response variables when approaching the limit of the state space, the original proportion values were logit transformed prior to the analysis (logit $p_i = \ln[p_i/\{1-p_i\}]$). Before transformation, zero values (not defined on the logit scale) were substituted with 0.01, which is 2.8 times lower than the minimum achievable value larger than 0 (1 of 36 items = 0.028).

The development in mean strategy use from November to April in year one was analysed by modelling the proportional use in the second assessment interview as a linear regression function of proportional use in the first assessment interview. The mean change in proportional strategy use from interview one can be derived as the vertical distance from a reference line y = x indicating no change in mean strategy use from assessment one to assessment two. As outlined in Figure 4.3, the location of the regression line in relation to the reference line indicates the level of increase or decrease in proportional strategy use. It is possible to test if the change in the use of a strategy is conditional on how much it was used at the first assessment by testing if the slope of the regression line is significantly different from 1. If the slope of the regression line deviates significantly from 1, there is evidence that students' progression from time 1 to time 2 were conditional on their level at time 1 (Figure 4.3).

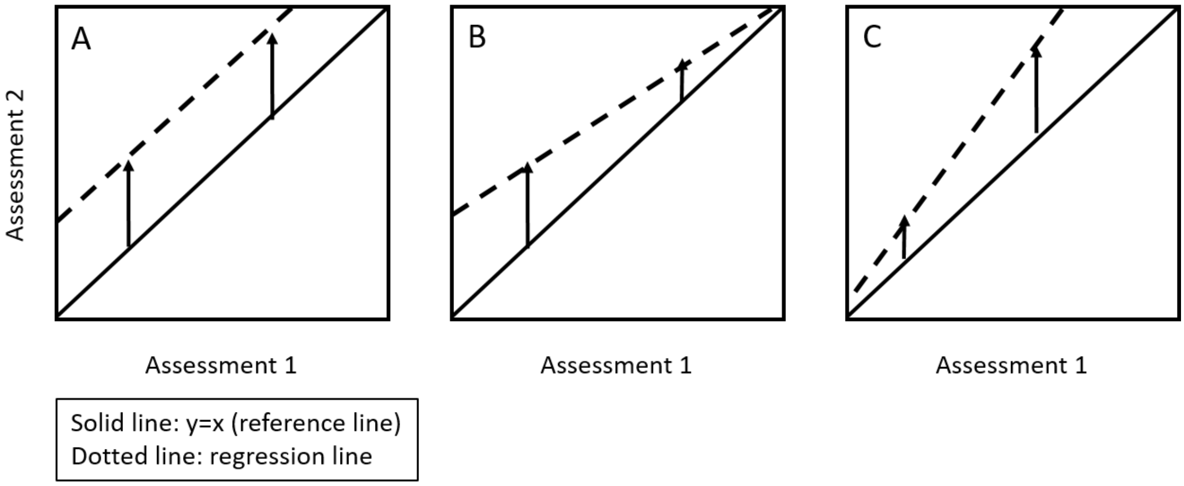


Figure 4.3 Conceptual illustration of the advancement patterns of low and high achieving students between subsequent assessments, analysed as achievement at assessment two, modelled as a linear regression function of the score achieved at assessment one. The deviation from the reference line y = x (indicating status quo) depicts how much students with a given baseline score at assessment one had developed, on average, between assessment one and assessment two.
A. The regression line is above the reference line and the slope is not different from 1, meaning that high and low achievers at assessment one advanced equally, on average, from assessment one to assessment two. This is illustrated as the same vertical difference between the regression line and the reference line y = x, irrespective of x.
B. The regression line is above the reference line of y = x, and the slope is < 1, meaning that the students with the lowest scores at assessment one expressed higher progress and thus appeared to catch up with their higher scoring fellow students between assessment one and assessment two.
C. The regression line is above the reference line of y = x, and the slope is > 1: Students who scored highest on assessment one increased their score even more at assessment two, than did students who had achieved lower scores at assessment one. The gap between the lowest and the highest achieving students thus increased from assessment one to assessment two.

Analysis of variance components

To quantify how much of the variation in proportional use of each strategy was explained by different variance components, the sum of square for the different factors were compared to the total type-3 sum of squares in a general linear model, with all 155 interview results as observation units and the factors and interaction terms as fixed effects. Statistical significances were derived from F-tests. The factors entered in the model were time (difference between interview one and interview two), student identity nested within class (representing the systematic differences between students over the two interviews within a given class), class (the systematic variation between students in different classes), and the interaction between class and time (systematic differences in how much students from different classes changed between interviews one and two). The interaction term class*time was included because it encompassed all variation that could be explained by differential teaching practices since it represents the amount of time a student has received instruction in a specific classroom. This approach has also been applied to reflect differences in the schooling of students in different countries as a nation*time interaction (Geary et al., 1996). As Geary et al. (1996) argued, this interaction term reflects not only differences in the in-class (or in-school) teaching practices, but also other factors associated with the specific entity or class, for example, peer influence.

From a statistical point-of-view, the aforementioned model can be criticized for being too simplistic as it states student identity and class identity as fixed (universal) and not random (representing a sample of units) effects (Littell et al., 2006) that they truly were since the students and classes indeed consisted of a replaceable sample of a larger population of students and classes. The rationale for nevertheless analysing the data with the simpler generalized linear model (GLM), consisting of fixed effects only, rather than using a GLMM that (more statistically correctly) stated student and class as random effects, was technical. Only the former model type returns sum of square statistics, whereas GLMMs are based on likelihood estimates from which variance components cannot be derived. In order to illustrate the relative distribution of variance components of the four strategy use types, the GLM solution was thus the only possible choice.

4.3.3 Quantitative analysis for Paper III

The focus of Paper III was to analyse the association between students' strategy use in year one and their later mathematics achievements, assessed three years later in year four, and to analyse if information on strategy use in year one provided additional predictive information on mathematical achievement in year four that is not available from a standardized mathematical achievement test. The data for these analyses were from the six classes in Study B.

The data used from year one were strategy use scores (test results used as observation unit with logit transformed proportional strategy use as response variable, as described for Paper II [4.3.2]) and mathematical achievement test score (MAT1) obtained in the first assessment round (November 2015). Because the five strategy use measures (logit transformed proportion of frequencies of error, counting all, counting on, direct retrieval, and derived facts) were negatively auto-correlated (the proportion of frequencies summed up to 1) and therefore statistically interdependent, the information from the five proportions was condensed to four principal component analysis (PCA) axes that, by definition, were statistically independent.
The data from year four were scores obtained in autumn 2018 from the mathematical achievement test (as described in section 4.2.2) covering the subtests number and arithmetic (NUMAR), proficiency with fractions (FRAC) and word problems (WP).

To get an overview of how well the different metrics on strategy use and general mathematical achievement from year one predicted performance in different fields of mathematical achievement in year four, a correlation matrix was constructed. The raw predictive potential of each match of variables between year one and year four were assessed as Pearson's r^2 (coefficient of determination). Correlation matrices were constructed for the entire sample, and also for boys and girls separately to make it possible to spot (although not to test outright) if boys and girls appeared to express markedly different correlations between a year one predictor and a year four response variable.

In a second analytical step, for NUMAR, FRAC, and WP, testing was done to explicitly examine whether strategy use measures obtained in early year one contained *additional* predictive information about a student's mathematical achievement in year four after information from the MAT1 test had been included as a predictor. Three types of predictive models were created: (A) a full model including the MAT1 test score and all four principal component analysis scores as predictors, (B) a model with the MAT1 test score as a mandatory variable, to which a step-wise selection procedure added those predictors that most improved the model's overall fit, and (C) a model similar to model (B), except that the MAT1 test score was not forced into the model. As post-hoc operations, testing was done for possible interactions between sex and the covariate included in model (C). On the basis of these model combinations, it was possible to identify which combinations of year one predictors explained most of the variation in the three types of year four mathematical achievement and how large the combined predictive power of the year one variables on the year four response variables was. Finally, to also address the predictive potential of the predictor variables within classes (for some response variables, there was a statistically significant between-class variation), a version of the predictive models was produced that included class identity as a random effect.

4.3.4 Qualitative analysis of interviews for Papers IV and V

Papers IV and V are both concerned about the teacher's perspectives on the learning and teaching of number and addition in year one. Paper IV investigated what teachers expect their students to know about number and addition by the end of year one, and Paper V focussed on the extent to which teachers' perspectives on the teaching and learning of number and addition in year one are aligned with established knowledge on the foundations of numerical and arithmetic competence. The papers draw on the data from the teacher interviews, and the same analytical approach was used in both papers.

The teacher interviews were analysed by means of a constant comparison analytical approach drawn from grounded theory (Strauss & Corbin, 1998). This facilitates the in-depth and elaborate description expected with case study analyses of complex educational settings (Yin, 2013). In brief, constant comparison, in this context, entailed a repeated reading of the data from the first case to identify categories of belief and practice. As each category was identified, the case data were reread to see if it had been missed earlier. On completion of this first pass, a second case was read for evidence of both the earlier categories and new ones. As each new category was identified, all previous case material was scrutinized again. Categorical definitions constantly refined as incidents were compared and contrasted. This process of continual comparison and refinement facilitates the integration of categories and identification of underlying structure, organising principle, or core category (Flick, 2014) and led to the identification of the different themes presented in Papers IV and V.

4.4. Ethical considerations

Throughout, I have followed the prevailing rules and regulations for the correct storing and handling of data according to the Danish Data Protection Agency (www.Datatilsynet.dk), and I have secured written informed consent from all parents of participating students and oral consent from the teachers, as described above (Section 4.1.2). In accordance with the Danish Code of Conduct for Research Integrity (Ministry of Higher Education and Science, 2014), I have striven to be transparent in describing the analytical approach so it is possible to redo and check the findings. Furthermore, the anonymized data on students' strategy use will be available on request.

Research ethics is more than just following rules and regulations, and throughout my research, I have encountered and dealt with different situations and issues that required reflecting on my role as a researcher and how best to act ethically. In the following, I will give examples of these situations with respect to reporting classroom studies, researching young students, and conducting classroom observations, along with how I have dealt with them.

When reporting small-scale studies, there is always a risk that anonymity will be violated (Goodchild, 2013). Although I have used pseudonyms for the teachers, the local community and colleagues, in addition to the participating teachers themselves, will know which schools I have visited and thus may be able to recognize themselves or others. I tried to address this by providing teachers with the opportunity to read the interview transcripts as well as manuscripts. None of the teachers took advantage of this offer.

When working with young students, it can be difficult to know if the student feels uncomfortable or intimidated. During the individual assessment interviews, the students were told that they could just ask me to stop if they did not want to participate. None of the students asked to stop, but on two occasions, I chose to stop the interview because I felt the student was uncomfortable. As explained elsewhere, I made it clear to students that I was not interested in evaluating how good they were with addition, but rather, I wanted to learn from them how young children do addition. Furthermore, I always ended the interviews by thanking the student for the interesting explanations and for having the opportunity to learn from them.

When conducting classroom observations, one must be aware that unexpected situations can occur. For that reason, teachers were informed that I would only use videorecordings for research purposes, and they could ask me to stop the recordings at any moment during a lesson or ask me to delete the recordings without having to justify why. I also made it clear that I would not hand over recordings to either school authorities or parents. Because of my own background and experience as a teacher, I felt it was important that the teachers were explicitly informed of these facts and were given that opportunity because situations occur in the classroom where students or teachers can feel intimidated by knowing a recording exists.

Furthermore, there is a risk that the behaviour of participants will change as a result of the presence of the researcher and camera. This can influence the research as well as the learning opportunities for the students. I observed some students found the camera very interesting the first time, but when they learned that I was interested not in them but in their teacher, they soon lost interest. During the teacher interviews, I asked the teachers if my presence and the camera had influenced their own or their students' behaviour. They all reported the same observations as my own about the students' behaviour. Furthermore, the teachers claimed they had taught the class in the same manner as they typically did. One teacher said with a smile, “Well, I made sure to put on a clean T-shirt this morning, but otherwise, everything was as any other day”. Another teacher said to me after one of the recordings, “I completely forgot you were there in the back of the class”.

On one occasion, a teacher told me she had felt uncomfortable by my presence and the camera. We immediately talked through the situation, and she explained that she had felt pressured because she failed to explain the mathematics for a student who was struggling with one of the activities. I told her what I had observed and offered to let her see the videorecording. She declined, saying that it was enough for her to share the experience with me. For me, this was an example of the importance of gaining the teachers’ trust and being honest and open about observations and research purposes. It is important that participants feel they have authority over the data and can trust that I will treat them and the data with respect.

An ethical aspect I had not foreseen was that the teachers expected more from me than I felt I could provide. On several occasions, different teachers asked me what I thought of their teaching, whether “it was OK” or if they should teach differently. One teacher said at the end of the project, “I have given you so much information, so now I want to know what you as an expert think about my teaching”. In retrospect, I think I had not made clear that I would not and cannot judge their teaching and that my research was merely exploratory and descriptive. However, I think this also tells me that the teachers had confidence in me. I arranged a meeting with the teachers after the final student assessments where I shared information about their students’ test results and during which we discussed possible interventions for students at risk for difficulties in learning mathematics.

5. Presentation and discussion of papers

The thesis is based on five individual papers, which are attached. This chapter summarises their aims, main results and discussion points and, where relevant, discusses these in their broader context. Full details of methods can be found in Chapter 4.

5.1 Paper I: Sex Differences in Mental Strategies for Single-Digit Addition in the First Years of School

5.1.1 Aim and research question

The aim of this paper was to describe and quantify the development and variation in Danish students' use of strategies for single-digit addition during the first four years of school. As well as the effect of time (school age), the effect of sex and class was investigated. The study addressed the following research question.

How does strategy use in single-digit addition develop with age, and is there any systematic between-subject or between-class variation in patterns of strategy use among boys and girls? (RQ 1, Chapter 3)

5.1.2 Methods

In this longitudinal study involving 147 students from a single school (2012–2015) (Study A) and 83 students from six classes and three schools (2015–2016) (Study B), participants' strategy use was assessed and categorised (as described in Chapter 4). Data from both studies were analysed using Generalised Linear Mixed Models (GLMM), and data from Study B were used to validate the results from Study A.

5.1.3 Results

Overall, frequency of use of all strategies varied as a function of age because, with increasing age, error and counting strategies became less frequent, and derived fact and direct retrieval strategies were used more often. Apart from error, there were substantial sex differences in the use of different strategies with age. Girls used counting (that is, counting all as well as counting on) significantly more often than boys, with an odds ratio of approximately 3:1 (F:M), while girls used direct retrieval and derived fact strategies significantly less than boys, with odds ratios of approximately 1:3 and 1:2, respectively. This difference corresponded to 2.5 years of development. These sex differences were strongest at the beginning of year one, levelling out for counting by year four and for direct retrieval and derived fact by year two. These patterns were confirmed by the results from study B. Both studies (A and B) showed no significant influence of class on strategy use. When counting was split into counting all and counting on, a post hoc analysis of sex differences showed F:M odds ratios of approximately 3:1 and 2:1, respectively.

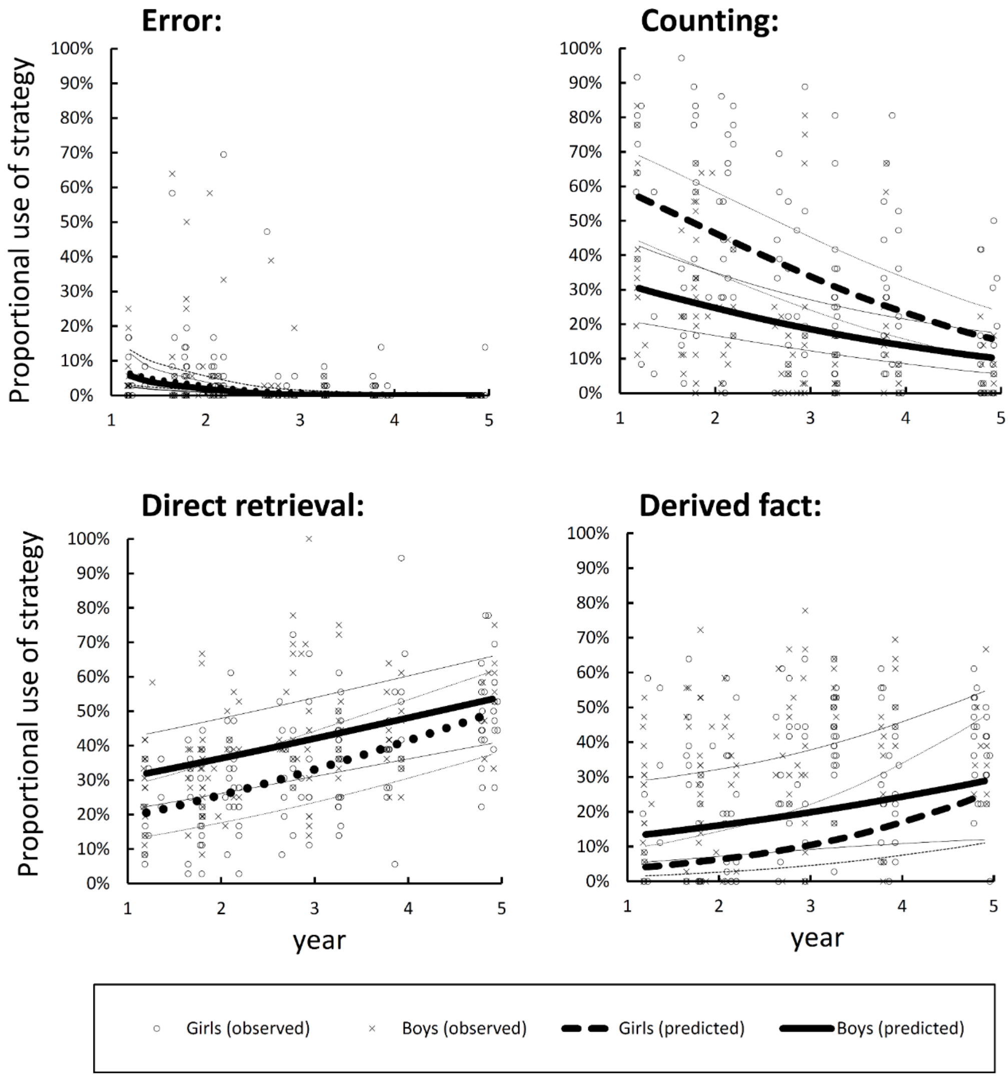


Figure 5.1. Proportional use of different addition strategies plotted against school year for girls and boys. Thick lines indicate the predicted functions for each sex; thin lines demarcate 95% confidence zones (Figure sourced from Paper I).

5.1.4 Discussion

Strategy use changed with age; specifically, the use of derived fact strategies increased while the use of counting strategies decreased. This pattern aligns with the existing evidence (Carpenter & Moser, 1984; Sarama & Clements, 2009). The girls participating in the present study showed patterns of strategy use that were 2–3 years behind the boys; as this developmental pattern was highly significant across schools and classes, it seems ubiquitous in the Danish primary school context. These findings align with several studies reporting that girls use counting more often than boys while boys use direct retrieval and derived fact strategies more often than girls (Bailey et al., 2012; Carr & Alexeev, 2011; Imbo & Vandierendonck, 2007; Shen et al., 2016). The finding that girls are 2–3 years behind boys aligns with Bailey at al. (2012), whose US study found that boys' use of direct retrieval was approximately two school years ahead of girls in years 1–6.

Individual as well as sex differences in strategy use have been attributed to cognitive (e.g. Foley et al., 2017; Laski et al., 2013) and cultural factors, including instructional practices (e.g. Hyde & Mertz, 2009; Penner & Paret, 2008; Shen et al., 2016). While sex differences seem to level out around year four, this may be a ceiling effect, as single-digit addition may not be a good diagnostic tool for the average student by year three, as also reported by Shen at al. (2016).

The magnitude of reported sex differences highlights the importance of investigating the occurrence and possible causes of sex differences in arithmetic patterns. For instance, it was difficult to relate our findings to results from other studies because sex differences are not tested for or described in adequate detail in many cases. We conclude that more knowledge is needed about the factors that influence sex differences, including individual differences and sociocultural factors such as teaching and educational practices.

5.2 Paper II: Development and variance components in single-digit addition strategies in year one

5.2.1 Aim and research question

This paper quantified the development of strategy use among Danish year one students over a five-month period (from November to April). Conditional development was analysed by modelling students' strategy use in April as a regression function of their initial strategy use in November. We also investigated the contribution of different components of variance, including student ID, time (i.e. the interval between assessments) and different teaching practices (i.e. the interaction of class with time) to the overall variation in strategy use. The paper addressed the following research questions.

To what extent does the individual student's strategy use develop during a half-year of formal mathematics teaching from their baseline level at the beginning of year one?
(RQ 2, Chapter 3)

How much do the different components of variance (initial strategy use, time, teaching practice) contribute to the overall variation in year one students' development of strategies?
(RQ 3, Chapter 3)

5.2.2 Methods

In study B, data on strategy use among the 83 participating year one students included 155 assessment interviews. As documented in the teacher interviews and video observations, teaching practices differed substantially across the six classrooms.

Development of strategy use was analysed as strategy use in the second assessment interview (April), modelled as a linear regression function of strategy use in the first assessment interview (November). Mean change in proportional strategy use was analysed by comparing the regression line for each strategy with a reference line $y = x$, representing the theoretical status quo, where students had achieved similar results in assessments 1 and 2. The vertical difference between the regression line of the function and the reference line $y = x$ indicated mean development in strategy use from November to April for students' given use in November (x-value). The R^2-value indicated how much of the variation in students' strategy use in April could be explained on the basis of their initial strategy use in November.

The total variation in proportional strategy use explained by the different explanatory factors was quantified as the sum-of-squares of the individual factors as compared to the total sum-of-squares in the general linear models, with the 155 test results (November and April) as observation units and the six classes, the students (n = 83) nested within the six classes, time of testing and time-by-class interactions as fixed effects. Statistical significance was established using F-tests.

5.2.3 Results

Students' proportional strategy use at assessment interview 2 was highly conditional on strategy use at assessment interview 1 (Figure 5.2), with R^2-values ranging from 0.26 (error) to 0.69 (derived fact). For error and direct retrieval, regression line slopes were significantly less than 1. In relation to error, this indicates that the students with most errors in assessment 1 showed the highest reduction in errors from assessment 1 to assessment 2. In the case of direct retrieval, students who used this strategy least often in assessment interview 1 showed the highest increase in use in assessment interview 2. Those students with the highest proportional use in interview 1 showed no mean increase in interview 2. For counting and derived fact strategies, the slopes of the regression lines did not differ significantly from 1, meaning that change in use of these two strategies was constant, irrespective of initial proportional use.

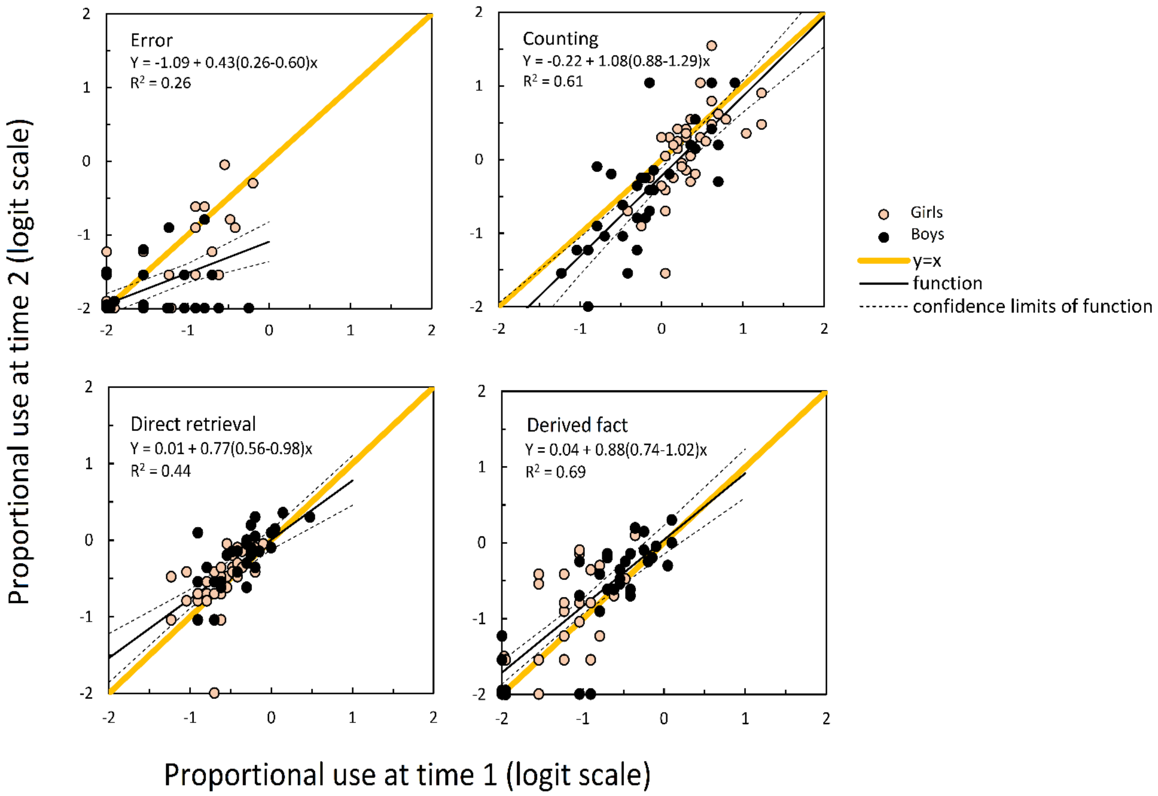


Figure 5.2 Proportional use (logit-scale: zero-values replaced by 0.01 before transformation) of strategies among 72 year one students in spring (April) plotted against their results in the same assessment interview in the preceding autumn (November). In the equations for the linear regression lines, 95% confidence limits of the slopes are shown in brackets (Figure sourced from Paper II).

The effects of the different components of variance showed that 60–83% of the total variation was explained by student identity while only 3–6% was explained by time (the five-month interval) and 3–12% by class. The interaction term *class by time* contributed less than 1% of the explained variation and was not statistically significant for any strategy.

5.2.4 Discussion

Counting is generally viewed as a less sophisticated strategy, and its excessive use has been linked to mathematical difficulties (Ostad, 1997a). In contrast, the use of derived fact strategies is associated with high arithmetic ability (Dowker, 2014). From a teaching perspective, one should therefore aim for an increase in the use of derived fact strategies over the course of year one at the expense of counting. On that basis, three main conclusions can be drawn from the analyses.

First, there was considerable spread in the frequency of students' use of different strategies, and this persisted over the five-month interval between interviews 1 and 2; this was particularly the case for counting and derived facts. While mean use of all strategies changed significantly over those five months, these changes were modest in relation to the considerable variation between students, confirming that variations in strategy use patterns within a class correspond to several years of education. This result emphasises the importance of acknowledging the substantial individual variation in strategy use in the first half of year one, which remain consistent throughout that year. Accordingly, unless overall strategy use patterns are established within the first three months of year one, most of the between-student variation in strategy use must be accounted for by factors unrelated to the year one teaching environment.

Secondly, students who made least use of direct retrieval strategies at assessment 1 exhibited the highest mean increase in use of this strategy between November and April. In contrast, all students exhibited a similar mean change in use of counting and derived facts strategies, regardless of their initial level. These findings emphasise the importance of taking account of students' differing starting levels (e.g. of strategy use) when analysing such changes.

From a learning perspective, where the aim is to reduce errors and excessive use of counting, one might hope that the students who made most errors and counted most, with most room for improvement, would also exhibit the greatest progress from interview 1 to interview 2, 'catching up' with those who had already reduced their dependence on counting. This was clearly the case in relation to error, as the majority of students with the highest error rates in interview 1 acquired the necessary skills to solve problems correctly when tested again five months later. However, this was not the case for counting, as the vertical deviation between the regression line and the reference line (as defined by the slope of the regression line) was either constant or decreased with increasing x-values (i.e. high use of counting in interview 1). In short, there was no indication of a 'catching up' effect of counting as was the case for error, which might suggest that reducing excessive use of counting is harder than reducing a high error rate.

From a methodological perspective, other studies might also find it useful to parameterize conditional progression in achievement (in whatever topic) as the vertical deviation between the estimated regression function (not necessarily a linear function, but easily modelled as a quadratic or cubic polynomial function) and the nil-expectation line of $y = x$. In that respect, the test of the nil-expectation of a slope = 1 is an easily applied and

comprehensible method of testing the nil-expectation that all students have made equal progress during a survey or intervention interval, whether their initial level of achievement was low or high.
Third, and finally, there was effectively no difference between the six classes in terms of how much the students' strategies changed between November and April. As any effect of different teaching practices is covered by the class-by-time interaction, it must be concluded that this had no measurable effect on changes in strategy use in the five months between interviews 1 and 2. This does not mean that teaching practice does not influence individual students' development; it shows only that differences in teaching practice had no measurable effect in the period from November to April in year one. Teaching practices in preschool or early in year one may have had some effect before the first assessment interview in November. There may also have been a delayed effect of teaching practice that did not appear until after the second assessment in April.

5.3 Paper III: Single-digit addition strategies in year one predicts achievement in year four

5.3.1 Aim and research question

This paper investigated the extent to which specific strategy use in year one predicts mathematical achievement in arithmetic and proficiency with fractions and word problems in year four. The study addressed the following research question.

How does strategy use among year one students relate to subsequent mathematics achievement (assessed in year four), and does information about strategy use in year one provide additional predictive information beyond a standardised test of mathematical achievement? (RQ 4, Chapter 3)

5.3.2 Methods

The data related to students in study B; because a number of students from the original study had changed school, we analysed only data for a subsample of students who also attended that class and school in year four. Early strategy use in year one was assessed and categorised as described in Chapter 4, along with students' achievement in a standardized test in year one (MAT1) and scores in number and arithmetic, fractions and word problems in year four.

To circumvent problems of interpretation caused by autocorrelation of the five measures of strategy use, the information was reduced to four principal component axes (PRIN1, PRIN2, PRIN3 and PRIN4). To identify which single year one variable predicted most variation in year four mathematical achievement, we first correlated all nine strategy use measurements and the mathematical achievement test result from year one with the achievement test results from year four. Model selection procedures were then used to test whether strategy use measures from year one contributed additional predictive information about year four achievement when year one mathematics achievement was included as predictor.

5.3.3 Results

At November in year one, the most frequently used strategy was counting on, followed by counting all, direct retrieval, derived facts and error. There was considerable between-student variation in frequency of use of the different strategies (Figure 5.3).

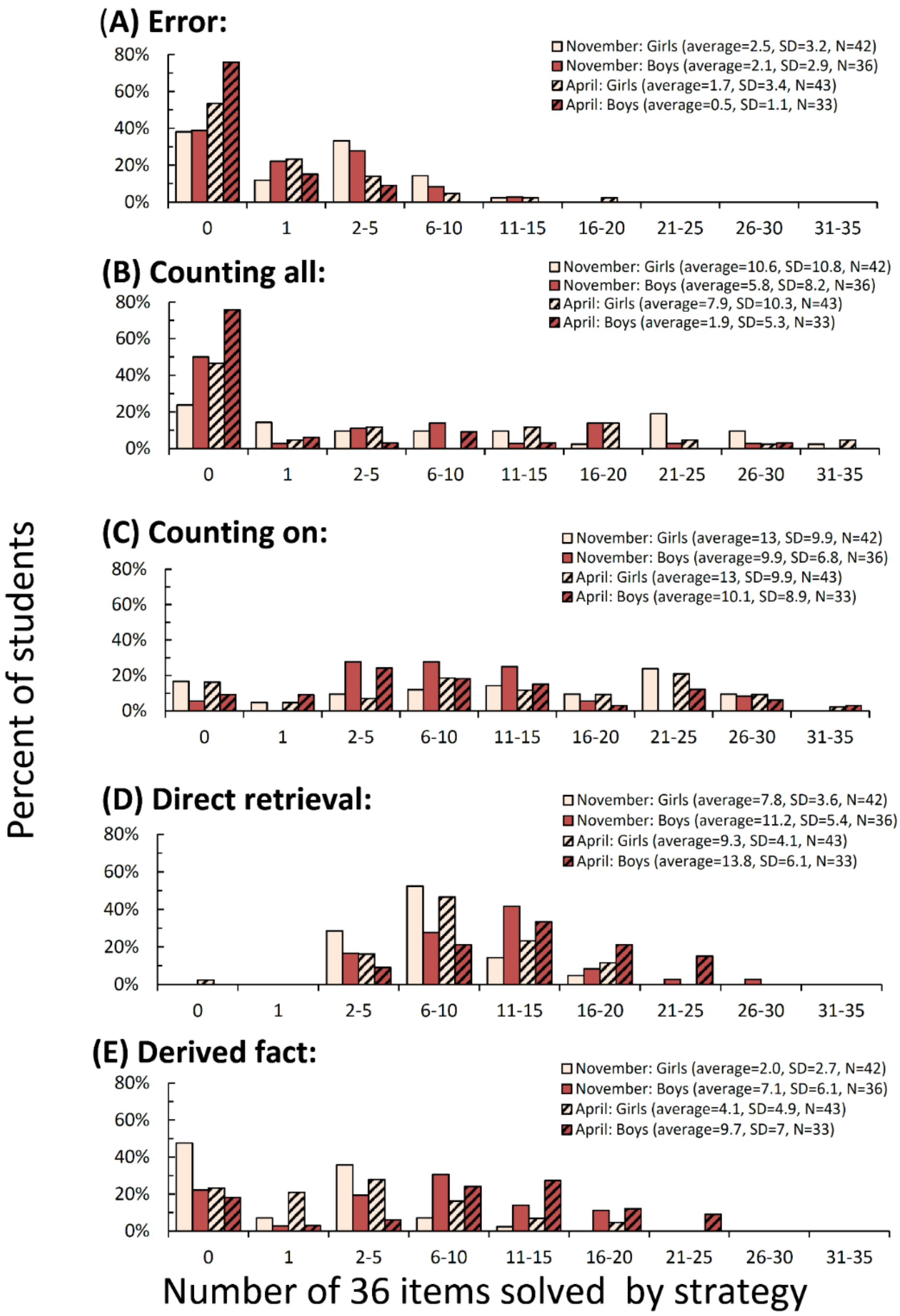


Figure 5.3 Frequency distribution of strategies used to solve the 36 single-digit addition items in the first (November) and second (April) assessments by sex (Figure sourced from Paper III).

As most of the five measures of year one strategy use correlated internally, they could not be regarded as independent if entered as predictors in the same model. In particular, *count all* was significantly negatively correlated with all of the other strategy use variables. *Direct retrieval* and *derived facts* correlated positively. Of the four PCAs, PRIN 1, 2 and 4 appear to represent the different gradients of variation in advanced strategy use. In this regard, PRIN1 was correlated with all strategies, most negatively with *count all* (r = -0.91) and most positively with derived fact (r = 0.78). Similarly, PRIN2 correlated most negatively with *count on* (r = -0.86) and most positively with direct retrieval (r = 0.56) and derived fact (r = 0.35). PRIN4 captured an apparent difference between more use of *direct retrieval* (r = 0.51) and less use of *derived fact* (r = -0.43). PRIN3 correlated strongly positively with error (r = 0.88), indicating that it mainly represents a gradient of low general mathematic achievement, expressed as an increasing error rate.

All types of mathematics problem in year four correlated significantly with one or more strategy use variables in year one. Specifically, use of *error* and *count all* in year one correlated negatively with achievement scores in year four while frequent use of *count on*, *direct retrieval* and *derived fact* correlated positively with year four achievement. High correlations of genuine strategy use variables and mathematical achievement scores were found between count all and word problem solving and between count all and fractions. When the latter analysis was split by sex, the correlations were stronger and highly statistically significant for boys but weaker and not statistically significant for girls.

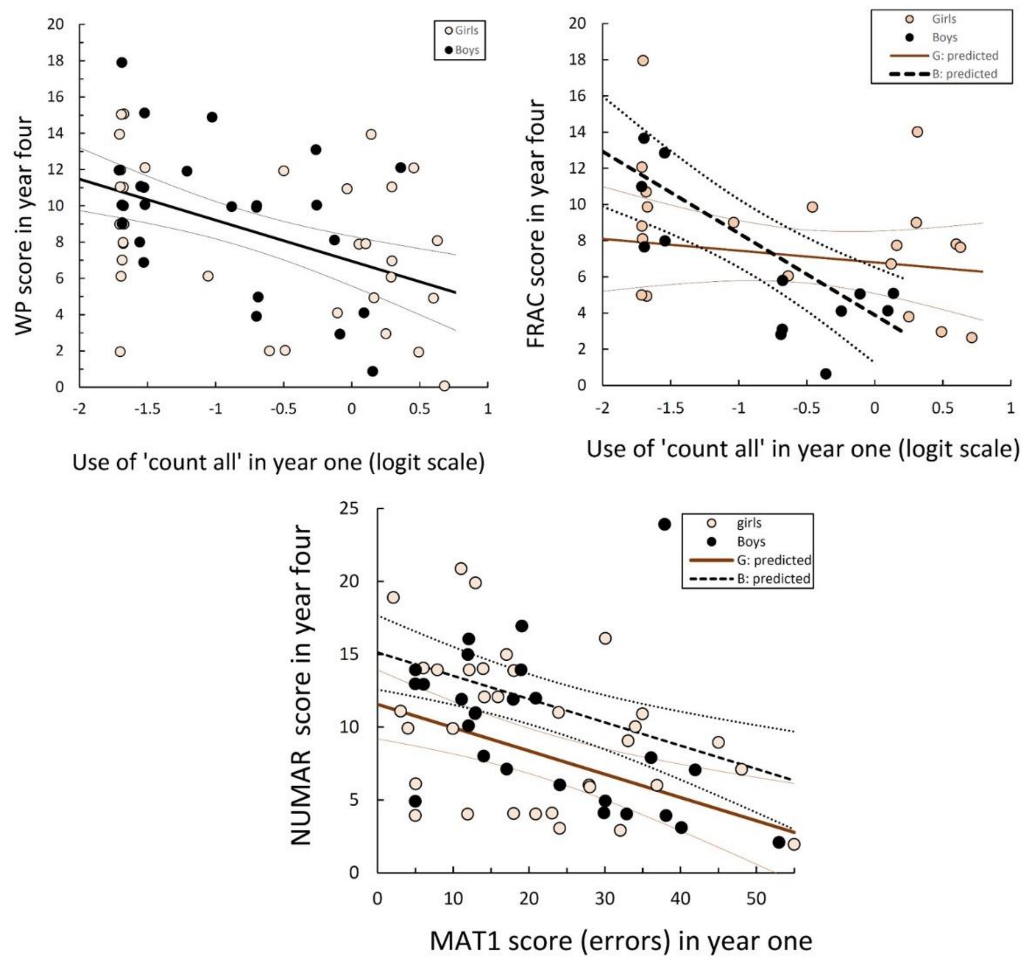


Figure 5.4. Achievement scores for word problems (WP), number and arithmetic (NUMAR) and fraction knowledge (FRAC) in year four plotted and regressed against the test predictors from year one that explained most variation in these response variables in conjunction with other predictors. Achievement scores for WP were best explained as the combined negative function of count all (ΔR^2=24%) and the MAT1 test result for general achievement (ΔR^2 = 17%, combined model R^2 = 31%), the achievement score for NUMAR as a combined function of MAT1 (ΔR^2=16%) and sex (ΔR^2 = 12%, combined model R^2 = 28%), and FRAC-score as an interactive function of count all (ΔR^2=23%) and sex (ΔR^2 = 17%, combined model R^2 = 39%) (Figures sourced from Paper III).

Class differences were identified for year four achievement in number and arithmetic and word problem solving but not for fractions. There were sex differences in year four number and arithmetic achievement, with boys scoring higher on average than girls (11.5 vs 8.5 points).

The model selection procedures revealed that after accounting for mathematical achievement test scores in year one, use of count all in year one predicted additional variation in achievement scores for fractions and word problem solving—that is, higher proportional use of counting all in year one was associated with lower achievement scores in year four. In the case of fractions, this negative relation was stronger for boys than for girls (Figure 5.4).

5.3.4 Discussion

The study revealed that about one third of the variation (28–39%, depending on problem type) of the variation in year four achievement response variables could be explained by fixed-effect predictors from the first half of year one. In that respect, measures of strategy use alone (especially those associated with use of counting all) explained between 13% of the variance in number and arithmetic and 24% of the variance in word problems and fraction knowledge. These results align well with previous evidence that early quantitative and arithmetic competence are predictive of later mathematical achievement (e.g. Geary, 2011; Jordan et al., 2009; M. M. M. Mazzocco & Thompson, 2005).

From both scientific and screening perspectives, it is worth noting that, for two of three problem types in year four achievement (word problems and fraction knowledge), measures of year one strategy use (related to use of count all) explained significant variation in achievement in year four after measures of general achievement in year one had been accounted for in the statistical models. In relation to knowledge of fractions, use of counting all explained all of the significant variation in year four achievement. For word problems and knowledge of fractions, single-digit addition strategy use at the outset of their mathematical education in year one was more predictive of their year four ability in this regard than the results from the year one general achievement test.

From a scientific perspective, the correlations between strategy use in year one and mathematical achievement in year four indicate that much of the variation in problem solving strategies in year one relates in some way to learning trajectories in the following years (although the underlying mechanisms remain to be clarified). From a learning perspective, the obvious follow-up question is whether strategy use in year one forms the basis for acquisition of more complex mathematical skills or merely indicates an underlying ability to comprehend and solve mathematical problems. Future investigations should assess how and to what extent later achievement in mathematics can be improved through teaching in developing and using more advanced single-digit addition strategies—that is, the causal relationship between active training of early number sense and strategy use and later mathematical achievement. In the present data, the significant between-class variation in almost all parameters of year four achievement, and the fact that this accounted for some of the partial effects of year one strategy use on year four achievement, suggests that teaching environment is a significant component of student achievement.

From a screening perspective, the result indicate that assessment of strategy use may provide useful additional information to identify students at risk of becoming low performers. In this study, for instance, frequent use of count all (or low values of PRIN1 that were strongly negatively correlated with count all use and strongly positively correlated with derived fact use) predicted subsequent proportionally low achievement in fraction and word problem solving that was not captured by the MAT1-test. A plausible explanation is that the mathematically simple problems in the MAT1 test for year one students could be efficiently solved using the unsophisticated count all strategy. Students that relied excessively on the count all strategy in their problem solving in year one without achieving any deeper understanding of number and arithmetic (for whatever reason) could therefore perform reasonably well on the general mathematical achievement test. If that was the case, their possible difficulties in understanding number and arithmetic would therefore be less apparent than if the information from their strategy use patterns had been included.

Proficiency with fractions is based on well-developed number sense and an understanding of part-whole relations and partitioning. Derived fact strategies build on a more complex understanding of numbers than counting all. Some facts must be automatized (e.g. Threlfall, 2002), and one needs to understand that numbers can be partitioned in many ways, as well as understanding relations between operations (Ambrose et al., 2003; E. M. Gray & Tall, 1994; Verschaffel et al., 2007). The use of derived fact strategies has been shown to relate to primary students' understanding of part-whole concepts (Canobi, 2004). Counting does not require an understanding of part-whole relations; an understanding of numbers as standalone entities will suffice (Fuson, 1992). As outlined by Van Dooren, Lehtinen and Verschaffel (2015), part of the explanation of the whole number bias is that whole number counting and an understanding of numbers as discrete entities is not applicable to rational numbers.

It follows that the association between use of these two strategies in year one (as expressed in proportional use of counting all and the PCA-score for PRIN1) and fractional knowledge may be related to the student's underlying number sense. Students exhibiting high proportional use of counting all may have a less developed understanding of numbers, part-whole relations and relations between operations, which are important prerequisites for later proficiency with fractions.

On this view, when testing students at such an early stage of their formal education, understanding how they solve single-digit addition problems (not just how well they solve them) may enhance the teacher's ability to assess the student's understanding of number and arithmetic and their risk of developing mathematical difficulties or becoming low performers. A final point of interest here is that boys and girls appeared to differ in terms of how their use of at least one strategy in year one correlated with achievement in year four. The negative correlation between frequency of use of count all in year one and year four achievement in knowledge of fractions was significantly stronger in boys than in girls. Given the modest sample size and the numerous models tested here, some caution is needed in interpreting this result. However, if the observed pattern reflects a true underlying interactive difference, the obvious follow-up question is what causes these differences between boys and girls in terms of strategy use (i.e. use of count all) in year one and achievement in certain types of mathematical problem solving three years later.

One hypothesis might be that while girls are more inclined than boys to use count all to solve a given problem even when capable of using retrieval strategies, boys uses count all only if they fail to master retrieval strategies. In that case, the frequency with which students use count all becomes a much more sensitive indicator of number skills in boys than in girls, where the correlation is blurred by a higher proportion of students exhibiting excessive use of count all relative to their number skills. If this is the case, one would predict that among students who most often use count all, girls should on average outperform boys who make similar use of count all in the fraction knowledge test in year four. This is supported by the data, but this prediction needs to be tested on a larger data set before any firm conclusions can be drawn about the generality of sexual differences in how strategy use in year one relates to achievement in later school years (and the underlying reasons for this).

5.4 Paper IV: Danish teachers' expectations of year one pupils' additive competence.

5.4.1 Aim and research question

The aim of this paper was to explore what six Danish teachers expected their students to learn in year one, and to what extent these expectations aligned with different learning pathways. The study addressed the following research question.

What do teachers expect their students to know about number and addition by the end of year one? (RQ 5, Chapter 3)

5.4.2 Methods

Six year one teachers from study B were interviewed in October 2015. This analysis involved only questions from the interview guide that related to teachers' expectations for their year one students. The interview transcripts were analysed by means of a constant comparison approach, which yielded three key themes: curriculum objectives, mathematical content and differentiation. The teachers' expectations were then analysed in terms of these three themes.

5.4.3 Results

In general, the teachers did not refer to the Danish curriculum. In terms of the mathematical content, their expectations for their students differed by 1–2 years of learning. For example, with regard to digit-range in addition, one teacher stated that students should be proficient in single-digit addition. In contrast, another expected all students to be proficient in two-digit addition without bridging ten, and it was expected that 20% of students would be proficient in bridging ten operations. The teachers also differed in terms of their focus on procedural or conceptual knowledge.

5.4.4 Discussion

The six year one teachers' expectations for their students differed substantially; according to established learning trajectories, those differences equate to one to two years of learning (Sarama & Clements, 2009). The Danish national curriculum does not provide detailed guidance for teachers on such matters as when students should be proficient in two-digit addition bridging ten, and this is one possible explanation for the differences in teachers' expectations. If teachers teach on the basis of their expectations, students are likely to have very different teaching and learning opportunities. Given the evidence for a link between early numerical and arithmetic competence and later mathematical achievement (e.g. Jordan et al., 2009; Landerl, Bevan, & Butterworth, 2004), it is therefore important to establish the extent to which teachers' expectations align with their actual teaching, and to what extent this influences students' development.

5.5 Paper V: Teaching addition: Teachers' perspectives on teaching and learning number and addition in year one

5.5.1 Aim and research question

This paper analysed teachers' perspectives on teaching and learning of number and arithmetic in year one and compared those perspectives with known foundational factors for this kind of learning. The paper addressed the following research question.

To what extent do teachers' perspectives on the teaching and learning of number and addition in year one align with established knowledge concerning the foundations of numerical and arithmetic competence? (RQ 6, Chapter 3)

5.5.2 Methods

The data were collected in the October interviews from study B. The semi-structured interviews were analysed to categorise the teachers' answers, and two themes were identified: *knowing numbers* and *doing addition*.

5.5.3 Results

Within the theme *knowing numbers*, the subcategory *quantity* was mentioned by only four of the teachers, and *estimation* was only mentioned by a single teacher. *Counting skills* was mentioned by all of the teachers, as was *partitioning of numbers*. However, the latter was explicated in a very different way; for example, one teachers said 'when we reach ten, we put two numbers together' while others talked about partitioning in 'tens and ones' and 'grouping in tens'.

Analysis of the teachers' perspectives on *doing addition* identified two groups. Two teachers focused on counting and standard written algorithms and did not mention derived fact strategies. The other four teachers emphasised both counting and derived fact strategies in year one; the written algorithm, when mentioned, was seen as a long-term goal.

5.5.4 Discussion

All of the teachers' responses in relation to number and arithmetic in year one mentioned many of the basic components of number sense, including *number name, number symbol, quantity, counting skills* and *partitioning of numbers*. However, it is noteworthy that only one teacher mentioned *estimation*, and none mentioned *number comparison*, both of which are considered important for the development of arithmetic competence (Booth & Siegler, 2008; De Smedt, Noël, Gilmore, & Ansari, 2013; Gilmore, McCarthy, & Spelke, 2007). These findings align with Sayers and Andrews (2015), whose analysis of 18 episodes from lessons across six different European classrooms found no examples of estimation activities and only two episodes from a single teacher introducing quantity discrimination.

The interviews revealed distinct differences of emphasis in relation to derived fact strategies and the written standard algorithm. Using number-based or mental strategies for multi-digit arithmetic affords opportunities for students to develop their number sense and arithmetic understanding (e.g. Kilpatrick, Swafford, & Findell, 2001). A premature focus on standard written algorithms without first providing opportunities to develop derived fact strategies may result in a less flexible approach to arithmetic. When students are first introduced to standard algorithms, it requires some effort to maintain their strategies for mental computation (Torbeyns & Verschaffel, 2016) and their flexibility in working with numbers and arithmetic.

The teachers showed a general awareness of the foundations of arithmetic competence. However, it became apparent that only one teacher mentioned estimation while none mentioned quantity discrimination skills. Only two of the six teachers regarded the standard algorithmic approach as an integral part of teaching in year one; the rest regarded this as a long-term goal.

Table 5.1: Overview of the two themes: knowing numbers and doing addition. X indicates that the teacher's utterances in the interview were assigned to that sub-category. (A: Allan, B: Bettina, C: Carl, D: Dan, E: Else and F: Frida) (Table sourced from Paper V).

Knowing numbers							
Category	**Sub-category**	**A**	**B**	**C**	**D**	**E**	**F**
Number symbol		X	X	X	X	X	X
Number name		X	X	X	X	X	X
Quantity	Link between number and quantity	X	X	X		X	
Estimation						X	
Counting skills	Number sequence	X	X	X	X	X	X
	Number before and after			X		X	
	Use number line	X	X	X	X	X	X
	Skip counting/times tables	X	X	X	X	X	X
Partitioning numbers	Base ten knowledge		X	X	X	X	X
	Decomposing/composing numbers		X	X	X	X	X
Doing addition							
Category	**Sub-category**	**A**	**B**	**C**	**D**	**E**	**F**
Counting strategies		X	X	X	X	X	X
Direct retrieval	Automatisation of single-digit sums	X		X		X	
	Knowing friends of ten	X		X	X	X	
Derived fact strategies			X	X	X	X	
Written algorithms		X		X	X		X

6. General discussion

In this chapter, I will first discuss the overall results not already addressed in the associated papers. I will then discuss some methodological considerations and limitations of relevance to my conclusion. After reflecting on the implications of these results for teaching practice, I conclude by identifying new research questions arising.

6.1 Discussion of results

The overarching aim was to investigate patterns in the development and use of strategies for single-digit addition among Danish students in relation to 1) differences in teaching practices, 2) later mathematical achievement and 3) teachers' perspectives on teaching and learning of number and arithmetic. The baseline pattern of strategy use in single-digit addition as a function of age and sex and variation across classes and individuals from early year one to late year four is statistically described for the first time in a Danish context.

The key findings reported in the five papers can be summarised as follows.

1. From the beginning of year one, substantial individual variation in strategy use totally overrode individual development over a five-month period (Papers I–III).
2. Despite clear differences among teachers' perspectives on teaching (Papers IV–V), this was not reflected in developmental patterns of strategy use among year one students (Paper II).
3. Strategy use in year one was significantly correlated with mathematical achievement in year four (Paper III).
4. Boys and girls differed in terms of strategy use patterns and how year one patterns correlated with mathematical achievement in year four (Papers I and III).

Overall, these results confirm the relevance of students' early understanding of number and arithmetic (i.e. strategy use), both as indicators of later achievement and as an important focus for intervention and targeted teaching to support student development in all cases.

The general trends in these findings are discussed in relation to individual characteristics, socio-cultural context and later mathematical achievement, as well as overall patterns of sex difference and observed strategy use.

6.1.1 Patterns of development, individual characteristics, socio-cultural context and mathematical achievement

From year one to year four, Danish students' use of counting strategies decreased, and their use of fact-based strategies increased (Paper I). Strategy use patterns varied widely between individuals (Papers I–III) but less so between classes (Papers I–II), and differences in teaching practice were not reflected in any developmental differences across classes (Paper II). The span in strategy use across students was more or less constant throughout year one, with no indication that students who most often used counting reduced their use of this unsophisticated strategy more rapidly over time than those who used this method less often from the start (Paper II). The variation in students' strategy use was not surprising in itself, but the persistence of that variation and the apparent lack of influence of different teaching practices was unexpected (Paper II). Measures of strategy use in year one, especially frequency of use of counting all, correlated significantly with mathematical achievement in year four. For two of the three measures of mathematical achievement (knowledge of fractions and word problem solving), strategy use patterns in year one accounted for the variation not explained by a standard achievement test in year one (Paper III). In other words, how a student performed single-digit addition in year one was a better predictor of later achievement in fraction and word problem solving than how well the student performed addition in year one (Paper III).

Taken together, these results indicate that habits of strategy use are deeply rooted within the individual child and must have been established either before or at the very beginning of formal schooling, changing only slowly and gradually over time. It is not possible to reliably infer the deeper reasons for these apparently hardwired individual 'profiles' of strategy use, but it seems likely that they relate to biological or socio-cultural factors prior to commencement of formal mathematic teaching in year one.

From a research perspective, the persistence of individual variation highlights the importance of analysing developmental patterns across different student groups (e.g. Carr & Alexeev, 2011), including different patterns of strategy development and the relationship between strategy development and specific aspects of later mathematical achievement. Knowledge about the relationship between early strategy use and the individual's approach to new arithmetic problems (e.g. in multiplication) can enhance our understanding of relations between strategy use, number sense and understanding of operations. For example, one might hypothesise that students who rely on counting for single-digit addition in year one are more likely to adopt an additive multiplication strategy (e.g. skip counting) and are less likely to use a strategy based on partitioning (e.g. $6 \cdot 8 = 2 \cdot 3 \cdot 8$). In this regard, Vasilyeva et al. (2015) identified a relationship between first graders' knowledge of single-digit addition number facts and arithmetic strategies and their approach to more complex and less familiar double-digit addition problems. It would be especially interesting to investigate which aspects of intervention and targeted teaching have lasting effects on specific mathematical achievement for students with different developmental profiles.

The persistence of individual variation becomes clear in the association between year one strategy use and later mathematical achievement. Several studies have identified a link between number sense, arithmetic competence and later achievement (e.g. Geary, Hoard, Nugent, & Bailey, 2013; Jordan et al., 2013; Jordan, Kaplan, Ramineni, & Locuniak, 2009; Nguyen et al., 2016), and the present findings aligns with Geary (2011) and Nguyen et al. (2016) in that year one strategy use is predictive of later (year four) mathematical achievement. Nguyen et al. (2016) found that preschool competences in counting—specifically, advanced counting with cardinality and counting forward and backward from a given number—was much more predictive of year five mathematical achievement than basic counting (i.e. verbal 'rote' counting, number recognition and one-to-one correspondence).

Geary (2011) found that year one use of decomposition and more advanced counting (i.e. frequent and accurate use of the min procedure) predicted year five mathematical achievement.

In studies like those mentioned above, later achievement are commonly based on a 'general score' in either an arithmetic test or a conglomerate test of different aspects of mathematics. Less is known about the relation between individual variation in strategy use early in year one and specific mathematical achievement several years later (e.g. fraction knowledge). The present study shows that specific strategy use in single-digit addition in year one—in this case, proportional use of counting all—is more predictive for complex mathematics (such as competences with fractions and solving word problems) than for arithmetic competence in general.

These results add to Bailey, Siegler and Geary's (2014) finding that year one students' whole number arithmetic proficiency (number of correct arithmetic tasks and fast direct retrieval of single-digit additions) predicted year seven knowledge of fraction arithmetic and year eight knowledge of fraction magnitude. The present findings regarding the relationship between use of counting on and year four fraction knowledge may indicate a relationship between strategy use and number sense, as well as conceptual understanding of addition and arithmetic in general. If understanding part-whole relations beyond cardinality is a prerequisite for developing derived fact strategies (and counting on), early use of these strategies may indicate better conceptual understanding of number and arithmetic. If so, this suggests that while assessing arithmetic skills (i.e. number of correct answers) has some predictive power for students' later achievement, their strategies for doing arithmetic may predict general achievement three years later, as well as fractional knowledge, which is known to be predictive of subsequent mathematical achievement (e.g. Bailey, Hoard, Nugent, & Geary, 2012; Siegler et al., 2012).

There was no evidence of any effect of teaching practice on developmental patterns, or of any (random) variation in those patterns between classes (Paper II). This result was somewhat surprising in light of the teacher interviews, which revealed differences of focus and perspective in relation to key aspects of teaching and learning number and arithmetic (Papers IV–V). Based on established learning trajectories (e.g. Clements & Sarama, 2004), their expectations for students' arithmetic competence by the end of year one differed by one to two years of learning (Paper IV). Their approach to strategies and calculation procedures also differed; while one group of teachers emphasised counting and standard algorithms, the other group emphasised derived fact strategies and flexible number-based methods of calculation (Paper V). It could therefore be expected that these differences of perspective in relation to teaching and learning of number and arithmetic would manifest in their students' development and use of strategies.

It is worth noting that in their comparative study of Dutch and Flemish year three to six students, Torbeyns, Hickendorff and Verschaffel (2017) reported an unclear relation between students' strategy use in multi-digit subtraction and instructional differences. They suggested a longitudinal design and use of classroom observation to provide more valid information on the enacted curriculum in the classroom. The present study provided data of this kind on students' development of strategies for single-digit addition, although only for a 5-month timespan. The results suggest that further analysis of classroom teaching and learning opportunities on the present scale would add nothing to the findings, as the contribution from teaching to the total explained variation did not exceed 1% (Paper II).

However, it is possible that Study B (where this analysis was possible) was too short, or that year one was too early for such differences to emerge.

The differences in teaching practice (as expressed in the interviews, Papers IV–V) indicate a lack of consistent knowledge and guidance for teachers in relation to providing opportunities for students to learn and develop arithmetic skills and proficiency. To provide that direction, future studies should seek to determine why strategy use in year one predicts mathematical achievement in year four.

6.1.2 Reflections on sex differences in strategy use

Although sex differences were not initially a primary focus here, their emergence in several response variables (Papers I and III) warrants reflection. On multiple occasions, sex explained much of the variation in response variables (i.e. strategy use from year one [Paper I] and achievement in year four [Paper III]), either as a main effect or through interaction with other predictors. The inclusion of sex as a predictor, where statistically justified, leads to more precise description and prediction of overall patterns for single observations. The sometimes substantial differences between boys' and girls' response patterns also invite reflection, as these overall mean differences point to underlying processes that remain to be elucidated. Identifying these drivers of sex differences are likely to be of both academic and didactic interest because this may tell us more about the underlying mechanisms.

As noted in the discussion of Paper I, it became clear during the search for comparable results in the literature that many papers do not explicitly address sex as an explanatory factor in their analyses of strategy use. As a result, it was difficult to assess whether the results we reported were unique in an international context or part of a general pattern across cultures. The observation that sex differences are often neglected indicates that sex-specific response patterns may be more widespread than the literature suggests, simply because they are not searched for. As a humble proposal for future research on strategy use and mathematical achievement, sex should always be included as a potential predictor, and the result, positive or negative, should always be reported (Paper I).

Systematic and stringent reporting of the extent to which sex differences appear in studies worldwide would provide a fuller picture of the prevalence of sex differences in mathematical achievement and strategy use and would enhance the basis for meta-analyses of the extent to which sex differences are universal or culturally conditioned. As discussed by Halpern et al. (2007), it is practically impossible to separate the impact of biological and environmental factors. However, if sex differences are entirely culturally driven ('social constructs'), the appearance and direction of such differences should vary widely across countries and cultures and should ultimately be explained by the specific socio-cultural context. Similarly, if sex differences in a given response variable are purely biological, they should remain similar in direction when investigated in a standardised way, regardless of socio-cultural context. Finally, if sexual variation is partly a matter of underlying biological variables while also influenced by the socio-cultural context in which children are raised and learn mathematics, sex differences are likely to appear in some situations and not in others but in the same direction overall. As discussed in Paper I, this appears to be the case for single-digit addition strategies, as any observed sex differences always point in the same direction (Bailey, Littlefield, & Geary, 2012; Carr & Alexeev, 2011; Carr & Davis, 2001; Imbo & Vandierendonck, 2007; Shen, Vasilyeva, & Laski, 2016).

As previous research has indicated a relationship between the use of different counting strategies and conceptual knowledge about counting and addition principles (Canobi, Reeve, & Pattison, 1998; Geary, Bow-Thomas, & Yao, 1992), it has been suggested that differences in boys' and girls' preferences for counting all could relate to differences in conceptual understanding (Carr & Davis, 2001). The present findings suggest that strategy use is more predictive for boys' later achievement than for girls' (see discussion of Paper III); this may indicate that the reported sex differences in early years strategy use is not necessarily or entirely related to their understanding of number and arithmetic per se but may reflect girls' greater inclination to use counting even when they have mastered retrieval strategies. One possible explanation relates to differences in domain-general cognitive factors (e.g. visuospatial skills and preferred orientation pathways), where women are known to adopt a more serial strategy (Coutrot et al., 2018; Pintzka, Evensmoen, Lehn, & Håberg, 2016). If that is so, the inclination to use a counting all strategy to solve single-digit addition problems in the first half of year one should correlate positively with psychological test measures of inclination to use serial problem solving strategies and negatively with spatial skills when controlling for general cognitive function.

In conclusion, although technically and statistically challenging (Halpern et al., 2007), it should be possible to statistically pinpoint sex differences in mathematical achievement and strategy use to obtain a deeper understanding of the mechanisms underlying sex-differences (both in this study and more generally), given at least some relevant measures for the individuals in question. From a didactic perspective, quantifying and explaining sex differences offers a means of understanding the drivers of individual behaviour and achievement.

6.2 Methodological considerations and limitations

The following discussion addresses whether the present research might have reached different or more robust conclusions if I had made different methodological choices. In my view, the most relevant methodological issues in this regard are (i) the validity and reliability of the data gathered using the chosen interview methods, (ii) how strategy use was categorized and quantified, and (iii) registration of relevant cognitive variables as predictors of individual variation in strategy use and achievement.

6.2.1 Validity and reliability of interview methods

Coding of interview data

Reliability is central to the replicability of research and can be enhanced in several ways when collecting and analysing interview data (Cohen, Manion, & Morrison, 2011). With respect to the analysis (i.e. coding) of responses, interrater reliability is a common measure for ensuring consistency of scoring. In the present study, resource constraints precluded any formal testing of interrater reliability of teacher interviews or student assessment interviews. While relevant in both cases, the process and complexity of scoring is somewhat different for each.

The interpretation and categorisation of teachers' utterances in the semi-structured interviews is far more complex than the categorisation of students' solution methods in the structured assessment interview (see section 4.2.1). When analysing the teacher interviews, I discussed examples of categories and utterances with my co-author. However, this would not suffice to ensure consistent scoring of the interview data, and interrater reliability should ideally have been formally assessed for both teacher and student assessment interviews.

Measuring strategy use

Students' choice of strategy was not scored for incorrect answers, which would have provided more detailed information about strategy development and strategy efficiency (e.g. Lemaire & Siegler, 1995). Furthermore, because the analytical approach involved analysing the proportional strategy use of a single strategy against all others (including the error category), there is a risk that strategy use was underestimated. However, as discussed in Paper I, it is likely that this led to an underestimation of the girls' use of counting, as Laski et al. (2013) found that girls were least accurate when using counting as compared to other strategies.

As addressed in Chapter 2 (and statistically demonstrated in Paper I, Table 3), strategy choices are widely dependent on the type of problem to be solved. When comparing the present results with those from other studies, it is important to keep in mind that overall strategy patterns also depend on the 36 items chosen for the strategy assessment interviews. As noted in section 4.2 I, the items were restricted to single-digit addition tasks only; it is possible that using more complex problems or imposing time restrictions might have prompted more sophisticated strategy use or provided more detailed information on individual variation and change in strategy use over time (e.g. Carr & Alexeev, 2011; Vasilyeva et al., 2015), especially among students proficient in single-digit addition.

By using a broad array of 36 single-digit items from among the available single-digit problems, ranging from the simple (2 + 3) to the more complicated (8 + 5 or 7 + 8), I was able to obtain a broad spectrum of the strategy repertoire of individual year one students, assessed by frequency of strategy use. As demonstrated in Paper III, the multitude of strategy use frequencies provided information that could be further condensed in the form of 1–3 principal component axes. To date, the predictive potential of strategy use profiles for mathematical achievement in subsequent years has only been tested for the first half of year one, where the strongest predictor was high use of counting all (Paper III). The extent to which strategy use profiles in years two, three and four can provide similar information (and for which strategies) remains to be investigated.

6.2.2 Categorisation and quantification of strategy use

I initially employed seven strategy use categories, of which five were genuine problem solving strategies and two (give up and error) related to attempts that did not result in a correct answer. As described in section 4.2.1 (Table 4.6), the two counting strategies (counting all and counting on) were pooled for methodological reasons in Papers I and II. It was not possible to establish statistical models for the counting all data from study A because of convergence problems, as most students did not use this strategy after year one. For the purposes of analysis, merging the two categories was justified by the observation that for the combined data from Study A and B (which enabled separate analysis of each category), girls in year one made equal use of both counting strategies more often than boys (Paper I). From this result, it seemed clear that the two counting strategies were two sides of the same coin, and the combined counting category was also used in the analysis of Paper II.

During the data analyses in Paper III, which commenced after Paper II was finally accepted, it became clear that the proportional use in early year one of counting all and counting on correlated negatively and yielded opposing predictions for mathematical achievement three years later. In contrast to the high use of counting all, high use of counting on correlated positively with all three types of mathematical achievement in year four. Similarly, the two counting strategies correlated differently with the primary PCA axis expressing the gradient of sophistication in strategy use. On that basis, counting on in the first half of year one should not be considered an indicator of unsophisticated strategy use and arithmetic comprehension but rather the opposite. As a result, the combined general counting strategy (summing counting all and counting on) had limited predictive power in relation to mathematical achievement in year four simply because the two counting strategies had opposing effects. That being so, it may well be that the results in Paper II would have been more nuanced if counting had been split into two categories.

6.2.3 Cognitive factors

In analysing the development and effect of specific mathematical competences and skills, it is relevant to include cognitive variables as predictors to explain individual variation in the response variable of interest. In relation to mental arithmetic strategies, relevant cognitive factors might include intelligence, visuospatial memory and memory span. In Study B, I initially collected cognitive development data using the Danish Children's Problem Solving (CHIPS) test (Hansen, Kreiner, Petersen, & Hansen, 1993). This test, which is widely used by teachers in Danish primary schools, assigns the child to one of four groups: pre-operational (1 and 2), concrete operational or formal operational. Unfortunately, as I later realized, this widely used test is not scientifically validated as an assessment tool for cognitive function per se and could not therefore be considered adequate for this purpose.

As several factors has been shown to be related to students arithmetic competence, for example number comparison skills (e.g. Linsen, Verschaffel, Reynvoet, & De Smedt, 2015; Vanbinst, Ghesquière, & De Smedt, 2012), and estimation skills (e.g. Booth & Siegler, 2008; Cowan & Powell, 2014), other relevant covariates would have been measures of number sense. However, the amount of testing will always be a trade-off between valuable teaching time and discomfort for children and teachers and pursuit of a predictor variable; for these reasons, no rigorous measures of cognitive function and number sense were applied. Retrospectively, given the high level of individual variation revealed here, it is regrettable that no such auxiliary predictors were available to explain the underlying reasons for this variation, including differences between boys and girls.

6.3 Implications for teaching

I believe the present findings have implications for teaching at two levels. First, students' strategy use 'profiles' from the assessment interviews can indicate a student's risk of developing mathematical difficulties or becoming a low achiever. In that respect, excessive use of counting strategies (especially counting all) in the early years of school may have particular value as a screening parameter. Early detection of students who need special attention would mean that targeted teaching can be initiated before potential difficulties arise.

Second, if the strong correlation between strategy use patterns in early year one and achievement in year four reflects a causal relation (which remains to be investigated), teaching practice should focus on developing number sense and flexible strategy use for arithmetic. Given the clear and substantial individual variation in strategy use at the beginning of formal teaching in year one, number sense development should probably be addressed in kindergarten or preschool. Targeted and differentiated teaching should be designed for different developmental profiles and not just for different achievement levels. When focussing on the explicit teaching of strategies it is important to be aware that the teaching do not become instrumental as the teaching of specific strategies can lead to these strategies becoming 'default' methods (Blöte, Klein, & Beishuizen, 2000; Threlfall, 2002).

6.4 Future directions

As is usual in primary research, creating new knowledge generates new questions, and a number of research questions remain unanswered or have been generated by the present findings. In my view, the following questions are of particular interest for future research.

1. What are the early underlying drivers of individual variation (including sex differences) in year one strategy use?
2. To what extent is the link between strategy use in early year one and later mathematical achievement causal?
3. How do differences in teaching practices influence strategy use for improved subsequent mathematical achievement?

The first and second questions arise from the finding that students varied considerably in strategy use at the earliest time of measurement (i.e. beginning of year one), and that this variation was linked to later mathematical achievement. The third question relates to the first in the sense that my results so far suggest that the foundation for students' strategy use is established prior to year one—in other words, strategies are hardwired at individual level and are not influenced by differences in teaching style. Despite the evidence, one might ask whether that can really be—that is, have I missed something in the present research? Question 3 also links to question 2 in the sense that improving students' mathematical achievement by teaching them strategies to develop their number skills will only matter if the link between early strategy use and later mathematical achievement is causal and subject to differential training.

To answer the first two questions, more detailed data are needed in relation to individual students' cognitive functions, strategy use and achievement in different types of mathematics over an even longer timespan—ideally, from kindergarten to the end of school—than in this investigation. Answering the third question will require carefully designed intervention studies. As essential information on cognitive function in year one was not collected and targeted intervention programs for year one or earlier have not yet been designed or implemented in the present context, a new study on a new cohort of students, ideally followed from kindergarten, will be needed to fully address these emergent research questions. At least in principle, some additional information of relevance to question 2 may be acquired by following the students in Study B in the years to come.

References

Adriansen, H. K. (2012). Timeline interviews: A tool for conducting life history research. *Qualitative Studies*, *3*(1), 40–55. https://doi.org/10.7146/qs.v3i1.6272

Ambrose, R., Baek, J.-M., & Carpenter, T. P. (2003). Children's invention of multidigit multiplication and division algorithms. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise.* (pp. 305–336). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Andrews, P., & Sayers, J. (2015). Identifying Opportunities for Grade One Children to Acquire Foundational Number Sense: Developing a Framework for Cross Cultural Classroom Analyses. *Early Childhood Education Journal*, *43*(4), 257–267. https://doi.org/10.1007/s10643-014-0653-6

Arcavi, A. (2000). Problem-driven research in mathematics education. *The Journal of Mathematical Behavior, 19*(2), 141–173. https://doi.org/10.1016/S0732-3123(00)00042-0

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review*, *2*(3), 213–236. https://doi.org/10.1016/0273-2297(82)90012-0

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*(1–2), 75–106. https://doi.org/10.1016/0010-0277(92)90051-I

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition*, *1*(1), 3–34. https://doi.org/10.1145/1959826.1959829

Aunio, P., Ee, J., Lim, S. E. A., Hautamäki, J., & Van Luit, J. (2004). Young children's number sense in Finland, Hong Kong and Singapore. *International Journal of Early Years Education*, *12*(3), 195–216. https://doi.org/10.1080/0966976042000268681

Aunio, P., & Niemivirta, M. (2010). Predicting children's mathematical performance in grade one by early numeracy. *Learning and Individual Differences*, *20*(5), 427–435. https://doi.org/10.1016/j.lindif.2010.06.003

Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The codevelopment of skill at and preference for use of retrieval-based processes for solving addition problems: Individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology*, *113*(1), 78–92. https://doi.org/10.1016/j.jecp.2012.04.014

Ball, D. L. (1988). *Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education*. Doctoral dissertation, Michigan State University.

Baroody, A. J. (1987). The Development of Counting Strategies for Single-Digit Addition. *Journal for Research in Mathematics Education*, *18*(2), 141–157. https://doi.org/10.2307/749248

Baroody, A. J. (1989). Kindergartners' Mental Addition with Single-Digit Combinations. *Journal for Research in Mathematics Education*, *20*(2), 159–172. https://doi.org/10.2307/749280

Baroody, A. J., & Dowker, A. (Eds.). (2003). *The development of arithmetic concepts and skills: Constructing adaptive expertise*. https://doi.org/10.4324/9781410607218

Baroody, A. J., & Wilkins, J. L. M. (1999). The development of informal counting, number, and arithmetic skills and concepts. In J. V. Copley (Ed.), *Mathematics in the early years.*

(pp. 48–65). Washington, DC, US: National Association for the Education of Young Children.
Bjorklund, D. F., & Rosenblum, K. E. (2002). Context Effects in Children's Selection and Use of Simple Arithmetic Strategies. *Journal of Cognition and Development*, *3*(2), 225–242. https://doi.org/10.1207/S15327647JCD0302_5
Blöte, A. W., Klein, A. S., & Beishuizen, M. (2000). Mental computation and conceptual understanding. *Learning and Instruction*, *10*(3), 221–247. https://doi.org/10.1016/S0959-4752(99)00028-6
Booth, J. L., & Siegler, R. S. (2008). Numerical Magnitude Representations Influence Arithmetic Learning. *Child Development*, *79*(4), 1016–1031. https://doi.org/10.1111/j.1467-8624.2008.01173.x
Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: SAGE.
Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General*, *130*(2), 299–315. https://doi.org/10.1037/0096-3445.130.2.299
Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development*, *19*(1), 81–93. https://doi.org/10.1016/j.cogdev.2003.10.001
Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology*, *34*(5), 882–891. https://doi.org/10.1037/0012-1649.34.5.882
Carpenter, T. P., Hiebert, J., & Moser, J. M. (1981). Problem structure and first-grade children's initial solution processes for simple addition and subtraction problems. *Journal for Research in Mathematics Education*, *12*(1), 27–39. https://doi.org/10.2307/748656
Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education*, *15*(3), 179–202. https://doi.org/10.2307/748348
Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology*, *103*(3), 617–631. https://doi.org/10.1037/a0023864
Carr, M., & Davis, H. (2001). Gender Differences in Arithmetic Strategy Use: A Function of Skill and Preference. *Contemporary Educational Psychology*, *26*(3), 330–347. https://doi.org/10.1006/ceps.2000.1059
Carr, M., Jessup, D. L., & Fuller, D. (1999). Gender Differences in First-Grade Mathematics Strategy Use: Parent and Teacher Contributions. *Journal for Research in Mathematics Education*, *30*(1), 20–46. https://doi.org/10.2307/749628
Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, *18*(1), 61–75. https://doi.org/10.1016/j.lindif.2007.04.005
Carraher, T. N., Carraher, D. W., & Schliemann, A. D. (1985). Mathematics in the streets and in schools. *British Journal of Developmental Psychology*, *3*(1), 21–29. https://doi.org/10.1111/j.2044-835X.1985.tb00951.x
Casey, B. M., Lombardi, C. M. P., Pollock, A., Fineman, B., & Pezaris, E. (2017). Girls' Spatial Skills and Arithmetic Strategies in First Grade as Predictors of Fifth-Grade Analytical Math Reasoning. *Journal of Cognition and Development*, *18*(5), 530–555. https://doi.org/10.1080/15248372.2017.1363044
Casey, B. M., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse

samples. *Developmental Psychology*, *31*(4), 697–705. https://doi.org/10.1037/0012-1649.31.4.697

Chang, A., Sandhofer, C. M., & Brown, C. S. (2011). Gender biases in early number exposure to preschool-aged children. *Journal of Language and Social Psychology*, *30*(4), 440–450. https://doi.org/10.1177/0261927X11416207

Cheng, Y.-L., & Mix, K. S. (2014). Spatial Training Improves Children's Mathematics Ability. *Journal of Cognition and Development*, *15*(1), 2–11. https://doi.org/10.1080/15248372.2012.725186

Chesney, D. L., McNeil, N. M., Matthews, P. G., Byrd, C. E., Petersen, L. A., Wheeler, M. C., … Dunwiddie, A. E. (2014). Organization matters: Mental organization of addition knowledge relates to understanding math equivalence in symbolic form. *Cognitive Development*, *30*(1), 30–46. https://doi.org/10.1016/j.cogdev.2014.01.001

Chu, F. W., VanMarle, K., & Geary, D. C. (2015). Early numerical foundations of young children's mathematical development. *Journal of Experimental Child Psychology*, *132*, 205–212. https://doi.org/10.1016/j.jecp.2015.01.006

Clements, D. H. (1999). Subitizing: what is it? Why teach it? *Teaching Children Mathematics*, *5*(7), 400–406.

Clements, D. H., & Sarama, J. (2007). Early childhood mathematics learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 461–555). Greenwich, CT: Information Age Publishing.

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (6th ed.). New York, NY: Routledge.

Coutrot, A., Silva, R., Manley, E., de Cothi, W., Sami, S., Bohbot, V. D., … Spiers, H. J. (2018). Global Determinants of Navigation Ability. *Current Biology*, *28*(17), 2861-2866.e4. https://doi.org/10.1016/j.cub.2018.06.009

Cowan, R. (2003). Does it all add up? Changes in children's knowledge of addition combinations, strategies, and principles. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise.* (pp. 35–74). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Cowan, R., Donlan, C., Shepherd, D.-L., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation proficiency and mathematics achievement in elementary school children. *Journal of Educational Psychology*, *103*(4), 786–803. https://doi.org/10.1037/a0024556

Cowan, R., & Powell, D. (2014). The contributions of domain-general and numerical factors to third-grade arithmetic skills and mathematical learning disability. *Journal of Educational Psychology*, *106*(1), 214–229. https://doi.org/10.1037/a0034097

Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed method approaches* (2nd ed.). Thousand Oaks, CA: SAGE.

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Dehaene, S. (2011). *The number sense: How the mind creates mathematics.* (Rev. and u). New York, NY: Oxford University Press.

Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences*, *16*(4), 351–367. https://doi.org/10.1016/j.lindif.2006.12.006

Dowker, A. (1998). Individual differences in normal arithmetical development. In C. Donlan (Ed.), *The development of mathematical skills.* (pp. 275–302). Hove, UK: Psychology Press/Taylor & Francis (UK).

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development*, *24*(4), 401–410. https://doi.org/10.1016/j.cogdev.2009.09.005
Dowker, A. (2014). Young children's use of derived fact strategies for addition and subtraction. *Frontiers in Human Neuroscience*, *7*, 924. https://doi.org/10.3389/fnhum.2013.00924
Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., … Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428
Ellis, T. J., & Levy, Y. (2008). Framework of Problem-Based Research: A Guide for Novice Researchers on the Development of a Research-Worthy Problem. *Informing Science*, *11*, 17–33.
Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis. *Psychological Bulletin*, *136*(1), 103–127. https://doi.org/10.1037/a0018053
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002
Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A Longitudinal Study of Gender Differences in Young Children's Mathematical Thinking. *Educational Researcher*, *27*(5), 6–11. https://doi.org/10.3102/0013189X027005006
Flick, U. (2014). *An introduction to qualitative research* (5th ed.). Thousand Oaks, CA: SAGE.
Foley, A. E., Vasilyeva, M., & Laski, E. V. (2017). Children's use of decomposition strategies mediates the visuospatial memory and arithmetic accuracy relation. *British Journal of Developmental Psychology*, *35*(2), 303–309. https://doi.org/10.1111/bjdp.12166
Frank, L. (2019). Giftig viden. *Weekendavisen*, p. A16.
Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., … Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29
Fuson, K. C. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics.* (pp. 243–275). New York, NY, England: Macmillan Publishing Co, Inc.
Fuson, K. C., & Burghardt, B. H. (2003). Multidigit addition and subtraction methods invented in small groups and teacher support of problem solving and reflection. In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise* (pp. 267–307). https://doi.org/10.4324/9781410607218
Gaidoschik, M. (2012). First-Graders' Development of Calculation Strategies: How Deriving Facts Helps Automatize Facts. *Journal Für Mathematik-Didaktik*, *33*(2), 287–315. https://doi.org/10.1007/s13138-012-0038-6
Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology*, *47*(6), 1539–1552. https://doi.org/10.1037/a0025510
Geary, D. C. (2013). Early Foundations for Mathematics Learning and Their Relations to Learning Disabilities. *Current Directions in Psychological Science*, *22*(1), 23–27. https://doi.org/10.1177/0963721412469398
Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of Arithmetical Competencies in Chinese and American Children: Influence of Age,

Language, and Schooling. *Child Development*, *67*(5), 2022–2044. https://doi.org/10.1111/j.1467-8624.1996.tb01841.x
Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88*(2), 121–151. https://doi.org/10.1016/j.jecp.2004.03.002
Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Child Psychology*, *113*(1), 49–65. https://doi.org/10.1016/j.jecp.2012.03.003
Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLOS ONE*, *8*(1), e54651. Retrieved from https://doi.org/10.1371/journal.pone.0054651
Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*(4), 293–304. https://doi.org/10.1177/00222194050380040301
Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*, 589–591. https://doi.org/10.1038/nature05850
Goodchild, S. (2013). Reporting classroom research: a moral dilemma. In B. Grevholm, P. S. Hundeland, K. Juter, K. Kislenko, & P. E. Persson (Eds.), *Nordic research in didactics of mathematics: Past, present and future.* (pp. 199–224). Oslo, Norway: Cappelen Damm.
Grabner, R. H., & De Smedt, B. (2011). Neurophysiological evidence for the validity of verbal strategy reports in mental arithmetic. *Biological Psychology*, *87*(1), 128–136. https://doi.org/10.1016/j.biopsycho.2011.02.019
Gray, D. E. (2014). *Doing research in the real world* (3rd ed.). London: SAGE.
Gray, E. M. (1991). An analysis of diverging approaches to simple arithmetic: Preference and its consequences. *Educational Studies in Mathematics*, *22*(6), 551–574. https://doi.org/10.1007/BF00312715
Gray, E. M., & Tall, D. G. (1994). Duality, Ambiguity, and Flexibility: A "Proceptual" View of Simple Arithmetic. *Journal for Research in Mathematics Education*, *25*(2), 116–140. https://doi.org/10.2307/749505
Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, *79*(4), 329–343. https://doi.org/10.1037/h0032950
Hailikari, T., Nevgi, A., & Komulainen, E. (2008). Academic self-beliefs and prior knowledge as predictors of student achievement in Mathematics: a structural model. *Educational Psychology*, *28*(1), 59–71. https://doi.org/10.1080/01443410701413753
Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "number sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*(5), 1457–1465. https://doi.org/10.1037/a0012682
Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The Science of Sex Differences in Science and Mathematics. *Psychological Science in the Public Interest*, *8*(1), 1–51. https://doi.org/10.1111/j.1529-1006.2007.00032.x
Hatano, G. (2003). Foreword. In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise* (pp. xi–xiii). Mahwah, NJ: Lawrence Erlbaum Associates.
Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing School Learning Using Insight From Expertise Research. *Educational Researcher*, *32*(8), 26–29. https://doi.org/10.3102/0013189X032008026

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice.* New York, NY, US: Guilford Press.

Hickendorff, M., Torbeyns, J., & Verschaffel, L. (2018). Grade-related differences in strategy use in multidigit division in two instructional settings. *British Journal of Developmental Psychology*, *36*(2), 169–187. https://doi.org/10.1111/bjdp.12223

Hopkins, K. B., McGillicuddy-De Lisi, A. V., & Lisi, R. De. (1997). Student gender and teaching methods as sources of variability in children's computational arithmetic performance. *Journal of Genetic Psychology*, *158*(3), 333–345. https://doi.org/10.1080/00221329709596672

Hutchison, J. E., Lyons, I. M., & Ansari, D. (2019). More Similar Than Different: Gender Differences in Children's Basic Numerical Skills Are the Exception Not the Rule. *Child Development*, *90*(1), e66–e79. https://doi.org/10.1111/cdev.13044

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist*, *60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin*, *107*(2), 139–155. https://doi.org/10.1037/0033-2909.107.2.139

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(22), 8801–8807. https://doi.org/10.1073/pnas.0901265106

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology*, *96*(4), 284–309. https://doi.org/10.1016/j.jecp.2006.09.001

Jensen, P. E., & Jørgensen, I.-L. (2007). *Vejledning til mat-prøver: pædagogisk analyse af matematik*. Virum, Denmark: Dansk psykologisk Forlag.

Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number Sense Growth in Kindergarten: A Longitudinal Investigation of Children at Risk for Mathematics Difficulties. *Child Development*, *77*(1), 153–175. https://doi.org/10.1111/j.1467-8624.2006.00862.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850–867. https://doi.org/10.1037/a0014939

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: helping children learn mathematics*. Washington: National Academy Press.

Krinzinger, H., Wood, G., & Willmes, K. (2012). What Accounts for Individual and Gender Differences in the Multi-Digit Number Processing of Primary School Children? *Zeitschrift Für Psychologie*, *220*(2), 78–89. https://doi.org/10.1027/2151-2604/a000099

Lachance, J. A., & Mazzocco, M. M. M. (2006). A longitudinal analysis of sex differences in math and spatial skills in primary school age children. *Learning and Individual Differences*, *16*(3), 195–216. https://doi.org/10.1016/j.lindif.2005.12.001

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8–9-year-old students. *Cognition*, *93*(2), 99–125. https://doi.org/10.1016/j.cognition.2003.11.004

Laski, E. V., Casey, B. M., Yu, Q., Dulaney, A., Heyman, M., & Dearing, E. (2013). Spatial skills as a predictor of first grade girls' use of higher level arithmetic strategies. *Learning and Individual Differences*, *23*(1), 123–130. https://doi.org/10.1016/j.lindif.2012.08.001

Laski, E. V., Ermakova, A., & Vasilyeva, M. (2014). Early use of decomposition for addition and its relation to base-10 knowledge. *Journal of Applied Developmental Psychology*,

35(5), 444–454. https://doi.org/10.1016/j.appdev.2014.07.002
LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 216–230. https://doi.org/10.1037/0278-7393.22.1.216
Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, *124*(1), 83–97. https://doi.org/10.1037/0096-3445.124.1.83
Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology*, *35*(4), 940–949. https://doi.org/10.1037/0012-1649.35.4.940
Linsen, S., Verschaffel, L., Reynvoet, B., & De Smedt, B. (2015). The association between numerical magnitude processing and mental versus algorithmic multi-digit subtraction in children. *Learning and Instruction*, *35*, 42–50. https://doi.org/10.1016/j.learninstruc.2014.09.003
Littell, R. C., Milliken, G. A., Stroup, W. W., Wolfinger, R. D., & Schabenberber, O. (2006). *SAS for Mixed Models* (second ed.). Cary, NC: SAS Publishing.
Luke, D. A. (2004). *Multilevel Modeling*. https://doi.org/10.4135/9781412985147
Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, *17*(5), 714–726. https://doi.org/10.1111/desc.12152
Ma, L. (2010). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States: Second Edition* (second ed.). https://doi.org/10.4324/9780203856345
Mazzocco, M. M. M., & Thompson, R. E. (2005). Kindergarten Predictors of Math Learning Disability. *Learning Disabilities Research & Practice*, *20*(3), 142–155. https://doi.org/10.1111/j.1540-5826.2005.00129.x
Mazzocco, M., Murphy, M., Brown, E., Rinne, L., & Herold, K. (2013). Persistent consequences of atypical early number concepts. *Frontiers in Psychology*, *4*, 486. https://doi.org/10.3389/fpsyg.2013.00486
Ministry of Higher Education and Science. (2014). Danish Code of Conduct for Research Integrity. Retrieved March 16, 2017, from http://ufm.dk/publikationer/2014/filer-2014/the-danish-code-of-conduct-for-research-integrity.pdf
Moè, A. (2018). Mental rotation and mathematics: Gender-stereotyped beliefs and relationships in primary school children. *Learning and Individual Differences*, *61*, 172–180. https://doi.org/10.1016/j.lindif.2017.12.002
Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. Retrieved from http://timssandpirls.bc.edu/timss2015/international-results/
Niiniluoto, I. (1993). The Aim and Structure of Applied Research. *Erkenntnis*, *38*(1), 1–21.
Nunes, T., Dorneles, B. V., Lin, P.-J., & Rathgeb-Schnierer, E. (2016). *Teaching and Learning About Whole Numbers in Primary School*. https://doi.org/10.1007/978-3-319-45113-8_1
OECD. (2010). *The High Cost of Low Educational Performance*. https://doi.org/10.1787/9789264077485-en
OECD. (2016). *PISA 2015 Results (Volume I)*. https://doi.org/10.1787/9789264266490-en
Ostad, S. A. (1997a). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology*, *67*(3), 345–357. https://doi.org/10.1111/j.2044-8279.1997.tb01249.x

Ostad, S. A. (1997b). Strategic competence: issues of task-specific strategies in arithmetic. *Nordic Studies in Mathematics Education*, *3*, 7–32.

Parsons, S., & Bynner, J. (2005). *Does Numeracy Matter More*. London: National Research and Development Centre for Adult Literacy and Numeracy.

Paul, J. M., & Reeve, R. A. (2016). Relationship between single digit addition strategies and working memory reflects general reasoning sophistication. *Learning and Instruction*, *42*, 113–122. https://doi.org/10.1016/j.learninstruc.2016.01.011

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research*, *37*(1), 239–253. https://doi.org/10.1016/j.ssresearch.2007.06.012

Pintzka, C. W. S., Evensmoen, H. R., Lehn, H., & Håberg, A. K. (2016). Changes in spatial cognition and brain activity after a single dose of testosterone in healthy women. *Behavioural Brain Research*, *298*, 78–90. https://doi.org/10.1016/j.bbr.2015.10.056

Powell, E. (2005). Conceptualising and facilitating active learning: teachers' video-stimulated reflective dialogues. *Reflective Practice*, *6*(3), 407–418. https://doi.org/10.1080/14623940500220202

Price, G. R., Mazzocco, M. M. M., & Ansari, D. (2013). Why Mental Arithmetic Counts: Brain Activation during Single Digit Arithmetic Predicts High School Math Scores. *Journal of Neuroscience*, *33*(1), 156–163. https://doi.org/10.1523/JNEUROSCI.2936-12.2013

Rechtsteiner-Merz, C., & Rathgeb-Schnierer, E. (2015). Flexible mental calculation and "Zahlenblickschulung" In Konrad Krainer; Naďa Vondrová (Ed.), *CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education* (pp. 354–360). Retrieved from https://hal.archives-ouvertes.fr/hal-01281864/

Rezat, S., & Ejersbo, L. R. (2018). Number sense in teaching and learning arithmetic. In T. Dreyfus, M. Artigue, D. Potari, S. Prediger, & K. Ruthven (Eds.), *Developing Research in Mathematics Education* (pp. 23–21). https://doi.org/10.4324/9781315113562

Rittle-Johnson, B., & Siegler, R. S. (1999). Learning to Spell: Variability, Choice, and Change in Children's Strategy Use. *Child Development*, *70*(2), 332–348. https://doi.org/10.1111/1467-8624.00025

Robson, C. (2011). *Real world research : a resource for users of social research methods in applied settings* (3rd ed.). Chichester, West Sussex: Wiley.

Ruspini, E. (2002). *An Introduction to Longitudinal Research*. https://doi.org/10.4324/9780203167229

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research : learning trajectories for young children*. New York: Routledge.

Sayers, J., & Andrews, P. (2015). Foundational Number Sense : Summarising the Development of an Analytical framework. In N. V. Konrad Krainer & N. V. Konrad Krainer (Eds.), *CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Prague, Czech Republic, February 2015* (pp. 361–367). Retrieved from https://hal.archives-ouvertes.fr/hal-01286842

Shen, C., Vasilyeva, M., & Laski, E. V. (2016). Here, but not there: Cross-national variability of gender effects in arithmetic. *Journal of Experimental Child Psychology*, *146*(50–65). https://doi.org/10.1016/j.jecp.2016.01.016

Siegler, R. S. (1987). The Perils of Averaging Data Over Strategies: An Example From Children's Addition. *Journal of Experimental Psychology: General*, *116*(3), 250–264. https://doi.org/10.1037/0096-3445.116.3.250

Siegler, R. S. (1988). Individual differences in strategy choices: good students, not-so-good students, and perfectionists. *Child Development*, *59*(4), 833–851. https://doi.org/10.1111/j.1467-8624.1988.tb03238.x

Siegler, R. S. (1989). Strategy Diversity and Cognitive Assessment. *Educational Researcher*, *18*(9), 15–20. https://doi.org/10.3102/0013189X018009015

Siegler, R. S. (1996). Emerging minds: The process of change in children's thinking. In *Emerging minds: The process of change in children's thinking.* New York, NY, US: Oxford University Press.

Siegler, R. S., & Jenkins, E. (1989). How children discover new strategies. In *How children discover new strategies.* Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Siegler, R. S., & Lemaire, P. (1997). Older and Younger Adults' Strategy Choices in Multiplication: Testing Predictions of ASCM Using the Choice/No-Choice Method. *Journal of Experimental Psychology: General*, *126*(1), 71–92. https://doi.org/10.1037/0096-3445.126.1.71

Siegler, R. S., & Shrager, J. (1984). A model of strategy choice. In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229–293). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.

Simons, H. (2009). *Case Study Research in Practice*. https://doi.org/10.4135/9781446268322

Steinberg, R. M. (1985). Instruction on derived facts strategies in addition and subtraction. *Journal for Research in Mathematics Education*, *16*(5), 337–355. https://doi.org/10.2307/749356

Sternberg, R. J. (2007). The importance of problem-driven research: Bringing Wachtel's argument into the present. *Applied and Preventive Psychology*, *12*(1), 37–38. https://doi.org/10.1016/j.appsy.2007.07.011

Stoet, G., & Geary, D. C. (2013). Sex Differences in Mathematics and Reading Achievement Are Inversely Related: Within- and Across-Nation Assessment of 10 Years of PISA Data. *PLoS ONE*, *8*(3), e57988. https://doi.org/10.1371/journal.pone.0057988

Stoet, G., & Geary, D. C. (2015). Sex differences in academic achievement are not related to political, economic, or social equality. *Intelligence*, *48*, 137–151. https://doi.org/10.1016/j.intell.2014.11.006

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory, 2nd ed.* Thousand Oaks, CA: SAGE.

Sunde, P. B., & Pind, P. (2016). Comparison of two test approaches for detecting mathematical difficulties. In L. Lindenskov (Ed.), *Special needs in mathematics education* (pp. 141–156). København: DPU, Aarhus Universitet.

Thevenot, C., Barrouillet, P., Castel, C., & Uittenhove, K. (2016). Ten-year-old children strategies in mental addition: A counting model account. *Cognition*, *146*, 48–57. https://doi.org/10.1016/j.cognition.2015.09.003

Thornton, C. A. (1978). Emphasizing Thinking Strategies in Basic Fact Instruction. *Journal for Research in Mathematics Education*, *9*(3), 214–227. https://doi.org/10.2307/748999

Threlfall, J. (2002). Flexible Mental Calculation. *Educational Studies in Mathematics*, *50*(1), 29–47. https://doi.org/10.1023/A:1020572803437

Timmermans, R. E., Van Lieshout, E. C. D. M., & Verhoeven, L. (2007). Gender-related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction*, *17*(1), 42–54. https://doi.org/10.1016/J.LEARNINSTRUC.2006.11.005

Torbeyns, J., Arnaud, L., Lemaire, P., & Verschaffel, L. (2004). Cognitive change as strategy change. In A. Demetriou & A. Raftopoulos (Eds.), *Cognitive developmental change: Theories, models and measurement* (pp. 186–216). New York, NY, US: Cambridge University Press.

Torbeyns, J., Hickendorff, M., & Verschaffel, L. (2017). The use of number-based versus digit-based strategies on multi-digit subtraction: 9–12-year-olds' strategy use profiles and task performance. *Learning and Individual Differences*, *58*, 64–74.

https://doi.org/10.1016/j.lindif.2017.07.004
Torbeyns, J., & Verschaffel, L. (2016). Mental computation or standard algorithm? Children’s strategy choices on multi-digit subtractions. *European Journal of Psychology of Education*, *31*(2), 99–116. https://doi.org/10.1007/s10212-015-0255-8
Torbeyns, J., Verschaffel, L., & Ghesquiere, P. (2004a). Efficiency and adaptiveness of multiple school-taught strategies in the domain of simple addition. *Proceedings of the 28th Conference of the International Group for the Psychology of Mathematics Education*.
Torbeyns, J., Verschaffel, L., & Ghesquiere, P. (2004b). Strategy Development in Childrenwith Mathematical Disabilities: Insights from the Choice/No-Choice Method and the Chronological-Age/ Ability-Level—Match Design. *Journal of Learning Disabilities*, *37*(2), 119–131. https://doi.org/10.1177/00222194040370020301
Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2002). Strategic competence: Applying Siegler’s theoretical and methodological framework to the domain of simple addition. *European Journal of Psychology of Education*, *17*(3), 275. https://doi.org/10.1007/BF03173537
Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2004). Strategic aspects of simple addition and subtraction: The influence of mathematical ability. *Learning and Instruction*, *14*(2), 177–195. https://doi.org/10.1016/j.learninstruc.2004.01.003
Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2005). Simple addition strategies in a first-grade class with multiple strategy instruction. *Cognition and Instruction*, *23*(1), 1–21. https://doi.org/10.1207/s1532690xci2301_1
Tournaki, N. (2003). The Differential Effects of Teaching Addition through Strategy Instruction Versus Drill and Practice to Students with and Without Learning Disabilities. *Journal of Learning Disabilities*, *36*(5), 449–458. https://doi.org/10.1177/00222194030360050601
UVM. (2011). PEU-hæfte matematik (Danish Ministry of Education). Retrieved February 25, 2019, from http://www.uvm.dk/folkeskolen/folkeskolens-proever/censur-og-evaluering/evaluering-af-proeverne/tidligere-evaluering-af-proeverne
Van Dooren, W., Lehtinen, E., & Verschaffel, L. (2015). Unraveling the gap between natural and rational numbers. *Learning and Instruction*, *37*, 1–4. https://doi.org/10.1016/j.learninstruc.2015.01.001
Vanbinst, K., Ghesquière, P., & De Smedt, B. (2012). Numerical Magnitude Representations and Individual Differences in Children’s Arithmetic Strategy Use. *Mind, Brain, and Education*, *6*(3), 129–136. https://doi.org/10.1111/j.1751-228X.2012.01148.x
Vanbinst, K., Ghesquière, P., & De Smedt, B. (2014). Arithmetic strategy development and its domain-specific and domain-general cognitive correlates: A longitudinal study in children with persistent mathematical learning difficulties. *Research in Developmental Disabilities*, *35*, 3001–3013. https://doi.org/10.1016/j.ridd.2014.06.023
Vasilyeva, M., Laski, E. V., & Shen, C. (2015). Computational fluency and strategy choice predict individual and cross-national differences in complex arithmetic. *Developmental Psychology*, *51*(10), 1489–1500. https://doi.org/10.1037/dev0000045
Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning.* (pp. 557–628). Greenwich, CT: Information Age Publishing.
Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, *24*(3), 335–359. https://doi.org/10.1007/BF03174765
Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-

analysis. *Psychological Bulletin*, *140*(4), 1174–1204. https://doi.org/10.1037/a0036620

Wu, S. S., Meyer, M. L., Maeda, U., Salimpoor, V., Tomiyama, S., Geary, D. C., & Menon, V. (2008). Standardized Assessment of Strategy Use and Working Memory in Early Mental Arithmetic Performance. *Developmental Neuropsychology*, *33*(3), 365–393. https://doi.org/10.1080/87565640801982445

Wylie, J., Jordan, J. A., & Mulhern, G. (2012). Strategic development in exact calculation: Group and individual differences in four achievement subtypes. *Journal of Experimental Child Psychology*, *113*(1), 112–130. https://doi.org/10.1016/j.jecp.2012.05.005

Yackel, E., & Cobb, P. (1996). Sociomathematical Norms, Argumentation, and Autonomy in Mathematics. *Journal for Research in Mathematics Education*, *27*(4), 458–477. https://doi.org/10.2307/749877

Yin, R. K. (2013). *Case study research: design and methods* (5th ed.). Thousand Oaks, CA: SAGE.

Paper I

Sex Differences in Mental Strategies for Single-Digit Addition in the First Years of School

Pernille B. Sunde[a,b], Peter Sunde[c], and Judy Sayers[d]

[a]Danish School of Education, Aarhus University, Copenhagen, Denmark; [b]VIA University College, Aarhus, Denmark; [c]Department of Bioscience, Aarhus University, Aarhus, Denmark; [d]School of Education, University of Leeds, Leeds, United Kingdom.

Strategy use in single-digit addition is an indicator of young children's numeracy comprehension. We investigated Danish primary students' use of strategies in single-digit addition with interview-based assessment of how they solved 36 specific single-digit addition problems, categorized as either 'error', 'counting', 'direct retrieval' or 'derived facts'. The proportional use of each strategy was analysed as multi-level functions of school age and sex. In a first study (260 interviews, 147 students) we found decreasing use of counting and increasing use of direct retrieval and derived facts through year 1-4, girls using counting substantially more and the other two strategies substantially less than boys, equal to more than two years' development. Similar results appeared in a subsequent study (155 interviews, 83 students), suggesting that the pattern is pervasive in Danish primary schools. Finally, we ask whether sex differences in strategy use is generally under-reported since many studies do not explicitly address them.

Keywords: sex differences; single-digit addition; strategies; years 1 to 4; mathematics.

Introduction

Appreciating sex differences of student competences in mathematics has long been a discussion across time and across borders (Carr, Steiner, Kyser, & Biddlecomb, 2008; Fennema, Carpenter, Jacobs, Franke, & Levi, 1998; Shen, Vasilyeva, & Laski, 2016). The topic continues to attract attention from education in general, mathematics education in particular, as well as more traditional psychological perspectives (Benbow, 1988; Browne, 2002; Connellan, Baron-Cohen, Wheelwright, Batki, & Ahluwalia, 2000; Halpern et al., 2007; Maccoby & Jacklin, 1974; Pinker, 2002). Research findings also attract attention for their diversity and contradiction, where either boys are found to outperform girls (Fennema et al., 1998; Pinxten, Marsh, De Fraine, Van Den Noortgate, & Van Damme, 2014), or where girls show higher achievement than boys (Voyer & Voyer, 2014).

Research in sex differences have been seen between age and achievement levels within different achievement groups (e.g. male overrepresentation in high achieving and low achieving groups (Penner & Paret, 2008). Moreover, high parental education has been reported to advantage young male students entering kindergarten, in not only middle and upper class families, but noted also in Asian families. All of which exhibit the largest male advantage at the top of the distribution. In contrast and of note, Penner and Paret (2008), report that Latino kindergarten girls have an advantage over boys at the top of the distribution. However, the effect sizes of such studies are often challenged (Hyde, 2005; Hyde, Fennema, & Lamon, 1990; Penner & Paret, 2008; Shen et al., 2016; Spelke, 2005). For example, Else-Quest, Hyde, and Linn's (2010) meta-analysis of TIMSS and PISA data showed very low effect sizes, but, they also suggested that there was a higher variability between the countries in those effect sizes in specific domains of mathematics. An example of the influence of educational and cultural context on young students' arithmetical strategies, is the study of Shen at al. (2016). They found that sex differences in the use of strategies for mixed- and double-digit addition and subtraction was present for American and Russian first graders but not for Taiwanese, indicating the importance of educational and cultural context.

Sex differences in mathematics ability related to number and arithmetic may in theory also be linked to underlying differences in cognitive factors that are known to correlate with mathematics ability, e.g. mental rotation skill (e.g. Laski et al., 2013; Moè, 2018) and working memory (e.g. Geary, Hoard, & Nugent, 2012). Sex differences in for example mental rotation skill has been shown for children as young as 4½ years (Levine, Huttenlocher, Taylor, & Langrock, 1999) as well as college students (Casey, Nuttall, Pezaris, & Benbow, 1995). Sex differences in cognitive factors does on the other hand not necessarily link to sex differences in calculation performance and number competences. For example, Moè (2018) found no sex differences in numbers and arithmetic ability of eight and ten year old children despite that boys scored higher than girls in mental rotation (2-D and 3-D), a factor that in turn correlated with mathematics performance at ten years' of age.

In recent years, attention has focussed on the importance of identifying individual differences children display in the early learning of number, in order to inform appropriate intervention planning and teaching for children (Dowker, 2005). Research also points to identifying what aspect of mathematics children struggle with. For example, Hornburg, Rieber, and McNeil (2017) analysed data from 14 studies (including 960 second and third graders) on boys' and girls' understanding of mathematical equivalence. They reported that in arithmetic, girls are more likely to construct narrow knowledge, i.e. memorise taught procedures and rely on that knowledge when encountering novel problems. Consequently, this causes differences between them and boys who demonstrate broader knowledge of number.

In light of the above, we appreciate that there are no simple distinctions between sex differences, or when students might present differences in different mathematical components. Thus, explanations for sex differences can be found in individual differences as well as sociocultural factors. Indeed, there is already a significant amount of research in these areas, which is receiving increasing recognition (e.g. Hirnstein, Andrews & Hausmann, 2014). The study presented here sought to investigate how young boys and girls in years 1 to 4 develop their strategies in single-digit addition in a Danish context.

Denmark

The study is of interest to a wider audience for a number of reasons. Firstly, little is published on Danish elementary school mathematics, its content and pedagogy. Secondly, according to Selter (2001, p. 147), standard methods of calculation are given a marginal position in the arithmetic curriculum focus. Lastly, and not insignificant, the Danish curriculum, and its assessment framework, are at the root of the international PISA testing instrument.

The teaching of mathematics at compulsory school in Denmark (age 6-15), has undergone much self-scrutiny in recent years, and teacher education has undergone some major transformations since 2001. A decade later, Niss and Højgaard (2011) recommended that the Ministry of Education "ensure that teacher training at all educational levels under the jurisdiction of the Ministry be designed and structured so that future teachers are equipped with the mathematical, didactic and pedagogical competencies presented in this report." (Section 1.5, p. 194). These competencies are now part of the official Danish curriculum, but Danish teachers enjoy great autonomy in the classroom (Skott, 2004) and are often reported to uncritically trust textbook content to deliver curriculum outcomes (Bremholm & Skott, 2019). The study here seeks to illuminate what arithmetic strategies Danish children are working with in single-digit addition. The operation that is reported to be the foundations of primary mathematics instruction (Carpenter & Moser, 1984).

Strategies in Single-Digit Mental Addition

Carpenter and Moser (1984) identified three basic levels of addition strategies based on 1) direct modelling strategy with fingers or physical objects where counting sequences are used by children to count all concrete objects; 2) Counting strategies, where children count on from the first or the larger number, used in developing mental counting and from counting all starting with the first addend to counting all starting with the larger addend and finally to counting on from the larger addend, in other words, counting-all is replaced by counting-on strategies. These strategies have, and since been reported to be used for some time before using strategies based on recalling number facts (Sarama & Clements, 2009). Finally, 3) Number fact strategies which involve either a) direct retrieval, or b) derived fact strategies. Direct retrieval (fact retrieval or just retrieval) is the direct recalling of the solution to an addition question from memory (Carpenter & Moser, 1984). These known facts are an important prerequisite of derived fact strategies, the focus of many recent studies (e.g. Dowker, 2014; Laski, Ermakova, & Vasilyeva, 2014). Derived fact strategies, where number facts are derived from a recalled number fact (Carpenter & Moser, 1984, p. 181), are also known as decomposition or partitioning strategies. These strategies build on known facts and involve transforming the addition question into more simple questions. Derived fact strategies fall into two distinct groups based on known facts or, base-10 knowledge (Dowker, 2014). An example of the former is to solve the unknown question of 7+8 by transforming into the known sum of 7+7 and then add 1. An example of the use of base-10 knowledge is solving the same question 7+8 by partitioning 7 into 2 and 5, adding 8 and 2 to get 10, and then add the remaining 5 to get the final sum 15.

Derived fact strategies has been associated with conceptual understanding (Canobi, Reeve, & Pattison, 1998) and number knowledge, e.g. base-10 knowledge (Laski et al., 2014) and are thus considered to be more advanced strategies than counting.

Strategy Use in Mental Addition and Relationship with Mathematics Achievement

The use of mental strategies in arithmetic has been shown to influence development of mathematics competence (e.g. Carr et al., 2008; Dowker, 2014; Fennema et al., 1998; Ostad, 1997) and different strategy use is thus a valid predictor of both later mathematical achievements and difficulties (Carr & Alexeev, 2011; Gersten, Jordan, & Flojo, 2005; Ostad, 1997; Price, Mazzocco, & Ansari, 2013). Yet, in the Nordic context Ostad's (1997) longitudinal study has shown that children without difficulties use "a richness of strategies" (p. 355) comprising of both direct retrieval and derived fact strategies, which increases as the numbers used confidently increases as they progress through school. Thus, children using retrieval strategies at an early age (6-7 years) are less likely to develop difficulties in mathematics. However, children with difficulties were found to use fewer and inadequate strategies, primarily backup strategies (counting), and fix their choice of strategies early in school with little subsequent progression (ibid.).

The debate of individual differences is related to strategy use and achievement increases, as more recently Bailey, Littlefield, and Geary (2012) reported that the preference for and skill at using a specific strategy related to a feedback loop, (a process for checking for and affirming understanding, thus can be negative as well as positive outcomes), predicts the later skill. In this case, early preference of strategies used for addition successfully which in turn predicts later preference, and so forth. Students preferring counting strategies will get less practice in using derived fact strategies and thus working flexibly with numbers. Consequently, they will have fewer opportunities to develop the more complex understanding necessary for advanced arithmetic (Gersten et al., 2005). However, Bailey et al. (2012) also found a variation in such influences between girls and boys across grades 1-6 in North America, indicating the importance of identifying such preferences in other contexts and cultural settings.

Sex Differences in Use and Development of Strategies in Addition

When considering strategies in addition specifically, research evidence continues to be ambiguous. Sex differences in use and development of strategies in single-digit addition has been addressed in several studies from different western countries. In a study on 6 and 7 year olds Dowker (2009) found no sex differences in addition performance or use of derived fact strategies. However, other studies report on consisting sex differences (e.g. Bailey et al., 2012; Carr & Alexeev, 2011; Carr & Davis, 2001; Imbo & Vandierendonck, 2007; Paul & Reeve, 2016; Shen et al., 2016).

Bailey et al. (2012) found that boys showed a bias towards the use of retrieval and girls showed a bias for the use of counting procedures from years 1–6, a result that was consistent with previous findings by Carr and Davis (2001), Carr and Jessup (1997), and Geary, Bow-Thomas, Liu, and Siegler, (1996). The developmental pattern for skilled retrieval is more complex. For example, in first grade, boys are reported to have a higher preference for retrieval than girls, but the difference in accuracy was insignificant at this age. But by the second grade, although girls were still shown to retrieve less than boys they were more accurate than boys when they did use it (Bailey et al., 2012). This reflects Siegler's (1988) research findings, resonating the identification of Carr and Alexeev's (2011) 'perfectionists': girls were less likely to adopt a retrieval strategy than boys, as they want to be sure of accuracy, even though this strategy is more time-consuming. Contrastingly, this early preference by boys, according to Bailey et al. (2012), was found to increase steadily from first to sixth grades, and at this point out-performed girls. It is important to note that Bailey et al.'s (2012) results indicate, that ability influences early skilled retrieval, but both practice and skill influence each other in a feedback loop later in development. Thus, there appears to be a change within grades 1-6.

Boys' early advantage for accurate retrieval was not due to IQ or 'central executive' according to Geary et al. (2012), and research suggests that boys generally take more risks than girls, especially in situations where their performance will be socially evaluated (Geary, 2010). Boys like to express themselves more freely in open settings that might encourage a competitive environment. Peterson and Fennema (1985) reported that boys tended to enjoy getting the answer before anyone else. However, Bailey et al. (2012) reported that boys had shorter reaction times than girls in grade one, but that girls begin to catch up with them, whereas boys' reaction times did not increase in the same way. They suggest this is because the competitiveness is about being seen to be first to answer, rather than accuracy. Which confirms previous studies that suggested that this competitive perspective may harm boys' addition accuracy in the early grades, but this early preference may put them at an advantage later on (Royer, Tronsky, Chan, Jackson, & Marchant, 1999).

The Current Study

This study parameterizes the use and development of strategies in single-digit addition by gender over time (age). We focus on the first years of school in a four year longitudinal perspective. Furthermore, we investigate the different strategy categories (counting, direct retrieval and derived facts strategies) separately.

Our research aim is to quantify the variation and development in students' use of different strategies for single-digit addition during the first four years (years one to four) of Danish school, including any possible sex differences. We apply a multilevel model approach to ensure the generalisability of the findings and test the model on an independent dataset.

Methods

Participants

Data consist of information from assessment interviews on 230 Danish primary students' unprompted strategy use for single-digit addition from two independent studies: A and B, where study B was used to validate the results from grade 1, in study A. Informed written parental consent was obtained for all students. Students were informed orally about the study and that participation was voluntary.

Study A

Study A consisted of 260 assessment interviews conducted on 147 students (77 girls) from eight classes from a single school from year one to four (age 7 to 11). Three classes were tested thrice in subsequent years; two classes twice; and three classes once (Table 1).

INSERT TABLE 1 ABOUT HERE

Study B

Study B consisted of 155 assessment interviews from 83 students (46 girls; age 7) from six classes from three other schools. Students were interviewed in the first and second half of year one (Table 2).

INSERT TABLE 2 ABOUT HERE

Assessment

Interviews on unprompted strategy use.

Strategy use was monitored in one-to-one assessment interviews by presenting the student with flashcards with the 36 addition tasks with numbers 2-9, including doubles. The addition tasks were presented with sums less than 10 first and then with increasing sum and difficulty. This was to avoid less confident students, as reported in Dowker 's (2003) study, giving up or only using counting because they were presented with difficult questions (e.g. 8+7) at the beginning of the interview. The students were not provided with any manipulatives or paper and pencil. The interviews lasted 10-30 minutes and were performed in a quiet room at the school. The solving time for the individual addition task was not measured.

In light of the children being so young, the interview script was cross checked for clarity and appropriateness with researchers and teachers of the children. The researcher stated to each student: "I will show you some single-digit addition tasks. First, I would like you to find the answer to the task, and then we talk about how you found the answer. There are many ways to find the answer to an addition task. Sometimes you might know the answer or count or perhaps you use other tasks to find the answer. I am interested in knowing how you find the answer." Then the student was presented for a flashcard, e.g. 4+5, and the interviewer asked: "what is the answer to four plus five?". If the student did not give an explanation following the answer, the interviewer asked: "how did you find the answer?" and if further prompting was needed: "did you count or did you just know the answer, or did you use some other tasks you know to find the answer?".

Categorisation of strategy use.

If the student gave up or miscalculated an answer, this was categorised as Error. Strategy use for incorrect answers were not categorised. Correct answers were categorised based on the student's self-report of strategy use and observations by the interviewer (e.g. visible signs of finger-counting or lip movements). The validity of young students self-reported strategy use has been established in other studies (e.g. Canobi et al., 1998) based on comparison with solution time. Strategy use of correct answers were categorised in Counting all (counting both addends and then all together), Counting on (counting on from one of the addends), Direct retrieval (reported just knowing the answer), Derived fact using addition (decomposing addends and calculating answers using automatized sums with subsequent use of addition e.g. 4 + 5 = 4 + 4 + 1), and Derived fact using subtraction (decomposing addends and calculating answers using automatized sums with subsequent use of addition e.g. 4 + 5 = 5 + 5 - 1). We did not distinguish between counting on fingers, verbal counting or self-report of mental counting. Initial analyses showed that the categories counting all and both derived fact categories were used relatively infrequently by the students. It was therefore difficult to establish robust statistical models for these rare categories. Therefore, we pooled the counting and derived fact strategies respectively to 1) provide more reliable results and 2) to reduce complexity of the analysis. The pooled categories of strategy use for correct answers used in the analysis was thus Counting, Direct retrieval and Derived fact.

Data Analysis

Development of strategy use (Study A).

Variation in the frequency by which a given solving strategy was used by students in study A (binary response variable: e.g. whether direct retrieval was used as opposed to all other strategies) to solve a specific addition task (e.g. '5+6') as a function of school year (number of years the student has attended school), sex and the interaction was analysed by means of Generalised Linear Mixed Models (GLMM: GLIMMIX procedure in SAS 9.4, SAS Inst. (Littell, Stroup, Milliken, Wolfinger, & Schabenberger, 2006)) with a logit link function and binomial error distribution.

We used the answers of the individual addition task (36 per assessment interview) as observation unit. The entire data set thus consisted of a total 9360 addition tasks based on 260 assessment interviews, nested within 147 students from eight classes. To account for the expected dependency of observations within this hierarchical data structure, we stated the following random effects in the model: (1) assessment interview nested within student ID (accounting for dependency of strategy use among task cases within the interview), (2) student ID nested within class (accounting for random variation between students within a class), and (3) class ID (accounting for random variation between classes). Finally, the identities of the specific addition task (e.g. '2+2', '2+3' etc.: 36 different task IDs in total) were also stated as random effects to account for any variation in problem solving strategy related to the specific task.

As fixed effects, we included sex (categorical variable) and school age (covariate representing the number of years the student had received school education: for example, if school year runs from 8 August to 26 June, a student tested in year one on 23 October had calculated school age of 1.24 and when tested again on 29 March in the same year a school age of 1.72) and the interaction term between these two factors.

The statistical significance of the fixed effects was based on p-values derived from t- and F-statistics of the models (the two test statistics rendered the same results, except a single case declared as such where the model had become over-parameterised because of inclusion of a superfluous interaction term, resulting in an unreliable t-statistics because of a poorly estimated SE). We used $p < 0.05$ as significance level for whether we would draw inferences from a result. Degrees of freedoms in the GLMM were calculated with the Satterthwaite approximation (Littell et al. 2006). Model predictions (predicted universal probability of using a given strategy for a given sex at a given school age with 95% confidence errors) were obtained as least square means.

Effect sizes were expressed as the odds ratios of the difference in strategy use by mid year one (school age = 1.5), derived as exp(b), where b is the coefficient of the gender effect on logit scale by school age = 1.5 (extracted from SAS as a least square means estimate specific to school age = 1.5).

Validation of results (Study B).

The same model structure was used as in study A. Model predictions (least square means estimates) for the two studies were compared for mid-year one (least square means predictions for year =1.5). The difference between estimates obtained in study A and B were tested on the t-statistics of the difference in parameter values (t(dfA+dfB) = [bA-bB]/[SEb(A)2 + SEb(B)2]0.5).

We also split the analysis in study B on the six different classes in order to visualize predictions if estimated separately for each class. We furthermore tested explicitly for between-class variation in sex specific strategy use by entering class ID as a fixed effect. Likewise, we tested for variation between classes in sex specific strategy use by entering a sex-by-class interaction term in the model's fixed effect statement.

Post-hoc analysis: counting divided on counting all and counting on.

After the overall result from study A had been replicated in study B, on the combined (and statistically more powerful) dataset, we quantified the sex differences in counting all and counting on, i.e. the two sub-categories comprising the overall category counting (same model structure as above).

INSERT FIGURE 1 ABOUT HERE

Results

Study A

The frequency by which all four strategy categories were used varied as function of school year, as 'error' and 'counting' were used less and the more advanced strategies 'derived fact' and 'direct retrieval' were used more with increasing school year (Figure 1, Table 3). Girls used "counting" significantly more (F:M odds ratio [95%CI] by mid year one = 2.9 [1.9-4.3]) and direct retrieval (0.57 or 1:1.8 [0.45-0.73]) and derived fact strategies (0.30 or 1: 3.3 [0.16-0.58]) significantly less often than boys. For counting, the effect size of sex corresponded to 2½ years development (Table 3); or in other words: the average use of counting by boys in the start of year one equalled the use of girls by mid year three (Figure 1). Same patterns applied for direct retrieval and derived fact (Table 3, Figure 1). On arithmetic scale, all sex differences diminished with school age. Hence, by year four, girls' strategy use in single-digit addition nearly equalled that of the boys.

The category error showed no sex differences (Table 3). For counting and derived fact strategies, sex differences were statistically significant as main effects only (Table 3).

This means that the relative difference between boys and girls was constant irrespective of school year on logit scale, but diminished with increasing school age when back-transformed to arithmetic scale (Figure 1) due to a reduced space within which the relative difference between sexes could be expressed in terms of percentage points when the proportional use of these strategies converged towards 0 and 1, respectively (boundary effect).

For direct retrieval, a statistically significant sex-by-school year interaction (Table 3) suggested that not only absolute (arithmetic scale: Figure 1) but also the relative sex difference (logit scale) in the inclination by which boys and girls used this strategy decreased with increasing school year.

In addition to the afore described mean trends in strategy use as function of school year and sex, considerable variation in strategy use could be attributed to the individual test situation (nested within student) and student identity (nested within class). This variation is apparent as the scatter around the predicted functions of the mean (Figure 1) as well as covariance parameter estimates of test and student identity as random effects (Table 3). In contrast, modest covariance parameter values of class identity relative to those for test and student identities indicated that most individual variation in strategy use occurred on test and student level rather than at class level. Lack of statistical significance of the class identity covariance parameters (Table 3) further suggested that the influence of class environment on strategy use was minor if present at all. The magnitude and statistical significance of the covariance parameter values of addition task identity also suggested that considerable variation existed in the probability by which the strategy in question was used to solve the individual addition tasks.

INSERT TABLE 3 ABOUT HERE

INSERT FIGURE 2 ABOUT HERE

Study B

Study B produced similar predictions for the difference between boys and girls in strategy use as well as strategy use specific to sex (Figure 2, Table 4 and 5) to study A: girls used counting significantly more often and derived fact and direct retrieval significantly less often than boys (Figure 2, Table 4). With exception of ‘error’ that occurred more often in study A than study B (Figure 2, Table 4), the proportional use of the different strategies was also equal in the two studies (Figure 2, Table 4).

The sex specific differences in the proportional use of ‘counting’, ‘derived fact’ and ‘direct retrieval’ estimated were also apparent if estimated separately for the six different classes in study B (Figure 2), and the magnitude of the sex effects did not vary significantly between the classes (Table 5).

Post Hoc Analysis of Counting All and Counting On

On the combined data from study A and B, by mid-year one, the sex difference (F:M odds ratio) in counting all and counting on was 3.2 (95%CI: 1.5-6.8, $t_{213.1}=3.05$, $P=0.0027$) and 2.1 (95%CI: 1.5-3.2, $t_{188.6}=3.27$, $P=0.0014$), respectively.

INSERT TABLE 4 ABOUT HERE

INSERT TABLE 5 ABOUT HERE

Discussion

This study demonstrates substantial mean differences in development patterns in the use and development of strategies for single-digit addition of girls and boys in the first years of Danish primary school, with girls using counting strategies 3 times more and derived fact strategies 3 times less often than boys by mid year one. This difference that was equal to girls being 2-3 years behind boys in development of strategy use was highly statistically significant and consistent across schools and classes, suggesting that it was general in the Danish primary school context. The girls' predisposition to use counting strategies more often than did boys applied for the least advanced counting all strategy as well as for the slightly more advanced counting on strategy.

Our results appear to be comparable to those of Bailey et al. (2012), where sex differences were found in the use of direct retrieval equal to approximately two school years in a longitudinal study (year 1-6) in USA. Our results also resembles that of Carr and Jessup (1997) who found that boys in year one in USA used retrieval more often than did girls. However, unlike us they found an increasing sex difference from October (where boys and girls used retrieval equally) to April (where boys were using retrieval almost twice as often as the girls). In Flemish Belgium, Imbo and Vandierendonck (2007), found that sex differences explained 4-5% of the variation in year four to six students' frequency of use of retrieval, also after controlling for age, arithmetic skill, working memory load and processing speed. However, the estimated difference in sex-specific strategy use (e.g. odds ratios) was not reported. Finally, Geary et al. (2012) in USA (kindergarteners) and Paul and Reeve (2016) (Australia, kindergarten to year 3) reported that boys differed significantly from girls in arithmetic strategies in the same direction as we observed, but gave no details on the size of these differences.

Other studies have reported differences in strategy use to be related to cognitive factors (Foley, Vasilyeva, & Laski, 2016; Laski et al., 2013), and indeed cultural factors such as different instructional practices, demonstrated to be apparent in several cross-country studies (e.g. Hyde & Mertz, 2009; Penner & Paret, 2008; Shen et al., 2016). Relating individual variation (including sex differences) in strategy use to other parameters such as general mathematical achievement or cognitive development is beyond the scope of this paper (a thorough analysis of data from study B that also include longitudinal data on mathematical achievements by the end of fourth grade will be published later). Nevertheless, we can reveal that year one students' general mathematical achievement was similar for boys and girls, but correlated negatively with proportional use of error and counting all but positively with proportional use of derived facts, suggesting some association between strategy use and general mathematical achievement in year one.

Implications

The marked sex-difference in development of strategy may have implications for research as well as teaching practice. Regarding research of variation and development of strategy use, our results suggest that sex may be a significant component of variation that should not only be tested for (or neutralised by balanced study designs), but incorporated in the statistical analyses instead of appearing as apparent statistical background noise. Sex differences in strategy use (or any other behaviour or skill) may also be of interest in its own right in order to identify differences between boys and girls in a given age or cultural context.

For teaching practice, knowledge of the existence of substantial variation in different students' development of strategy use, much of which connecting to the individual's sex may also be of importance. For instance, students who overly rely on counting in single-digit problem solving (of which girls are overrepresented) may need explicit teaching and encouragement to memorise number facts and patterns in number bonds. Thus, if instruction in the early years of school focus on counting strategies and procedural knowledge in the teaching of arithmetic this would disadvantage the girls as they are more prone to construct narrow knowledge and rely on the taught algorithms (Hornburg et al., 2017). Furthermore, Laski et al. (2013) points out that girls' persistent counting provide them with fewer opportunities to practice derived fact strategies and retrieval and that might lead to girls' poorer accuracy in these strategies.

Considering the findings of this study, it would be relevant to investigate the effects of different aspects of instructional practices on sex specific development of strategies in arithmetic in the early years. For example, boys are known to be more willing to take risks in a competitive environment and it could be that teaching promoting a competitive learning environment may be more beneficial for boys, whereas girls might experience a reinforcement of the feedback loop (Bailey et al., 2012).

Limitations

We are aware of difficulties in researching with young children through one-to-one interviews (Punch, 2002), and that gender effects could be mediated by personality characteristics (Hornburg et al., 2017) such as impulsivity and inhibition which may affect strategies the student used in the interviews. Furthermore, using only a choice method, where the students are free to choose the strategy to solve each item, cannot necessarily provide information on the students' actual repertoire. There is a risk that the student's choice of strategy in the test situation is biased for different reasons as indicated above. To avoid such bias, a choice/no-choice method can be applied as discussed by Torbeyns, Verschaffel, and Ghesquière (2005). However, we believe this bias to be relatively small in this study, because of the students in general were used to this kind of test or interview situations in school, it took place in the students' familiar school settings and the students in most cases were familiar with the researcher.

The students' attempted category was not registered. As the error category was indirectly incorporated in the analysis of other strategy categories as each category was analysed against all other categories (including error) this could influence the result. However, as the proportional use of the error category was relatively small (less than 10% in year one) and decline rapidly with time the overall result would most likely not be influenced. Furthermore, Laski et al (2013) found that girls were least accurate when using counting strategies compared to direct retrieval and decomposition. If this is the case, then we would probably rather have underestimated the girls use of counting in year one.

The effect of sex in this study levelled out by year four in single-digit addition, as students move onto more challenging tasks by years two and three. This is not necessarily an indication that sex specific differences no longer are present. It is equally likely that it is caused by a ceiling effect similar to the ceiling effects reported by Shen et al. (2016). Thus, single-digit addition is probably not a good diagnostic tool for the average students beyond year three.

Conclusion and Perspectives

We have shown the existence of considerable specific sex differences in one of the foundational aspects of teaching number in the early years arithmetic, in the very first years of formal Danish schooling. Considering the lack of research on the learning and development of number and arithmetic in the early years of school in Denmark, as well as early sex differences, we find that it is important that these findings are reported.

The reasons behind these sex differences, equalling 2-3 year's, remain to be explicated. No matter the reasons behind these sex differences, their sheer magnitude in this as well as in at least one previous study (Bailey et al., 2012) may suggest that sex differences and their underlying causes deserves more attention than appear to be standard in most studies of arithmetic in early years of school.

Although a systematic review of the extent to which the presence and magnitude of sex differences in arithmetic research is beyond the scope of this paper, it seems that sex differences are sometimes not quantified in research papers, which means that sex differences in arithmetic patterns possibly could be underreported. The occurrence of differences in arithmetic patterns between boys and girls may be considered politically sensitive and vulnerable to partisan misinterpretations, but this does not change the fact that they sometimes exist and should be investigated, quantified and explained as any other predictor variable in education research. Knowledge on how individual differences and sociocultural factors interact to produce the resulting sex differences might thus prove useful to inform targeted teaching and educational practice that 'decrease' rather than unintentionally enhance sex differences no matter what have caused them.

Acknowledgements

We would like to thank The National PhD Council in Educational Research, Denmark for supporting the collection of this data.

References

Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The codevelopment of skill at and preference for use of retrieval-based processes for solving addition problems: Individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology, 113*(1), 78–92.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral & Brain Sciences, 11*, 169–232.

Bremholm, J., & Skott, C. K. (2019). Teacher planning in a learning outcome perspective: A multiple case study of mathematics and L1 Danish teachers. *Acta Didactica Norge, 13*(1), 1–22.

Browne, K. R. (2002). *Biology at work: Rethinking sexual equality*. New Brunswick, NJ: Rutgers University Press.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology, 34*, 882–891.

Carpenter, T. & Moser, J. (1984). The Acquisition of Addition and Subtraction Concepts in Grades One through Three. *Journal for Research in Mathematics Education*, *15*(3), 179–202.

Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology*, *103*(3), 617–631. doi: 2048/10.1037/a0023864

Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary Educational Psychology, 26*(3), 330–347. doi:10.1006/ceps.2000.1059

Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, *89*(2), 318–328.

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, *18*(1), 61–75.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, *31*(4), 697–705.

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior & Development, 23,* 113–118.

Dowker, A. (2003). Young children’s estimates for addition: The zone of partial knowledge and understanding. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructive adaptive expertise*. (pp. 243–266). Mahwah, NJ: Erlbaum.

Dowker, A. (2005). Early Intervention for students with mathematical difficulties. *Journal of Learning Disabilities*, *38*(4) 324–332.

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development*, *24*(4), 401–410. doi:10.1016/j.cogdev.2009.09.005

Dowker, A. (2014). Young children's use of derived fact strategies for addition and subtraction. *Frontiers in Human Neuroscience, 8,* 1–9. doi:10.3389/fnhum.2013.00924

Else-Quest, N., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103–127. http://dx.doi.org/10.1037/a0018053 . Retrieved 20.10.17

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher, 27*(5), 6–11.

Foley, A. E., Vasilyeva, M., & Laski, E. V. (2016). Children's use of decomposition strategies mediates the visuospatial memory and arithmetic accuracy relation. *British Journal of Developmental Psychology*, *35*(2), 303–309. doi:10.1111/bjdp.12166

Geary, D. C. (2010). *Male, female: The evolution of human sex differences.* Washington, DC: American Psychological Association.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development*, *67*, 2022–2044.

Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Child Psychology*, *113*(1), 49–65. doi:10.1016/j.jecp.2012.03.003

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*(4), 293–304.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1–51.

Hirnstein, M., Andrews, L. C., & Hausmann, M. (2014). Gender-stereotyping and cognitive sex differences in mixed-and same-sex groups. *Archives of sexual behavior, 43*(8), 1663-1673.

Hornburg, C. B., Rieber, M. L., & McNeil, N. M. (2017). An integrative data analysis of gender differences in children's understanding of mathematical equivalence. *Journal of Experimental Child Psychology, 163*, 140–150.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581–592. doi:10.1037/0003-066X.60.6.581

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139–155. doi: 10.1037/0033-2909.107.2.139

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences, 106*(22), 8801–8807. doi:10.1073/pnas.0901265106

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology, 96*(4), 284–309. doi:10.1016/j.jecp.2006.09.001

Laski, E. V., Casey, B. M., Yu, Q., Dulaney, A., Heyman, M., & Dearing, E. (2013). Spatial skills as a predictor of first grade girls' use of higher level arithmetic strategies. *Learning & Individual Differences, 23*, 123–130. doi:10.1016/j.lindif.2012.08.001

Laski, E. V., Ermakova, A., & Vasilyeva, M. (2014). Early use of decomposition for addition and its relation to base-10 knowledge. *Journal of Applied Developmental Psychology, 35*(5), 444–454.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35*(4), 940–949.

Littell, R. C., Stroup, W. W., Milliken, G. A., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models.* SAS institute.

Maccoby, E. E., & Jacklin, C. N. (1974). *Psychology of sex differences.* Stanford, CA: Stanford University Press.

Moè, A. (2018). Mental rotation and mathematics: Gender-stereotyped beliefs and relationships in primary school children. *Learning and Individual Differences, 61*, 172–180.

Niss, M., & Højgaard, T. (2011). *Competencies and mathematical learning: Ideas and inspiration for the development of teaching and learning in Denmark* (IMFUFA tekst). Roskilde: Roskilde University.

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology, 67*(3), 345–357.

Paul, J. M., & Reeve, R. A. (2016). Relationship between single digit addition strategies and working memory reflects general reasoning sophistication. *Learning and Instruction, 42*, 113–122.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37*(1), 239–253. doi:http://dx.doi.org/10.1016/j.ssresearch.2007.06.012

Peterson, P. L., & Fennema, E. (1985). Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics. *American Educational Research Journal*, *22*(3), 309–335.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York: Viking.

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, *84*(1), 152–174.

Price, G. R., Mazzocco, M. M., & Ansari, D. (2013). Why mental arithmetic counts: Brain activation during single digit arithmetic predicts high school math scores. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(1), 156–163. doi:10.1523/JNEUROSCI.2936-12.2013

Punch, S. (2002) Interviewing strategies with young people: the ‘secret box’, stimulus material and task based activities. *Children and Society 16*(1): 45–56.

Royer, J. M., Tronsky, L. N., Chan, Y., Jackson, S. J., & Marchant, H. III, (1999). Math-fact retrieval as the cognitive mechanism underlying gender differences in math test performance. *Contemporary Educational Psychology*, *24*(3), 181–266.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York: Routledge.

Selter, C. (2001). Addition and subtraction of three-digit numbers: German elementary children's success, methods and strategies. *Educational Studies in Mathematics*, *47*(2), 145–173.

Shen, C., Vasilyeva, M., & Laski, E. V. (2016). Here, but not there: Cross-national variability of gender effects in arithmetic. *Journal of Experimental Child Psychology*, *146*, 50–65. doi:10.1016/j.jecp.2016.01.016

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, *59*, 833–851.

Skott, J. (2004). The forced autonomy of mathematics teachers. *Educational Studies in Mathematics*, *55*(1–3), 227–257.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *American Psychologist*, *60*(9), 950–958.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2005). Simple addition strategies in a first-grade class with multiple strategy instruction. *Cognition and Instruction*, *23*(1), 1–21. Retrieved 21.4.18 from http://dx.doi.org/10.1207/s1532690xci2301_1

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204.

Table 1. Study A: Number of students assessed per class (Girls/Boys) at each year and mean dates for assessment interviews.

	Number of assessment interviews conducted per year										Students in total	
	Year 1		Year 2		Year 3		Year 4		Total			
Class	G/B	mean date	G/B	mean date	G/B	mean date	G/B	mean date	G/B	all	G/B	all
A1					11/6	2-Nov-13	13/6	30-May-15	24/12	36	13/6	19
A2			9/4	31-Aug-12	12/4	28-Oct-13	14/5	22-Apr-15	35/13	48	14/5	19
A3	0/1	1-Nov-12	11/6	29-Sep-13	14/8	21-Apr-15			25/15	40	14/8	22
A4	13/7	21-Feb-14	8/7	10-Apr-15	8/6	2-Jun-16			29/20	49	8/7	15
A5			7/10	1-Apr-15					7/10	17	7/10	17
A6	6/10	26-Apr-15	6/9	6-Jun-16					12/19	31	6/10	16
A7	6/9	21-Apr-15							6/9	15	6/9	15
A8	9/15	6-Oct-15							9/15	24	9/15	24
Sum	34/42		41/36		45/24		27/11		147/113	260	77/70	147

Table 2. Study B: Number and mean dates of assessment interviews conducted on students (Girls/Boys).

	Number and date of assessment interviews:						Students in total	
	round A		round B		Total			
Class	G/B	mean date	G/B	mean date	G/B	all	G/B	all
B1	9/6	5-Nov-15	8/5	14-Apr-16	17/11	28	9/6	15
B2	5/5	3-Nov-15	7/5	20-Apr-16	12/10	22	6/5	11
B3	7/5	2-Nov-15	7/5	14-Apr-16	14/10	24	7/5	12
B4	6/5	10-Nov-15	7/6	15-Mar-16	13/11	24	8/6	14
B5	7/6	5-Nov-15	6/6	4-Apr-16	13/12	25	7/6	13
B6	9/9	26-Oct-15	8/6	29-Mar-16	17/15	32	9/9	18
Total:	43/36	2-Nov-15	43/33	5-Apr-16	86/69	155	46/37	83

Table 3. Equations for GLMMs (logit link) of proportional use of the four different addition strategies in study A as interactive functions of year (range from mid year 1 to mid year 4) and sex, when accounting for random variation attributable to the individual, student specific assessment situation, student ID, class and addition task. Statistical significances (unless otherwise stated F- and t-statistics gave similar results): ns: $p \geq 0.1$, °: $p < 0.1$,*: $p < 0.05$, **: $p < 0.01$, ***: $p <0.001$, ****: $p < 0.0001$.

	Error					Counting				
Fixed effects	B	SE(B)	DF	t/z	Sign	B	SE(B)	DF	t/z	Sign
Intercept	-0.86	0.63	80	-1.35	ns	-0.39	0.32	57	-1.21	ns
Sex (S)	-0.40	0.71	199	-0.55	ns	1.31	0.33	254	3.99	****
Year (YR)	-1.63	0.26	178	-6.28	****	-0.36	0.09	207	-4.06	****
YR*S	0.45	0.30	235	1.50	ns	-0.17	0.10	227	-1.58	ns
Random effects (covariance parameters):										
Test(student)	0.62	0.27		2.27	*	0.38	0.07		5.41	****
student(class)	1.75	0.43		4.09	****	0.64	0.13		5.00	****
class	0.18	0.22		0.82	ns	0.21	0.13		1.53	ns
Addition task	1.39	0.42		3.30	***	0.45	0.11		4.03	****

	Direct retrieval					Derived fact				
Fixed effects	B	SE(B)	DF	t/z	Sign	B	SE(B)	DF	t/z	Sign
Intercept	-1.05	0.26	65	-3.99	***	-2.17	0.58	79	-3.76	***
Sex (S)	-0.74	0.19	254	-3.88	****	-1.64	0.53	210	-3.07	**
Year (YR)	0.24	0.05	197	5.29	****	0.26	0.14	139	1.82	°§
YR*S	0.12	0.06	197	1.99	*	0.30	0.17	183	1.74	°
Random effects (covariance parameters):										
Test(student)	0.06	0.02		2.56	*	1.43	0.24		5.96	****
student(class)	0.23	0.04		5.41	****	0.94	0.29		3.23	**
class	not estimable					0.30	0.24		1.27	ns
Addition task	1.82	0.43		4.20	****	4.45	1.14		3.89	***

§ in this model, the t-statistics for year yielded a p-value of 0.08, whereas F-statistics resulted in $p < 0.0001$. If the interaction term was removed from the model, the t- as well as the F-statistics suggested a strongly significant effect ($p < 0.0001$) of year.

Table 4. Comparison and tests for differences between study A and B in predicted relative differences of strategy use by girls and boys (logit scale) in mid year one as well as the mean probabilities (also on logit scale) by which girls and boys were using the different strategies to solve simple addition tasks. Statistical significances: ns: $p \geq 0.1$, °: $p < 0.1$,*: $p < 0.05$, **: $p < 0.01$, ***: $p <0.001$, ****: $p < 0.0001$.

		Sex difference (G-B)					Mean: Girls					Mean: Boys				
Strategy	Study	B	se	df	t	sign	b	se	df	t	sign	b	se	df	t	sign
error	A	0.14	0.43	189.5	0.34	ns	-3.03	0.36	34.7			-3.31	0.38	31.3		
	B	0.49	0.36	84.9	1.37	ns	-4.01	0.39	15.5			-4.50	0.42	20.8		
	Diff A-B	-0.35	0.56	274.4	-0.62	ns	0.99	0.53	50.2	1.85	°	1.19	0.56	52.1	2.11	*
counting	A	1.11	0.23	241.4	4.83	****	0.12	0.24	23.7			-0.94	0.25	24.8		
	B	1.37	0.31	74.5	4.50	****	0.61	0.31	76.6			-0.76	0.32	85.0		
	Diff A-B	-0.27	0.38	315.9	-0.70	ns	-0.49	0.39	100.3	-1.25	ns	-0.18	0.41	109.8	-0.43	ns
direct retrieval	A	-0.60	0.13	263.9	-4.43	****	-1.25	0.24	47.1			-0.68	0.24	46.3		
	B	-0.90	0.26	73.7	-3.45	****	-1.16	0.25	45.7			-0.66	0.25	47.8		
	Diff A-B	0.30	0.29	337.6	1.03	ns	-0.09	0.35	92.8	-0.25	ns	-0.03	0.35	94.1	-0.08	ns
derived fact	A	-1.29	0.37	217.1	-3.52	***	-2.98	0.47	49.3			-1.78	0.47	46.7		
	B	-1.49	0.37	65.0	-3.97	***	-2.44	0.24	38.0			-1.24	0.24	40.9		
	Diff A-B	0.20	0.52	282.1	0.39	ns	-0.54	0.53	87.3	-1.02	ns	-0.54	0.53	87.6	-1.03	ns

Table 5. Test statistics for between-class effects (6 classes) in study B in statistical models where class identity and sex-by-class identity were entered as fixed effects. The models also accounted for random effects of test case nested within student identity, student identity and addition task identity. For simplicity these results are not presented. Statistical significances: ns: $p \geq 0.1$, *: $p < 0.05$, **: $p < 0.01$, ***: $p <0.001$, ****: $p < 0.0001$.

	Error			Counting			Direct retrieval			Derived fact		
Fixed effect	F	df	Sign	F	df	Sign	F	df	Sign	F	df	Sign
Sex	0.51	1,70.2	ns	19.64	1,67.2	****	10.34	1,65.1	**	13.1	1,57.1	***
Year	19.13	1,66.5	****	17.12	1,68.7	****	22.63	1,63.9	****	19.31	1,50.0	****
classID	2.62	5,68.4	*	0.7	5,67.4	ns	0.62	5,64.8	ns	1.47	5,57.2	ns
Sex*classID	0.65	5,68.2	ns	0.33	5,67.3	ns	0.37	5,64.7	ns	0.71	5,57.1	ns

Figure 1: Proportional use of different addition strategies plotted against school year for girls and boys. Thick lines indicate the predicted functions for each sex with thin lines demarcating 95% confidence zones.

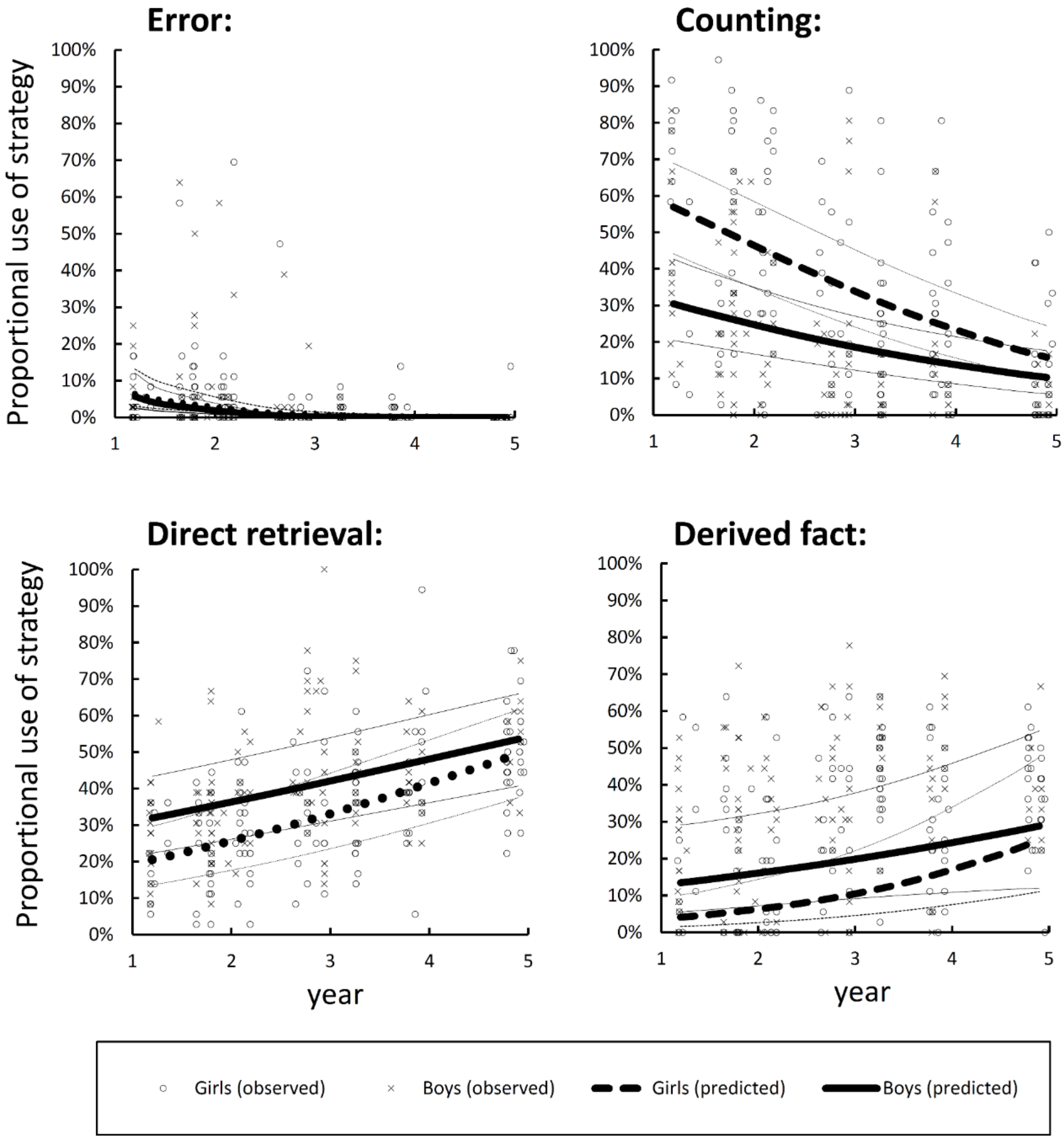
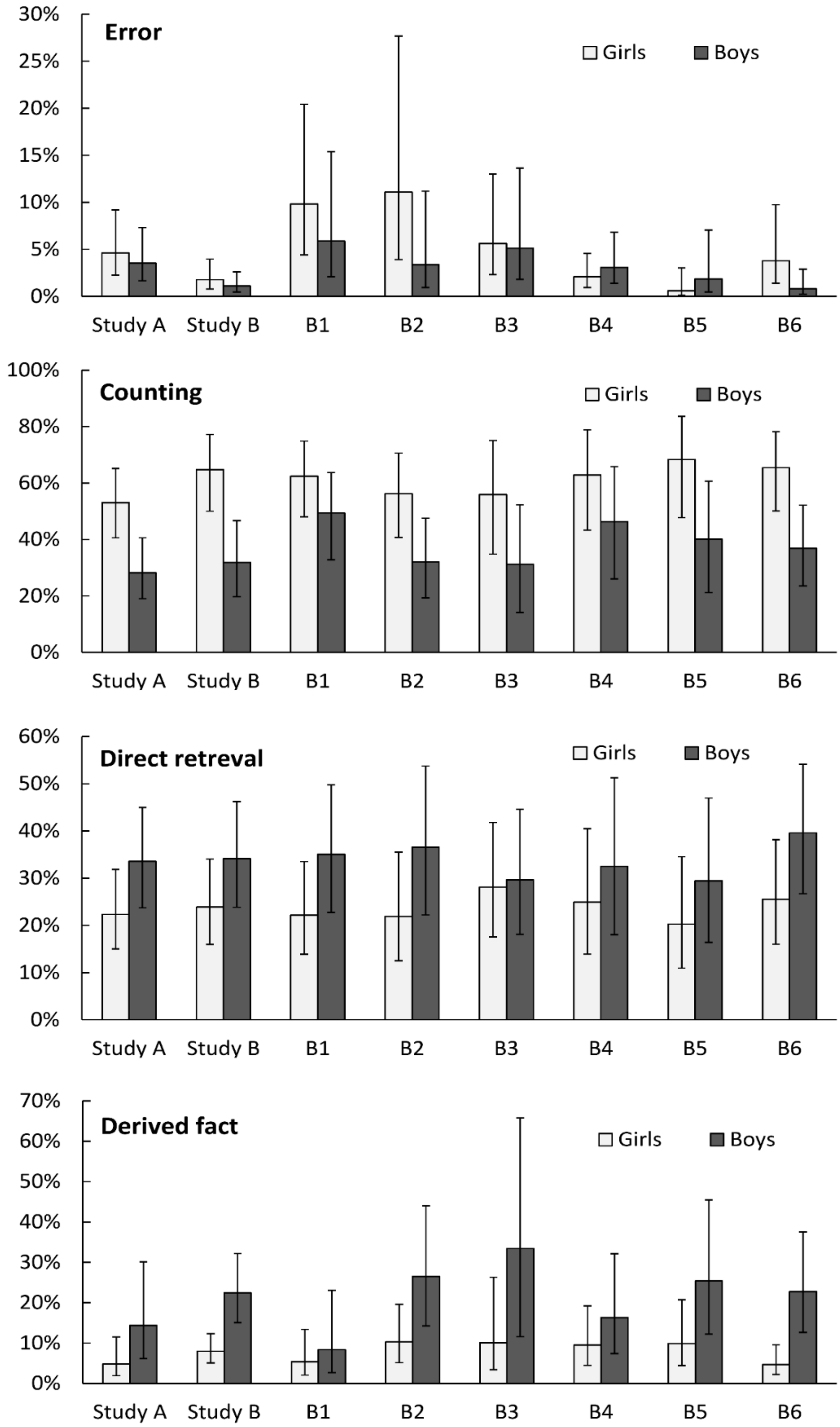


Figure 2. Strategy use by mid year one (least square means that account for random variation: error bars indicate 95% confidence intervals) by girls and boys divided on study and class within study B.



Paper II

Development and variance components in single-digit addition strategies in year one

Pernille B. Sunde[1, 2] and Peter Sunde[3]

[1]Danish School of Education, Aarhus University, [2]VIA University College, Denmark; pbos@via.dk

[3]Department of Bioscience, Aarhus University, Denmark; psu@bios.au.dk

In this paper, we aim to quantify students' development of mental strategies in single-digit addition over a five months period in six Danish year one classes dependent on their initial stage based on assessment interviews of 83 year one students' strategy use when answering 36 single-digit addition tasks. Students were interviewed twice (autumn and spring). Students using direct retrieval the least at interview 1 had the highest increase in proportional use at interview 2, whereas students using it the most showed no mean increase. For counting and derived fact, the change from interview 1 to 2 was constant irrespective of initial use. Analysis of how much of the variation in change in proportional use of different strategies was explained by different factors showed that 60-83% of the total variation was explained by the students' initial strategy use. Factors attributable to the specific learning environment and instructional practice did not explain any significant variation.

Keywords: Mental computation, addition, primary education, individual differences.

Background

Whole number arithmetic is a foundational subject in early years' mathematics. The first arithmetic operation students encounter in school is addition and simple arithmetic, which is related to number sense according to Gersten, Jordan and Flojo (2005). Thus, knowledge on students' early use and development of strategies in addition contributes to the understanding of students' development of arithmetic proficiency in general (e.g. Laski, Ermakova, & Vasilyeva, 2014). Furthermore, students' early qualitative use of strategies in mental arithmetic has been shown to influence development of mathematics competence and to be a valid predictor of later mathematical achievement and difficulties (Carr & Alexeev, 2011; Dowker, 2014; Gersten et al., 2005; Ostad, 1997).

Strategies for single-digit mental addition can be divided into three categories as described by Carpenter and Moser (1984): (1) Direct modelling strategies with fingers or physical objects, where counting sequences are used to count all concrete objects. (2) Counting strategies, where mental counting is used in counting-all or counting-on from either the first or, later, the largest addend. (3) Number fact strategies, which involve either a) direct retrieval, which is the direct recalling of the sum, or b) derived fact strategies, where known facts are used to derive the answer. Interesting to note that derived fact strategies has been the focus for many recent studies (e.g. Dowker, 2014; Gaidoschik, 2012; Laski, et al., 2014).

Students strategy use in arithmetic is influenced by many factors related to the individual student, for example working memory (e.g. Imbo & Vandierendonck, 2007) and gender (e.g. Carr, & Alexeev, 2011), as well as the learning environment, such as instructional practice (e.g. Timmermans, Van Lieshout, & Verhoeven, 2007), and parents and teachers' strategy preferences (Carr, Jessup, & Fuller, 1999). Thus, Carr et al. (1999) found that boys would use the strategies their parents use, which include the use of retrieval strategies. These strategies were also those their teachers were more likely to teach them.

Furthermore, students tend to use the strategies they are confident with in a feedback loop as described by Bailey, Littlefield, and Geary (2012), where early preference for a strategy predicts later use, which in turn predicts later preference. Students, who primarily use counting in early arithmetic at the expense of derived fact strategies, will get less practise in working flexibly with numbers. As a consequence, they will have fewer options to develop the more complex comprehension required in advanced arithmetic (Gersten et al., 2005). Thus, students change in strategy use are found to be dependent on individual factors, such as gender and initial strategy use, as well as factors related to the learning environment and learning opportunities, such as instructional practice.

The study of Gaidoschik (2012) is an example of a study where students' development of strategies for single-digit addition in year one are investigated in relation to both, students' prior knowledge and teachers' instructional practice. Gaidoschik (2012) found that strategy development was significantly related to students' prior number knowledge. Furthermore, he found in a qualitative analysis that although teachers' instructional practice focussed on counting strategies in year one, many students developed derived facts and automatized sums. This finding indicates instructional practice had minor influence on these students development of strategies. Although several studies have documented effects of students' prior knowledge and instructional practice, there is less evidence on the relative contribution of individual differences and instructional practice for students' development of strategies.

The current study

The data presented here are part of a larger project on the relationship between students' development of mental strategies in single-digit addition, teachers' perceptions of the learning of number and arithmetic in year one, and classroom instruction. In this paper, we aim to quantify students' development of mental strategies in single-digit addition over a five month period in six Danish year one classes dependent on their initial strategy use. Furthermore, we investigate how much different variance components contribute to the overall variation in year one students' development of strategies. We are specifically interested in quantifying how much of the total variation in the students' strategy use at time 2 that could be explained by their strategy use at time 1, instruction time and possible differences between classes in general as well as in interaction with time as the latter would include all variation due to possible effects of instruction.

Methods and analyses

The study and participants

The study consisted of 155 assessment interviews on strategy use for single-digit addition conducted on 83 year one students (46 girls; age 7) from six Danish year one classes. In each class, between 11 and 18 students were selected randomly though balanced with respect to gender. 72 students were interviewed in November 2015 and April 2016. 11 students were interviewed only once due to absence. To evaluate differences and similarities in the learning environment of the six classes, the six teachers (three female and three male) were interviewed on the teaching and learning of arithmetic and number. Furthermore, their instructional practice was documented through classroom video observations. These data will be analysed and discussed in detail elsewhere, but initial analysis indicated the teachers differed with respect to emphasis on teaching mental strategies for arithmetic and focus on procedural as opposed to conceptual knowledge. For example, two teachers emphasised rote learning of facts and procedural learning of counting strategies, two teachers focussed on students' skilled use of derived fact strategies and taught these explicitly, and two teachers were conscious of both rote learning of facts and derived fact strategies but did not emphasise one over the other. Furthermore, a previous study showed that the teachers' expectations for their year one students' development in number and arithmetic differed substantially (with one to two years of learning) (Sunde and Sayers, 2017). Thus, with respect to differences in teachers approach to teaching and learning arithmetic the six classes represents a continuum from high level of focus on explicit instruction of strategies with conceptual understanding to an instructional practice focussing on automatization of facts and procedural understanding.

Assessment of students' strategy use

Students' use of mental strategies for single-digit addition was assessed in one-to-one assessment interviews where students were presented with flashcards of the 36 addition tasks with numerals 2-9, including doubles. Manipulatives, paper and pencil were not accessible for the student. The flashcards were presented with tasks of increasing difficulty to avoid less confident students to give up or only use counting if presented with difficult questions at the beginning of the interview. Interviews took place in a quiet room at the school and each session lasted 10-30 minutes. We asked the student to explain how he or she found the answer. Based on students' explanations and the interviewer's observation, answers were categorised in four categories. 'Error' was used if the student gave up or miscalculated. For correct answers 'Counting' included all variations of counting procedures on fingers, verbal or self-report as mental counting, 'Direct retrieval' was used for automatized sum, and 'Derived fact' included all variations of decomposing addends and using automatized sums to calculate the answer (e.g. 4+5 = 4+4+1 or 5+5-1) (for more detail see Sunde, Sunde, & Sayers, 2019).

For each assessment interview, we calculated a strategy use profile on the basis of the proportion of answers (denoted 'proportional strategy use' hereafter) to the 36 addition tasks scored as 'error', 'counting', 'direct retrieval', and 'derived fact'. Hence, if a student used 'counting' to solve 6 of the 36 tasks, the proportion for counting, p_{counting}, was 6/36 = 0.1667. Original values were logit transformed prior to the analysis (logit $p_i = \ln[p_i/\{1-p_i\}]$). Before transformation, we substituted zero values (not defined on the logit scale) by 0.01, which is 2.8 times lower than the minimum achievable value larger than 0 (1 of 36 items = 0.028).

Analysis of change in proportional strategy use

Development in mean use of each of the four strategies over half a school year was analysed by modelling proportional strategy use in the second assessment interview (p_{i2}) as a linear regression function of the students' proportional strategy use in the first assessment interview (p_{i1}) five months earlier. From this regression function ($\hat{p}_{i2} = \beta_0 + \beta_1 p_{i1}$) the mean change in (logit-transformed) p_i from a given level in autumn to the next spring could be derived as the vertical distance between the regression line and the reference line y=x that indicated no change in mean strategy use between the two test sessions. Hence, the more students increases the proportional use of a given strategy from autumn to spring, the higher above the y=x line will the data points and the regression line be located and vice versa. From the slope of the regression line, it is furthermore possible to test whether the change in use of a given strategy is conditional on how much it was used in the first interview. If the mean change in proportional use (on logit scale) of a given strategy from interview one to interview two is equal for all students irrespective of how much they used the strategy on the first assessment interview, the slope of the regression line would not be significantly different from 1. On the contrary, a slope < 1 indicate a greater change in proportional use of a strategy by those students that used the strategy least in the first interview whereas a slope >1 would indicate the opposite, i.e. that those students that used a strategy least in interview one are also those that have developed relatively less from interview one to interview two.

Before the linear regression functions were established, we tested for possible quadratic and cubic relationships between p_{i2} and p_{i1}, as well as for possible effects of class (for the case development in strategy use varied between classes between interview one and two) and sex (for the case boys and girls differed in change of strategy use between interview one and two) without finding any significant effects of those predictors.

Analysis of variance components

We quantified how much of the variance in proportional use of each strategy was explained by different factors. The factors were time (difference between interview one and two), student identity nested within class (the systematic difference between students over the two interviews within a given class), class (the systematic difference between students in different classes) and the interaction between class and time (systematic differences in how much students from different classes changed from interview one to two). We calculated variance components as each of the aforementioned factors' sum-of-square compared to the total sum-of-squares in General linear models with all 155 interview results as observation units and the aforementioned variables and interaction terms as fixed effects. In this respect, the interaction term class*time encompasses all variation that could be explained by differential instructional practice because it represents the time a student has received instruction in a specific classroom. We derived statistical significances of each factor from F-tests.

Results

Change in proportional strategy use: Conditional on initial strategy use

As expected, mean use of all four strategies significantly changed from autumn to spring (all p-values ≤ 0.001) as the proportions of errors and counting decreased and direct retrieval and decomposition increased. The coefficients of determination (R^2) when modelling results from assessment interview two as functions of the results from interview one showed that the changes in the students' proportional use of a strategy was highly conditional on initial strategy use (Figure 1: R^2-values: 0.26-0.69). The slopes of the regression lines were significantly < 1 for 'error' ($t_{71} = 6.52$, $P < 0.0001$) and 'direct retrieval' ($t_{71} = 2.18$, $P = 0.03$), but did not differ from 1 for 'derived facts' ($t_{71} = 1.75$, $P = 0.08$) and 'counting' ($t_{71} = 0.6$, $P = 0.4$).

For ‘error’, the slope lower than 1 indicates that students with the highest error rates in interview one displayed the highest reduction in error rates between assessment interview one and two, probably as the mere result of the vast majority of the students solving all items correctly in general and especially in interview two. For ‘counting’ and ‘derived fact’ slopes not significantly different from 1 indicated that the mean change in use of these two strategies from interview one to two (decrease in counting, increase in derived fact) were constant no matter how much the strategies were used in the first assessment interview situation. For ‘direct retrieval’ a slope lower than 1 and inspection of the confidence zones of the regression line suggested that students using direct retrieval least in the first interview expressed the highest increase in use of this strategy in interview two. Whereas, students with the highest use of this strategy in interview one did not show any mean increase in use of this strategy in interview two.

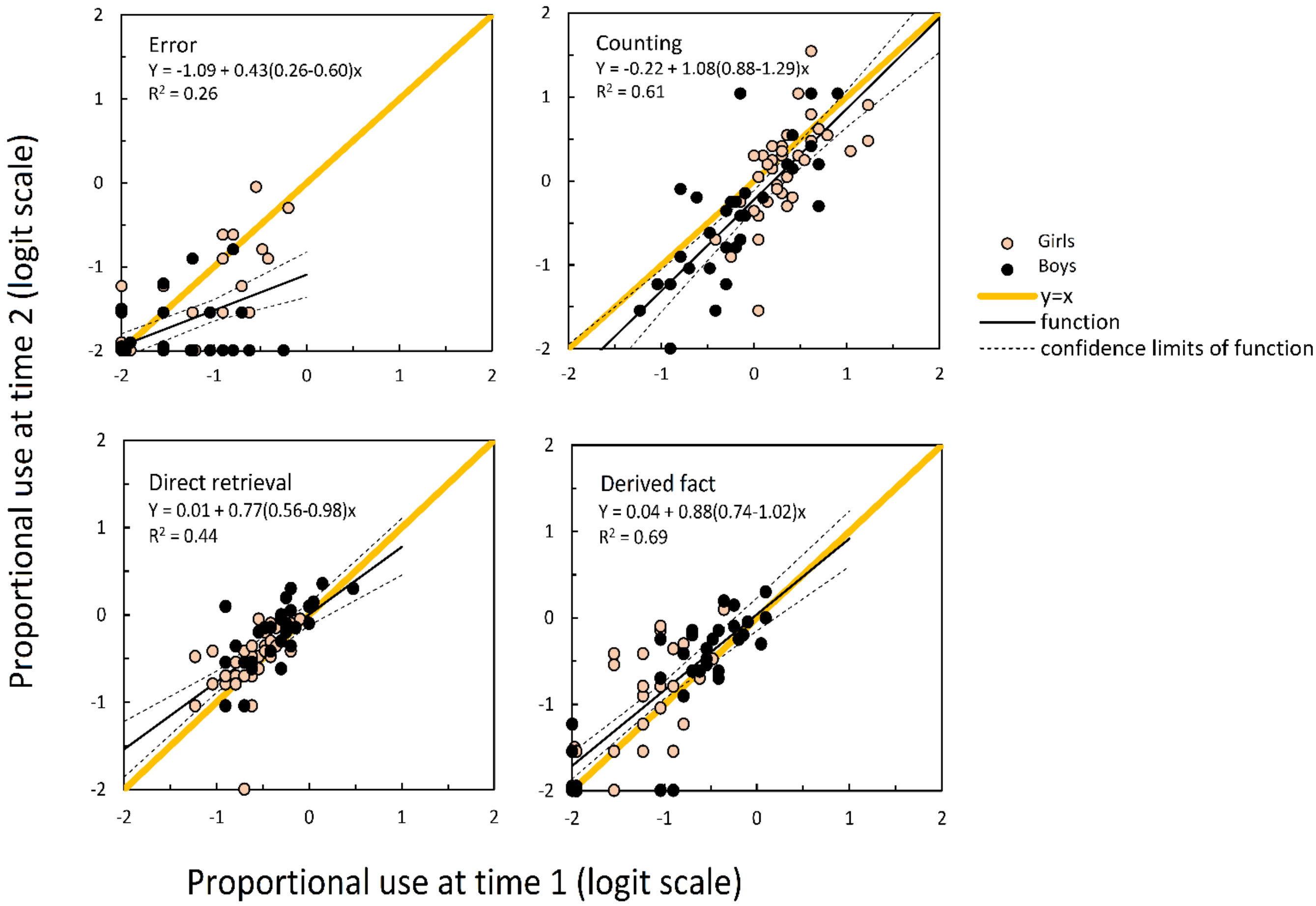
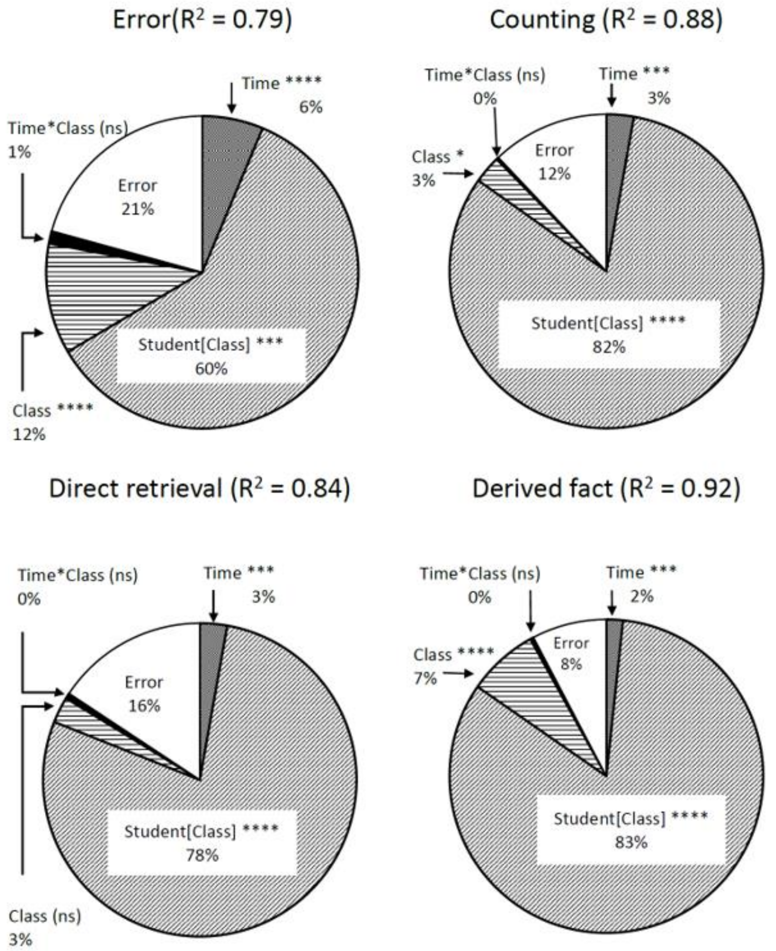


Figure 1: Proportional use (logit-scale: zero-values replaced by 0.01 before transformation) of strategies of 72 year one students in spring (April) plotted against their results achieved in the same assessment interview the preceding autumn (November)

From inspection of Figure 1, one can see that girls generally used counting substantially more than boys, and boys significantly used derived fact more than girls. These sex differences that were already present at interview one are dealt with elsewhere (Sunde et al., 2019). As regards changes from interview one to interview two, boys and girls did not differ for any of the strategies (additive effect of sex if included in regression models: all p-values > 0.2).

Variance components: Effect of individual differences, time and class

The half school year represented by the mean difference in strategy use between assessment interview one and two explained 3-6 % of the total variation in use of the four strategies, as compared to 60-83% explained by student identity nested within class (all p-values ≤ 0.001) and 3-12 % by class (statistically significant for ‘error’, ‘counting’ and ‘derived facts’: Figure 2). The time-by-class interactions (representing specific instructional practice) explained 0-1 % indicating no measurable class differences in mean progression from interview one to two (Figure 2).



Fixed effects:

Time: Difference between autumn and spring, i.e. 5 months, df=1

Student (Class): Student identity nested within class, df=77

Class: The specific class the student attend df=5

Time*Class: the interaction between class and time, df=5

ns: $p \geq 0.05$
*: $p < 0.05$
**: $p < 0.01$
***: $p < 0.001$,
****: $p < 0.0001$

Figure 2: Variance components in terms of Type 3 sum-of-squares of strategy use assessed twice in year one (autumn and spring), derived from general linear models

Discussion

This statistical investigation of strategy use in single-digit addition amongst year one students revealed the following three results worth noting. First, the slopes of the regression lines of strategy use at interview two as function of strategy use at interview one demonstrates that in at least two of four strategies the mean change in strategy use from interview one to two differed significantly between those students who used the strategy most and least. In the present case those students with highest proportion of errors in interview one expressed the largest reduction in error frequency in interview two five months’ later. Those who used direct retrieval least in the first interview expressed the highest increase in use of this strategy to interview two. For counting, the mean change in strategy use did not differ significantly between students with high and low initial use. This corresponds with Bailey et al.'s (2012) feedback loop of relationship between skills and preference for specific strategies.

Considering that use of derived fact strategies are related to high arithmetic ability (Dowker, 2014) and the relationship between excessive use of counting and mathematical difficulties (Ostad, 1997), it is concerning that the students with high initial proportional use of counting do not decrease the use of this strategy more than students with a low initial use. These findings underline the importance of analysing and reporting changes in strategy use (or any other achievement measure) as function of a given time period or intervention type differentially for student groups with different start level. In that respect, the here used graphical and statistical method based on testing the nil-expectation of a slope = 1, may provide an easily used and comprehendible method. One implication for practice of these results is that instruction should support students in developing a diverse strategy use.

Second, albeit mean use of all four strategies significantly changed over the five months between interview one and two (as expected), these changes, reflecting the fundamental 'effect' of learning and maturation, were modest compared to the variation in strategy use between students (evident as the high correlation coefficients of the scatter plots in Figure 1, as well as the amount of variance components explained by student ID in Figure 2). This illustrates the importance of acknowledging the long known but sometimes implicitly forgotten notion that students within a class may vary in strategy use patterns equalling several years' education time.

Third and finally, our analysis revealed that for Danish primary school children little variation (1-7%) appeared to exist between classes in general (main effect of class), and even less (0-1%) between classes and time (class-by-time interaction). Given that the five months' time span represents a big part of the students' first formal mathematics education, we find it interesting that the different classes in effect appeared to develop similarly between interview one and two despite considerable differences in teaching practice.

We want to stress that from this, one cannot conclude that differential teaching practices do not influence learning outcome, just that such differences were barely measurable over a time scale of five month's education time. A delayed effect of different instructional practices in year one on later arithmetic ability is thus possible. It is also important to notice that effects of differential teaching practices may have occurred before November, when interview one took place. Significant main effects of class for counting and especially derived fact indicate that this might have been the case.

In conclusion, Danish year one students' developmental change of strategy use for single-digit addition is primarily determined by their strategy use when entering year one. Development patterns were similar between classes, suggesting no influence of instructional practice, but mean change in error rate and use of direct retrieval differed for students with high or low initial use. The relatively small influence of five months teaching on strategy use patterns and the lack of difference between classes with different instructional practice is stunning. The findings call for qualitative research on what role different instructional practice play in year one students' development of strategy use.

References

Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The codevelopment of skill at and preference for use of retrieval-based processes for solving addition problems: Individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology, 113*(1), 78–92.

Carpenter, T. & Moser, J. (1984). The Acquisition of Addition and Subtraction Concepts in Grades One through Three. *Journal for Research in Mathematics Education, 15*(3), pp. 179–202.

Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology, 103*(3), 617–631.

Carr, M., Jessup, D. L., & Fuller, D. (1999). Gender differences in first-grade mathematics strategy use: Parent and teacher contributions. *Journal for Research in Mathematics Education, 30*(1), 20–46. doi:10.2307/749628

Dowker, A. (2014). Young children's use of derived fact strategies for addition and subtraction. *Frontiers in Human Neuroscience, 8,* 1-9. doi:10.3389/fnhum.2013.00924

Gaidoschik, M. (2012). First-graders' development of calculation strategies: How deriving facts helps automatize facts. *Journal für Mathematik-Didaktik, 33*(2), 287–315.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*(4), 293–304.

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology, 96*(4), 284–309. doi:10.1016/j.jecp.2006.09.001

Laski, E. V., Ermakova, A., & Vasilyeva, M. (2014). Early use of decomposition for addition and its relation to base-10 knowledge. *Journal of Applied Developmental Psychology, 35*(5), 444–454.

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology, 67*(3), 345–357.

Sunde, P. B., & Sayers, J. (2017). Danish teachers' expectations of year one pupils' additive competence. Paper presented at The Eighth Nordic Conference on Mathematics Education, Norma 17, Stockholm, SE.

Sunde, P. B., Sunde, P., & Sayers, J. (2019). *Mental strategies for single-digit addition in the first years of school: Girls count, boys retrieve.* Manuscript submitted for publication.

Timmermans, R. E., Van Lieshout, E. C., & Verhoeven, L. (2007). Gender-related effects of contemporary math instruction for low performers on problem-solving behavior. *Learning and Instruction, 17*(1), 42–54.

Paper III

Single-Digit Strategies in Year One Predicts Achievement in Year Four

Pernille B. Sunde[a,b] and Peter Sunde[c]

[a]Danish School of Education, Aarhus University, Copenhagen, Denmark; [b]VIA University College, Aarhus, Denmark; [c]Department of Bioscience, Aarhus University, Aarhus, Denmark.

Students who perform lower than their peers at the beginning of school in number and arithmetic are at risk of being low performers in later school years and for life. Early detection students at risk of becoming low achievers is therefore of paramount importance. We investigated the extent to which achievement within different types of more advanced mathematics (Number and arithmetic, proficiency with fractions, word problems) of 63 Danish primary school student in year four could be predicted from assessment results obtained in early year one of general mathematical achievement and strategy use in simple digit addition. In total, 28-39% of the variation in the three achievement response variables in year four could be explained by fixed-effect predictors obtained in the first half of year one. We found that the frequency by which certain strategies are used to solve 36 predefined addition problems in the first half of year one correlated significantly with achievement in number and arithmetic (r^2=0.13), word problems (r^2=0.22) and fraction knowledge r^2=0.22) in year four. For word problems and fraction knowledge, strategy use measures (i.e. the proportion of problems solved with 'count all') explained a statistically significant additional amount of variation in year four achievement that could not be accounted for by achievement scores from a standard achievement test for year one that is widely used in Danish primary schools. Boys generally expressed closer correlations between strategy use patterns in year one and achievement scores in year four than girls. For fraction knowledge this difference was apparent as a significant sex-by-strategy use interaction. From these results, we hypothesise that systematically obtained measures of how young students solve simple single-digit addition problems might be provide useful information about their foundational number knowledge that in turns may reveal how well they achieve in later school years.

Introduction

Students who perform lower than their peers at the beginning of school in number and arithmetic are at risk of being low performers throughout school (Duncan et al., 2007), with consequences for life after school as well (Parsons & Bynner, 2005). Early predictors of later mathematical achievement are therefore useful and important as tools to detect students at risk of developing mathematics difficulties and therefore need intervention (Geary, 2011; Gersten et al., 2005). Numerical and quantitative knowledge (Jordan et al., 2013, 2009; M. M. M. Mazzocco & Thompson, 2005) and early arithmetic ability and strategy use (Bailey et al., 2012; Carr et al., 1999; Geary, 2011; Shen et al., 2016) has been found to be related to students' general achievement in mathematics and could therefore be relevant predictors of later potential difficulties or low performance. For instance, Jordan et al. (2009) found that 66% of the variation in mathematical achievement in year three could be explained by number competence in the beginning of kindergarten. Likewise, Geary (2011) found that a number of quantitative skills and competences in year one predicted year five mathematical achievement beyond the contribution of several domain general abilities.

In this paper, we investigate the extent to which achievement within different types of more complex mathematics (Number and arithmetic, proficiency with fractions, word problems) of Danish primary school students in year four can be predicted from assessment results obtained in early year one of general mathematical achievement and strategy use in simple digit addition. First, and most basically we are interested in elucidating how much variation in year four achievement that can be explained by different assessment measures in isolation and in combination. Second, and more specifically, we want to investigate whether information on strategy use in single digit addition ('how students solve simple addition problems') might provide additional, predictive information about students' mathematical achievement three years later that is not already available from a standard achievement test (that assess how successful students solve simple addition problems). Our focus on strategy use as potential indicator of mathematic achievement later in school is rooted in a suspicion that because simple addition problems (on which students work in early year one) can be solved relatively successfully with simple, unsophisticated strategies without the student necessarily understanding why, students with poor arithmetic understanding may possibly perform quite well in achievement tests on simple problems if they have optimised primitive strategies such as counting. If that is the case, the frequency by which they use primitive versus more advanced strategies may possibly tell more about their arithmetic understanding than the number of very simple problems they solve correctly.

Mental strategies for single-digit addition are generally categorised in counting strategies and fact based strategies. Counting can be further divided into several subcategories (e.g. Baroody, 1989). Most common is to distinguish between counting both addends (count all) and counting on from one or the larger addend (count on). Counting strategies are reported to be the first strategies students are introduced to and use even before formal schooling (e.g. Baroody & Wilkins, 1999; Clements & Sarama, 2007). Fact based strategies are categorised in direct retrieval (the sum is retrieved directly from memory) and derived fact strategies, where the sum is calculated by decomposing the addends into other know sums and/or subsequent counting. This is a step-wise process where known addition facts are used to derive the result. An example of this process is finding the answer to 7+8 by adding one to the known fact 7+7.

In general, fact based strategies are thought of as more advanced strategies, not just because they are generally faster than counting strategies (Cowan, 2003; Cowan & Powell, 2014) but also because they built on more complex number understanding than counting. Counting as addition strategy built on an understanding of cardinality and one-to-one correspondence (Fuson, 1992). Fact based strategies are mental procedures that depends on prior knowledge of number facts (e.g. Geary, 2011; Threlfall, 2002) as well as the understanding of part-whole relations and partitioning of numbers (e.g. Ambrose et al., 2003; E. M. Gray & Tall, 1994). Furthermore, students using fact based strategies tend to have higher overall mathematics performance compared to students that rely on counting strategies (Carr & Alexeev, 2011; Carr et al., 2008; Geary et al., 2004; Gersten et al., 2005; Price et al., 2013).

The use of specific strategies are associated with different aspects of mathematical knowledge and components of number sense as for example symbolic numerical magnitude processing skills (Vanbinst et al., 2012, 2014), base-ten knowledge (Laski et al., 2014), and understanding of mathematics equivalence (Chesney et al., 2014).

Vanbinst et al. (2012, 2014) found that the frequency of direct retrieval as well as the speed of retrieving addition and subtraction facts were positively correlated with symbolic magnitude processing skills. However, they did not distinguish between counting and decomposing (derived fact) as both types of strategies were coded as ‘procedure’ as opposed to ‘retrieval’.

Laski et al. (2014) found that kindergarteners’ use of derived fact strategies based on base-10 decomposition was dependent on their knowledge of base-10 number structures and Chesney et al. (2014) found that the use of decomposition strategies has been found to be related to students understanding of mathematics equivalence. Students who understand that a number can be represented by many combinations of addends (e.g. 6=2+4, 6=1+5, 6=3+1+1+1 etc.) and organize their addition knowledge based on these equivalent values have a better understanding of mathematic equivalence. Thus, the use of derived fact strategies not only built on complex understanding of numbers but their use are also associated with understanding of more complex mathematics.

Torbeyns, Verschaffel, & Ghesquière (2004) showed that no matter the students’ general mathematic ability they may use a wide variety of strategies from counting to derived fact strategies. However, with respect to the relative use of strategies they found considerable differences between groups as well as individuals. Low achieving students use counting more often than fact-based strategies while high achieving students use fact-based strategies more often (Dowker, 2014; Torbeyns, Verschaffel, & Ghesquière, 2004).

These early differences in strategy use in arithmetic has been linked to later mathematic performance in general and arithmetic ability specifically (Gersten et al., 2005; Ostad, 1997a; Price et al., 2013). Studies have found that the specific strategy use in solving mental addition and subtraction influence the development of mathematics performance (e.g. Carr et al., 2008; Cowan et al., 2011; Fennema et al., 1998; Geary, 2011; Ostad, 1997a). Longitudinal studies show that with increasing school age students with mathematical difficulties tend to use more counting strategies persistently whereas students without difficulties shift to using more fact-based strategies (e.g. Ostad, 1997a; Vanbinst et al., 2014). Furthermore, Carr and Alexeev (2011) found that the use of cognitive strategies in second grade was important for development of mathematics performance in fourth grade while excessive use of counting strategies had negative influence on the students mathematics performance.

Although a relationship has been demonstrated between specific strategy use and later mathematical achievement, the knowledge of relationship between early strategy use and later achievement within specific mathematical knowledge areas are, to the best of our knowledge scarce.

This study will analyse the relationship of strategy use in year one with mathematic achievement in year four in a Danish context. In addition, the study will add to the current knowledge by analysing the specific contribution of strategy use in year one controlling for general arithmetic proficiency to year four mathematics achievement measured as number knowledge, arithmetic proficiency, proficiency with fractions, and word problem solving. Furthermore, we will investigate possible sex differences in these relationships.

The study will contribute with knowledge of the explanatory power of students' arithmetic proficiency and strategy use respectively in year one on their mathematics performance in year four.

Methods

Participants

The comparison between assessment results from year four as function of assessment results from year one is based on 63 students (36 girls, 27 boys) from six classes from three different schools. These students were interviewed on their strategy use and their general mathematical achievement in November in year one and in August in year four. The 63 students comprised as subsample of a total of initially 83 students in the six classes in year one of which 78 were tested in November and 76 in April in year one. For simplicity (and because our aim is to investigate the predictive power of assessment tests obtained early in year one) we only relate test results from year four with the results from the first (November) test, but provide the descriptive results from both test round to illustrate the general development in strategy use distributions among students in year one's first and second half, respectively.

Informed written parental consent was obtained for all students. Students were informed orally about the study and that participation was voluntary.

Assessment of strategy use year one

Strategy use was assessed in one-to-one interviews were the student was presented with flashcards with addition tasks. The set of tasks consisted of the 36 addition task with addends 2 to 9, including the doubles. The flashcards with sums less than 10 were presented first to avoid less confident students to give up or only use counting because they were presented with difficult tasks (e.g. 6+8) in the beginning of the interview. The interview took place in the student's school in familiar quiet settings and the student did not have access to manipulatives or paper and pencil. Each interview lasted 10-30 minutes.

The interviewer stated to each student: "I will show you some single-digit addition tasks. First, I would like you to find the answer to the task, and then we talk about how you found the answer. There are many ways to find the answer to an addition task. Sometimes you might know the answer or count or perhaps you use other tasks to find the answer. I am interested in knowing how you find the answer." Then the student was presented for a flashcard, e.g. 4+5, and the interviewer asked: "what is the answer to four plus five?". If the student did not give an explanation following the answer, the interviewer asked: “how did you find the answer?” and if further prompting was needed: “did you count or did you just know the answer, or did you use some other tasks you know to find the answer?”

Categorisation of strategy use.

During the interview, the student’s answers were categorised into five categories, one for incorrect answers or instances where the student gave up, ‘error’ which was not categorised further, and four for correct answers. Correct answers were categorised based on the student’s self-report of strategy use and observations by the interviewer (e.g. visible signs of finger-counting or lip movements).

Strategy use of correct answers were categorised in ’counting all’ (counting both addends and then all together), ‘counting on’ (counting on from one of the addends), ’direct retrieval (reported just knowing the answer), and ’derived fact’ (decomposing addends and calculating answers using automatized sums with subsequent use of addition e.g. 4 + 5 = 4 + 4 + 1 or subtraction, e.g. 4 + 5 = 5 + 5 - 1). We did not distinguish between self-report of mental counting and visible (e.g. on fingers) or audible counting.

The relative use of the different categories was assessed as the proportion of answers to the 36 addition tasks scored as being solved with strategy in question (denoted ‘proportional strategy use’ hereafter). Hence, if a student used ‘counting all’ to solve 6 of the 36 tasks, the proportion for counting all was 6/36 = 0.1667. For the case the two counting strategies (counting all, counting on) should represent the same type of underlying process and number sense, we also created a combined variable of the sum of the two, denoted ‘counting’. Original values were logit transformed prior to the analysis (logit $p = \ln[p/\{1-p\}]$). Before transformation, we substituted zero values (not defined on the logit scale) by 0.01, which is 2.8 times lower than the minimum achievable value larger than 0 (1 of 36 items = 0.028).

Assessment of arithmetic proficiency year one

The students’ arithmetic proficiency was tested with a curriculum-based Danish standardised test, MAT1 (Jensen & Jørgensen, 2007). MAT1 is a paper and pencil test and measures a wide range of mathematical skills within numbers, arithmetic, algebra, geometry, and applied mathematics. The test is not time restricted and time is not monitored. For further description of MAT test, see Sunde and Pind (2016). For the present analysis, we used the students’ scores (number of incorrect answers) in items on number and arithmetic.

Assessment of mathematical proficiency year four

This is a computerized test consisting of three subtests. The first subtest NUMAR comprises 40 items of calculation in the four operations and number knowledge, whole number and decimal number. The second subtest FRAC comprises 36 items on comparing fractions, addition and subtraction with fractions. The third subtest WP consists of 20 word problems in the four operations and fractions. In this test, the students were allowed to use calculator. Two of the tests were time restricted with 10 minutes for each test (NUMAR and FRAC) and one were not time restricted (WP). Only few students managed to solve all items in the time restricted tests.

The students were tested in the beginning of year four in the classroom using either tablets or computers. All students' were familiar with using computers or tablets for testing as well as in everyday lessons. The students' mathematics teacher administered the tests. For each subtest, the achievement score was calculated as the total number of correct answers. Of practical reasons, not all students tested in year one were retested on all problem types in early year four. This was particularly the case for the test on FRAC, where only three of the six classes (34 students) participated, whereas all six classes (63 and 61 students respectively) took the WP and NUMAR tests.

Statistical analyses

Principal component axes of proportional strategy use

To achieve statistically independent measures of strategy use (the five strategy use variables were based on proportional values that summed up to 1 for each student, wherefore their values per definition to some extent were dependent and negatively auto-correlated), with a principal component analysis (PCA) we condensed the information from the five variables to four eigenvalues. The PCA-scores were estimated from strategy use patterns of the tests conducted November in year one.

Correlations between test measures

To establish the extent to which the different measures of mathematical achievement in year four was statistically associated with strategy use and mathematical achievement in year one, we established a correlation matrix of the simple ('raw') between the five year four mathematical achievement scores (WP, NUMAR, FRAC) and the strategy variables, principal component values and the Mat1-test. We used Pearson's product moment correlation coefficient (r) as test statistics, as its squared value (r^2, 'determination coefficient') expressed the amount of variation in the dependent variable that can be explained by the variation in the other variable (and vice versa). From the correlation matrix it was possible to spot which test result variables from year one that correlated with which achievement variables from year four, and compare their apparent predictive power. We did correlations for the entire student sample and split on sex.

To test for possible class differences in year four mathematical achievement variables, and to quantify approximately how much of the variation in the test variables that could be explained by class affiliation, we ran simple General Linear Models (GML procedure in SAS 9.4) with class identity as fixed effect.

To test for possible gender differences in year four mathematical achievement variables, we ran General Linear Mixed Models (Mixed procedure in SAS 9.4) with sex as fixed effect and class ID as random effect.

Mathematics achievement in grade 4 as complex functions of 1st grade-results, class identity and sex

We created three types of predictive models: (A) a full model including MAT1-test and all four PCA-scores as predictors, (B) a model with MAT1-test as mandatory variable to which a step-wise selection procedure added those predictors (selected among the five strategy use variables, sex and second-order interactions with sex) that improved the model's overall fit, and (C) a model similar to model (B), except that MAT1-test was not forced into the model (i.e. MAT-test was eligible for the model on the same conditions as all other predictor variables). As post-hoc operations we manually tested for possible interactions sex and the covariate(s) included in model (C). If the combined additional effect sex plus the interaction term between sex and the selected covariate was statistically significant ($p < 0.05$), we presented this interaction model as model (D). In a case were sex came out as a main effect, we also presented model without this main effect as the model D.

From model A, it was possible to compare the relative, predictive influence of the general achievement (MAT1-test) and the four independent strategy use PCA-variables when forced into the same model, albeit accepting the risk of overfitting the model because of potential variable redundancy between the MAT-test predictor and the four principle component variables representing strategy use.

From model type B, it was possible to test outright, whether information from variation in strategy use, sex or the interaction between strategy use and sex represented additional, predictive variation to the general mathematical achievement test in order to explain variation in year four mathematical achievement. In that case, the selected model would include one or more strategy use variables in addition to the MAT1-test variable that was forced into the model.

From model type C, we selected the 'best' possible (the least complex, adequate) model to explain variation in year four achievement from all information sources on strategy use and mathematical achievement in early year one. The purpose of model C was to test whether strategy use information by early year one possibly could be sufficient to explain all relevant variation in year four mathematical achievement without inclusion of information from the MAT1-test. In that case, the selected model would include strategy use predictors only, but not the MAT1-test variable.

The multiple regression models were produced with the HPREG and GLM procedures in SAS. The two procedures result in identical model predictions, but provide different model building options and information outputs. We used the HPREG procedure to find the most parsimonious explanatory statistical models (the models that explained most variation in the dependent variables with the fewest possible predictor variables) with forward-stepwise selection procedure with the lowest AICc (Akaike's Information Criterion corrected for low sample size) as forward selection criterion and a critical p-value of 0.05 for inclusion as well as exclusion of predictors in the model.

By analysing the type-1 effects in the GLM model, we measured the additive effects and additive explanatory power of each variable when added to model in sequence.

For the full model, we expressed the partial effects of each variable in terms of absolute beta-coefficients (expressing how much the response value change on average when the predictor changes by one unit) and standardized beta-coefficients (expressing how much the response value change on average when the predictor changes by one SD unit: thereby one can compare relative effect sizes between different predictors).

Finally, to also address the predictive potential of the predictor variables within classes (for some response variables, there were statistical significant between-class-variation), we also produced a version of the predictive models that included class identity as a random effect with MIXED procedure in SAS.

Results

Strategy use in November and April of year one

In the first strategy use assessment by November in year one, the most frequently used strategy was counting on with an average of 13 items of 36 for girls and 10 items out of 36 for boys (Fig. 1), followed by counting all (11 and 6), direct retrieval (8 and 11), derived facts (2.0 and 7) and error (2.5 and 2.1)(Fig. 1). However, from visual inspection of the spread of observations (Fig. 1) as well as from standard deviations of the means and variance coefficients (CV=SD/$\bar{x}$) it was also clear that there was also a considerable variation between students in the frequency by which they used the different strategies. These overall patterns appeared to remain in the April assessment, albeit the frequency by with error and counting all occurred had been reduced and the use of direct retrieval and derived facts increased (Fig. 1).

Correlations between strategy use categories at the beginning of year one

Proportional strategy use correlated between several of the five main strategy use categories (Table 1). Most notably, use of the less advanced strategy 'count-all' correlated strongly, negatively (r = -0.47--0.66) with use of all other strategies except 'error' (Table 1). In comparison, the use of the two most advanced strategies, direct retrieval and derived fact strategies, were highly positively correlated (Table 1).

The eigenvalues of the first, second, third and fourth principal component axes explained 49, 24, 18 and 8% (cumulative: 99%) of the variance in the five strategy use variables. The first axis (PRIN1) correlated significantly with all strategy use variables, most negatively with count all (r=-0.91: Table 1) and most positively with derived fact (r=0.78: Table 1). The PRIN1-scores may therefore be considered to represent a gradient from the least to the most advanced strategies. The second PCA-score (PRIN2) correlated most negatively with 'count on' (r=-0.86: Table 1) and most positively with direct retrieval (r=0.56: Table 1) and derived fact (r=0.35: Table 1). The third PCA-score (PRIN3) correlated strongly positively with 'error' (r=0.88: Table 1), indicating that it mainly represented a gradient of (lack of) general mathematic achievement expressed as a high failure rate in problem solving. The fourth PCA-score (PRIN4) appeared to capture a difference between using direct retrieval much (r=0.51) and derived fact less (r=-0.43, Table 1).

Simple associations between mathematics achievement in year four and potential predictor variables

For all five main strategies, differential use in year one correlated with achievement in for at least one type of advanced mathematics problems in year four (Table 2). The same applied for two of the four derived PCA-scores (PRIN1 and PRIN3). Overall, use of error and count all in year one correlated negatively with achievement scores in year four, while high frequency of using count on, direct retrieval and derived fact correlated positively with achievement in year four. The magnitude and statistical significance of these correlations varied between strategy use scores at year one and problem types in the year four test (Table 2).

For the entire sample, particularly high correlations ($|r| \geq 0.45$ equal to an explained variation [r^2, determination coefficient] ≥ 20%) between genuine strategy use variables and mathematical achievement scores were found between count all-WP (r = -0.45) and count all-FRAC. When split on sex, these correlations were even stronger and highly statistically significant in boys, but weaker and not statistically significant in girls (Table 2).

The correlation patterns between PRIN1 and mathematical achievement scores were almost identical to those for count all (Table 2).

Of the achievement related test scores from year one assessments, determination coefficients ≥ 20% was found in girls between the MAT1-test score and WP (r=-0.48), AR (r = -0.65), and NUMAR (r = -0.52), respectively. Directions were similar, but weaker in boys (Table 2).

Statistically significant between-class differences in mathematical performance in the year four test (tested as fixed effect) appeared to exist for WP ($F_{5,55}$ = 13.17, P < 0.0001, R^2=0.55) and NUMAR ($F_{5,57}$ = 3.65, P = 0.006, R^2=0.24), but not FRAC ($F_{2,31}$ = 1.96, P = 0.16, R^2=0.11).

Statistically significant sex differences in mathematical performance in the year four test (adjusted for class) was found for NUMAR where boys on average scored 3.4 points more than girls (11.5 vs. 8.2 points, SE = 1.12, t_{56}=3.01, P = 0.004). Boys and girls performed similarly in all other mathematical performance categories (all p-values ≥ 0.26).

Mathematics achievement in grade 4 as complex functions of year one-results, class identity and sex

Word problem (WP)

When the full effects of MAT1 and strategy use (principle component axes 1-4) were forced into the same model, the five predictors combined explained 32% of the variation in WP, divided on ΔR^2=17% from the MAT1 when entered first, and ΔR^2=15% on the four PCA-variables when entered subsequently (Table 3). The partial effects were statistically significant for MAT1 (negative) and PRIN1 (positive). In addition, there was a borderline, negative significant effect of PRIN2. Standardized b-coefficients indicated a slightly higher partial effect size of PRIN1 than of MAT1. In addition, there was a highly significant between-class variation on WP-scores. If this between-class variation was included in the models, it reduced the effects PRIN1 (and PRIN2) that then no longer were statistically significant.

The constrained (model B: MAT1 forced in) as well as the unconstrained step-wise selection model (model C: all predictors equally eligible) comprised MAT1 and count all as predictors alongside, resulting in a combined explanatory power of 31% of the variation in WP. Of the two partial predictors, count all expressed the highest statistical significance, and was included in the model before MAT1 (Table 3). Both predictors were (about equally) significant when accounting for between-class variation (Table 3). Neither count all nor the effect MAT1 interacted with gender (additive effects of sex + sex*covariate; count all: $F_{2,53}$ = 0.91, P = 0.41; MAT: $F_{2,53}$ = 0.06, P = 0.94).

Number and arithmetic (NUMAR)

MAT1 was the only significant year one predictor, explaining 16% of the variance in NUMAR on which it had a negative effect (Table 4). In the full model (A), the effect of

PRIN1 (positive) was borderline significant and explained further 5% of the variance, while PRIN2-4 (all non-significant) explained additionally 2%. When forced into the full model (A) no partial effects were statistically significant, which demonstrate that the effects of MAT1 confounded strongly with the effects of PRIN1-4, resulting in an over-parameterized model.

Model B and C both selected MAT1 and sex as predictors. Girls averagely scored fewer points than did boys. The effects and significance of MAT1 and sex was similar if accounting for between-class variation. The effects of MAT1 and sex did not interact ($F_{1,56} = 0.28$, $P = 0.60$).

Fractions (FRAC:
MAT1 explained mere 6% of the variance in FRAC, an effect that was not statistically significant, while PRIN1 (positive effect) explained additionally 16% if entered after MAT1 in model A (Table 5). In model B, count all (negative effect) was selected with an additionally explained variance of 21%. Model C selected count all as the only necessary predictor with an explanatory power of 23% of the variance in FRAC (Table 5). Even though boys and girls scored equally on average, the effect of count all was significantly different in boys and girls ($F_{2,30} = 2.30$, $P = 0.02$), as the negative effect of count all on FRAC was stronger in boys than in girls (Fig. 3).

Discussion

The results from this analysis suggest that specific strategy use for single-digit addition assessed after few months' of formal mathematics education at 6-7 years' age, correlated significantly with the ability to solve all three types of problems such as proficiency with number and arithmetic (NUMAR), fraction knowledge (FRAC) and word problem solving (WP) three years later, with coefficients of determination of 13-22% of individual predictor variables. These findings are in line with Geary (2011), who found that early (year one) quantitative competences including skilled use of counting (high proportion of counting on), derived fact and direct retrieval strategies positively predicted later mathematics achievement in year five. Bailey, Siegler and Geary (2014) found that year one students whole number arithmetic proficiency (number of correct arithmetic tasks and fast direct retrieval of single-digit additions) predicted knowledge of fraction arithmetic in year seven and knowledge of fraction magnitude in year eight. Other longitudinal studies has also established a link between early quantitative and arithmetic competence and prediction of later mathematical achievement (Jordan et al., 2009; M. M. M. Mazzocco & Thompson, 2005) and Jordan et al. (2013) found that year three number line estimation skills made important unique contribution to prediction of year four fraction concepts and procedures.

More interestingly, for two of three problem types, WP and FRAC, measures of strategy use explained significant, additional variation in achievement of these problem types three years' ahead than did the results from the standard achievement test. For FRAC, measures of strategy use (i.e. the frequency by which students solved problems by counting all) even outperformed measures of achievement in order to explain all significant variation in achievement three years later. Hence, the proportional use of a single strategy category (i.e. counting all) was sufficient to account for all significant variation in FRAC.

From these results, it thus seems as Danish students' strategy use in single digit addition at the very start of their mathematical education, can tell as least as much about their ability to comprehend several types of more advanced mathematical topics of WP and FRAC three years later than can a general achievement test. This information might be useful for several reasons.

First, albeit the mere correlations between strategy use in year one and mathematical achievement in year four tell little about the underlying causal reasons behind this identified pattern, the r^2-values (13-22%) hints that much variation in problem solving strategies in year one relates to learning trajectories in the following years. The obvious follow-up question for future investigations is to pinpoint the extent to which achievement in advanced mathematics several years later can be improved through instruction aimed at the students' development and use of more advanced strategies in single-digit addition (i.e. a causal relationship between active training of early number sense and strategy use and comprehension of advanced mathematics).

In the present data, the incident that significant variation were apparent between classes in almost all parameters of year four achievement, and that this between-class variation accounted for some of the partial effects of year one strategy use on year four achievement, may suggest that the teaching environment and thus also teaching practice does explain a significant part of how well students achieve.

Second, from a screening perspective, where the aim is to identify students at risk of becoming low performers with potential difficulties, assessing strategy use may provide useful additive information in order to identify these students. In the specific cases in this study, frequent use of count all (or low values of principle component 1 that correlated strongly negatively with count all use and strongly positively with derived fact use) predicted proportionally low achievement in fraction and word problem solving three years later that was not captured by the standardised mathematical achievement score for MAT1-test. We think that the plausible explanation for this is, that the mathematically simple problems in the achievement test designed for year one students could be efficiently solved by use of the unsophisticated count all strategy. Students that relied excessively on the count all strategy in their problem solving in year one, but without having achieved any deeper understanding of number and arithmetic (for whatever reason) could thus perform reasonably well when tested in a general mathematical achievement test. Thereby their more fundamental problems of understanding number and arithmetic which influenced their achievement in more complex mathematics as for example fractions and word problem solving three years later were not revealed as clearly as if the information from their strategy use patterns had been included.

It is worth noting that the proportional use of the two counting strategies, count all and count on, correlated negatively with each other and yielded opposing predictions for mathematical achievement three years later: contrary to high use of counting all, a high use of counting on correlated positively with all three types of mathematical achievement in year four. Similarly, the two counting strategies correlated differently with PRIN1 that was the primary PCA-axis that expressed the gradient of sophistication in strategy us. From this, count on in the first half of year one should therefore not be considered as an indicator of unsophisticated strategy use and arithmetic comprehension, but rather the opposite. As a result, the combined, general count strategy, comprised by summing count all and count on had limited predictive power on the mathematical achievement in year four, simply because the two count strategies had opposing effects.

Proficiency with fractions is based on well-developed number sense, understanding part-whole relations and partitioning. Derived fact strategies build on a more complex understanding of numbers than do counting all. It requires that some facts are automatized (e.g. Threlfall, 2002) but also an understanding that numbers can be partitioned in many ways as well as an understanding of relations between operations (Ambrose et al., 2003; E. M. Gray & Tall, 1994; Verschaffel et al., 2007). The use of derived fact strategies has been found to relate to primary students' understanding of part-whole concepts (Canobi, 2004). Counting does not require the understanding of part-whole relations as an understanding of numbers as standalone entities will suffice (Fuson, 1992). As outlined by (Siegler, Thompson and Schneider (2011) and Van Dooren, Lehtinen and Verschaffe (2015) part of the explanation of the whole number bias is that whole number counting and understanding of numbers as discrete entities is not applicable to rational numbers.

Taking this into account, the association between the use of these two strategies in year one (as expressed in proportional use of counting all and the PCA-score for PRIN1) and fractional knowledge could be related to the students' underlying number sense, with students with a high proportional use of counting all might have less developed understanding of numbers, part-whole relations and relations between operations, important prerequisites for developing proficiency with fractions later. This could add to the findings of Bailey et al. (2014) as they did not assess the students strategy use but the students' general arithmetic proficiency and fast direct retrieval of single-digit additions by year one and the relationship with later fraction knowledge. They suggested that the found relation between arithmetic proficiency and later knowledge of fraction arithmetic could be explained by the fact that students skilled in whole number arithmetic will free memory working resources (Geary, 2011) and further make less errors in fraction arithmetic. The present study found that strategy use was a better predictor of later fraction knowledge than arithmetic proficiency. Thus indicating not only arithmetic proficiency but also number sense and understanding of arithmetic operations play a role in later understanding of fraction knowledge.

In that perspective, when students are tested at such an early stage of their formal education trail, information on how they solve single-digit addition problems and not just how well they solve them may potentially enhance the teacher's assessment of the student's understanding of number and arithmetic and thereby provide valuable information on students' at risk of developing mathematical difficulties or becoming low performers.

A third and final discussion point is the observation, that boys and girls appeared to differ for at least one strategy use (year one) – achievement (year four) correlation: Seemingly, the negative correlation between the frequency by which students use count all in year one and their achievement in FRAC in year four was stronger in boys than in girls. Given the modest sample size and the numerous models tested, we will be careful not to overemphasize the apparent sex interaction that is just modestly statistically significant. However, it is reflecting a factual, underlying interactive difference, one must also address that underlying mechanistic links between strategy use in year one and achievement of certain types mathematical problems later on may be systematically different for boys and girls. A possible explanation could be if girls are more inclined to use count all than boys even when they are capable of using retrieval strategies to solve a given problem, whereas boys only uses count all as strategy if not mastering retrieval strategies. In that case, the frequency by which boys are using count all will be a much more sensitive indicator of their number skills than it will be for girls, where the correlation will be blurred by a higher proportion of students with an excessively high use of count all compared to their number skills. If this is the case, one should predict that for the groups of students with highest use of count all, girls should on average outperform boys with similar use of count all in the fraction knowledge test three years later. Which was what we found. Before firm conclusions can be inferred on this point, this prediction should be tested on a larger data set. At this point, however, we consider the perspectives of using strategy use patterns in single-digit addition as an early indicator of young students' arithmetic understanding and foundational number knowledge and potential of facing difficulties with mathematics later in school promising.

References

Bailey, D. H., Littlefield, A., & Geary, D. C. (2012). The codevelopment of skill at and preference for use of retrieval-based processes for solving addition problems: Individual and sex differences from first to sixth grades. *Journal of Experimental Child Psychology, 113*(1), 78–92.

Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral & Brain Sciences, 11*, 169–232.

Bremholm, J., & Skott, C. K. (2019). Teacher planning in a learning outcome perspective: A multiple case study of mathematics and L1 Danish teachers. *Acta Didactica Norge*, *13*(1), 1–22.

Browne, K. R. (2002). *Biology at work: Rethinking sexual equality*. New Brunswick, NJ: Rutgers University Press.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology, 34*, 882–891.

Carpenter, T. & Moser, J. (1984). The Acquisition of Addition and Subtraction Concepts in Grades One through Three. *Journal for Research in Mathematics Education*, *15*(3), 179–202.

Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology, 103*(3), 617–631. doi: 2048/10.1037/a0023864

Carr, M., & Davis, H. (2001). Gender differences in arithmetic strategy use: A function of skill and preference. *Contemporary Educational Psychology, 26*(3), 330–347. doi:10.1006/ceps.2000.1059

Carr, M., & Jessup, D. L. (1997). Gender differences in first-grade mathematics strategy use: Social and metacognitive influences. *Journal of Educational Psychology*, *89*(2), 318–328.

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, *18*(1), 61–75.

Casey, M. B., Nuttall, R., Pezaris, E., & Benbow, C. P. (1995). The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, *31*(4), 697–705.

Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior & Development, 23,* 113–118.

Dowker, A. (2003). Young children's estimates for addition: The zone of partial knowledge and understanding. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructive adaptive expertise*. (pp. 243–266). Mahwah, NJ: Erlbaum.

Dowker, A. (2005). Early Intervention for students with mathematical difficulties. *Journal of Learning Disabilities*, *38*(4) 324–332.

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development*, *24*(4), 401–410. doi:10.1016/j.cogdev.2009.09.005

Dowker, A. (2014). Young children's use of derived fact strategies for addition and subtraction. *Frontiers in Human Neuroscience, 8,* 1–9. doi:10.3389/fnhum.2013.00924

Else-Quest, N., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin, 136*(1), 103–127. http://dx.doi.org/10.1037/a0018053 . Retrieved 20.10.17

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher, 27*(5), 6–11.

Foley, A. E., Vasilyeva, M., & Laski, E. V. (2016). Children's use of decomposition strategies mediates the visuospatial memory and arithmetic accuracy relation. *British Journal of Developmental Psychology, 35*(2), 303–309. doi:10.1111/bjdp.12166

Geary, D. C. (2010). *Male, female: The evolution of human sex differences.* Washington, DC: American Psychological Association.

Geary, D. C., Bow-Thomas, C. C., Liu, F., & Siegler, R. S. (1996). Development of arithmetical competencies in Chinese and American children: Influence of age, language, and schooling. *Child Development, 67*, 2022–2044.

Geary, D. C., Hoard, M. K., & Nugent, L. (2012). Independent contributions of the central executive, intelligence, and in-class attentive behavior to developmental change in the strategies used to solve addition problems. *Journal of Experimental Child Psychology, 113*(1), 49–65. doi:10.1016/j.jecp.2012.03.003

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38*(4), 293–304.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1–51.

Hirnstein, M., Andrews, L. C., & Hausmann, M. (2014). Gender-stereotyping and cognitive sex differences in mixed-and same-sex groups. *Archives of sexual behavior, 43*(8), 1663-1673.

Hornburg, C. B., Rieber, M. L., & McNeil, N. M. (2017). An integrative data analysis of gender differences in children's understanding of mathematical equivalence. *Journal of Experimental Child Psychology, 163*, 140–150.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581–592. doi:10.1037/0003-066X.60.6.581

Hyde, J. S., Fennema, E., & Lamon, S. J. (1990). Gender differences in mathematics performance: A meta-analysis. *Psychological Bulletin, 107*(2), 139–155. doi: 10.1037/0033-2909.107.2.139

Hyde, J. S., & Mertz, J. E. (2009). Gender, culture, and mathematics performance. *Proceedings of the National Academy of Sciences, 106*(22), 8801–8807. doi:10.1073/pnas.0901265106

Imbo, I., & Vandierendonck, A. (2007). The development of strategy use in elementary school children: Working memory and individual differences. *Journal of Experimental Child Psychology, 96*(4), 284–309. doi:10.1016/j.jecp.2006.09.001

Laski, E. V., Casey, B. M., Yu, Q., Dulaney, A., Heyman, M., & Dearing, E. (2013). Spatial skills as a predictor of first grade girls' use of higher level arithmetic strategies. *Learning & Individual Differences, 23*, 123–130. doi:10.1016/j.lindif.2012.08.001

Laski, E. V., Ermakova, A., & Vasilyeva, M. (2014). Early use of decomposition for addition and its relation to base-10 knowledge. *Journal of Applied Developmental Psychology, 35*(5), 444–454.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35*(4), 940–949.

Littell, R. C., Stroup, W. W., Milliken, G. A., Wolfinger, R. D., & Schabenberger, O. (2006). *SAS for mixed models.* SAS institute.

Maccoby, E. E., & Jacklin, C. N. (1974). *Psychology of sex differences.* Stanford, CA: Stanford University Press.

Moè, A. (2018). Mental rotation and mathematics: Gender-stereotyped beliefs and relationships in primary school children. *Learning and Individual Differences, 61*, 172–180.

Niss, M., & Højgaard, T. (2011). *Competencies and mathematical learning: Ideas and inspiration for the development of teaching and learning in Denmark* (IMFUFA tekst). Roskilde: Roskilde University.

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology, 67*(3), 345–357.

Paul, J. M., & Reeve, R. A. (2016). Relationship between single digit addition strategies and working memory reflects general reasoning sophistication. *Learning and Instruction, 42*, 113–122.

Penner, A. M., & Paret, M. (2008). Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research, 37*(1), 239–253. doi:http://dx.doi.org/10.1016/j.ssresearch.2007.06.012

Peterson, P. L., & Fennema, E. (1985). Effective teaching, student engagement in classroom activities, and sex-related differences in learning mathematics. *American Educational Research Journal, 22*(3), 309–335.

Pinker, S. (2002). *The blank slate: The modern denial of human nature.* New York: Viking.

Pinxten, M., Marsh, H. W., De Fraine, B., Van Den Noortgate, W., & Van Damme, J. (2014). Enjoying mathematics or feeling competent in mathematics? Reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, *84*(1), 152–174.

Price, G. R., Mazzocco, M. M., & Ansari, D. (2013). Why mental arithmetic counts: Brain activation during single digit arithmetic predicts high school math scores. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(1), 156–163. doi:10.1523/JNEUROSCI.2936-12.2013

Punch, S. (2002) Interviewing strategies with young people: the 'secret box', stimulus material and task based activities. *Children and Society 16*(1): 45–56.

Royer, J. M., Tronsky, L. N., Chan, Y., Jackson, S. J., & Marchant, H. III, (1999). Math-fact retrieval as the cognitive mechanism underlying gender differences in math test performance. *Contemporary Educational Psychology*, *24*(3), 181–266.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. New York: Routledge.

Selter, C. (2001). Addition and subtraction of three-digit numbers: German elementary children's success, methods and strategies. *Educational Studies in Mathematics*, *47*(2), 145–173.

Shen, C., Vasilyeva, M., & Laski, E. V. (2016). Here, but not there: Cross-national variability of gender effects in arithmetic. *Journal of Experimental Child Psychology*, *146*, 50–65. doi:10.1016/j.jecp.2016.01.016

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, *59*, 833–851.

Skott, J. (2004). The forced autonomy of mathematics teachers. *Educational Studies in Mathematics*, *55*(1–3), 227–257.

Spelke, E. S. (2005). Sex differences in intrinsic aptitude for mathematics and science?: a critical review. *American Psychologist*, *60*(9), 950–958.

Torbeyns, J., Verschaffel, L., & Ghesquière, P. (2005). Simple addition strategies in a first-grade class with multiple strategy instruction. *Cognition and Instruction*, *23*(1), 1–21. Retrieved 21.4.18 from http://dx.doi.org/10.1207/s1532690xci2301_1

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174–1204.

Table 1. Correlation matrix (Pearson's r) of proportional use (logit-transformed) of the five strategy use types and the combined category, counting (count all + count on), and the four derived principal component axes (PRIN1-4).

(N=63)	Error	Count all	Count on	Counting	Direct retr.	Derived fact
Error	-	0.17	-0.25*	-0.17	-0.19	-0.13
Count all	-	-	-0.66****	0.56****	-0.47****	-0.66****
Count on	-	-	-	0.16	-0.11	0.07
Counting	-	-	-	-	-0.79****	-0.8****
Direct retr.	-	-	-	-	-	0.53****
PRIN1	-0.42***	-0.91****	0.51****	-0.65****	0.67****	0.78****
PRIN2	0.2	0.24	-0.86****	-0.61****	0.56****	0.35**
PRIN3	0.88****	-0.31*	0.1	-0.41***	-0.01	0.21
PRIN4	0.03	0.02	-0.01	-0.05	0.51****	-0.43***

Levels of statistical significance: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; ****: $p < 0.0001$.

Table 2. Correlations (Pearson's r, n = number of students) between test variables of strategy use and general mathematical achievement in early year one and the students' test scores on different fields of mathematic achievement in year four. Correlations coefficients with numerical values ≥ 0.45 ($r^2 > 20\%$) are boldfaced.

		WP			NUMAR			FRAC		
		r	n	sign	r	n	sign	r	n	sign
Error	All:	-0.25	61		-0.37	63	**	-0.16	34	
	Girls:	-0.36	35	*	-0.36	36	*	-0.18	21	
	Boys:	-0.08	26		-0.40	27	*	-0.12	13	
Count all	All:	**-0.45**	**61**	*******	-0.24	63		**-0.47**	**34**	******
	Girls:	-0.33	35		-0.19	36		-0.18	21	
	Boys:	**-0.61**	**26**	*******	-0.18	27		**-0.79**	**13**	******
Count on	All:	0.37	61	**	0.09	63		0.22	34	
	Girls:	0.40	35	*	0.26	36		0.07	21	
	Boys:	0.36	26		-0.03	27		0.49	13	
Counting	All:	-0.16	61		-0.22	63		-0.31	34	
	Girls:	0.19	35		0.15	36		0.11	21	
	Boys:	-0.40	26	*	-0.18	27		-0.51	13	
Direct retrieval	All:	0.13	61		0.28	63	*	0.23	34	
	Girls:	-0.10	35		-0.03	36		-0.21	21	
	Boys:	0.32	26		0.33	27		0.55	13	*
Derived fact	All:	0.31	61	*	0.29	63	*	0.39	34	*
	Girls:	0.20	35		0.22	36		0.34	21	
	Boys:	0.42	26	*	0.09	27		0.45	13	
PRIN1	All:	0.44	61	***	0.36	63	**	**0.45**	**34**	******
	Girls:	0.35	35	*	0.28	36		0.17	21	
	Boys:	0.56	26	**	0.29	27		**0.74**	**13**	******
PRIN2	All:	-0.21	61		0.06	63		-0.05	34	
	Girls:	-0.41	35	*	-0.24	36		-0.10	21	
	Boys:	-0.01	26		0.14	27		-0.06	13	
PRIN3	All:	-0.01	61		-0.25	63		0.08	34	
	Girls:	-0.17	35		-0.24	36		-0.08	21	
	Boys:	0.18	26		-0.32	27		0.29	13	
PRIN4	All:	-0.12	61		-0.02	63		-0.17	34	
	Girls:	-0.19	35		-0.18	36		-0.44	21	*
	Boys:	-0.01	26		0.24	27		0.22	13	
MAT1	All:	-0.42	58	**	-0.40	60	**	-0.25	31	
	Girls:	**-0.48**	**32**	******	**-0.52**	**33**	******	-0.20	18	
	Boys:	-0.34	26		-0.39	27	*	-0.34	13	

Levels of statistical significance: *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; ****: $p < 0.0001$.

Table 3. Statistical model to predict values and amount of explained variation of WP in year four as multiple regression functions of test values of strategy use and mathematical achievement in early year one.

		Type 1 - effects					Partial (Type-3) effects:					Partial effects (Class as random effect)				
Model§	Variable†	F	df	p	ΔR^2	R^2	b	st,b	SE(b)	t	p	b	SE(b)	df	t	p
A	Intercept						10.7	.	1.04			10.7	1.44	5		
	MAT1*	11.19	1,52	**0.002**	17%	17%	-0.09	-0.27	0.04	-2.02	**0.049**	-0.09	0.03	47	-2.45	**0.018**
	PRIN1*	2.73	1,52	0.11	9%	26%	0.92	0.32	0.37	2.48	**0.017**	0.42	0.30	47	1.40	0.17
	PRIN2*	0.11	1,52	0.74	5%	31%	-0.87	-0.23	0.44	-1.98	0.053	-0.13	0.36	47	-0.37	0.71
	PRIN3*	2.55	1,52	0.12	1%	32%	0.40	0.09	0.55	0.72	0.47	0.63	0.41	47	1.53	0.13
	PRIN4*	0.23	1,52	0.63	0%	32%	-0.26	-0.04	0.79	-0.32	0.75	-0.30	0.61	47	-0.48	0.63
	*Class ID***	*9.7*	*5,47*	*<.0001*	*35%*	*67%*										
B	Intercept						9.22	.	1.19			9.45	1.47	5		
	MAT1*	13.79	1,55	**0.0005**	17%	17%	-0.09	-0.28	0.04	-2.29	**0.026**	-0.07	0.03	50	-2.34	**0.023**
	Count all	10.44	1,55	**0.002**	13%	31%	-1.98	-0.39	0.61	-3.23	**0.002**	-1.19	0.50	50	-2.40	**0.020**
C	Intercept						9.22	.	1.19			9.45	1.47	5		
	Count all	18.98	1,55	**<.0001**	24%	24%	-1.98	-0.39	0.61	-3.23	**0.002**	-1.19	0.50	50	-2.40	**0.020**
	MAT1	5.24	1,55	**0.026**	7%	31%	-0.09	-0.28	0.04	-2.29	**0.026**	-0.07	0.03	50	-2.34	**0.023**

l§ Model A: Model comprised only by variables forced into the model (no step-wise selection procedure), B: Model established with a step-wise selection procedure with MAT1 forced into model, C: Model established with step-wise selection procedure (sequence of variables show the order by which predictors were selected.
†Predictors eligible for model in step-wise selection procedures: MAT1, error, count all, count on, direct retrieval, derived fact, PRIN1, PRIN2, PRIN3, PRIN4 and sex.
*: Predictor forced into model
**: Type-1 effect of class identity if added to model A

Table 4. Statistical models to predict values and amount of explained variation of Number and arithmetic (NUMAR) in year four as multiple regression functions of strategy use and mathematical achievement in early year one.

		Type 1 – effects					Partial (type-3) effects:					Partial effects (Class as random effect)				
Model l§	Variable†	F	df	P	ΔR^2	R^2	b	st,b	SE(b)	t	p	b	SE(b)	df	t	p
	Intercept						11.81	0.00	1.27			11.48	1.45	5		
A	MAT1*	11.52	1,54	**0.0013**	16%	16%	-0.09	-0.24	0.05	-1.73	0.09	-0.08	0.05	49	-1.42	0.16
	PRIN1*	3.34	1,54	0.07	5%	21%	0.90	0.27	0.45	1.99	0.052	0.91	0.46	49	1.99	0.052
	PRIN2*	0.3	1,54	0.59	0%	21%	0.36	0.08	0.55	0.66	0.51	0.51	0.54	49	0.93	0.36
	PRIN3*	2.16	1,54	0.15	3%	24%	-0.96	-0.18	0.65	-1.47	0.15	-0.71	0.62	49	-1.15	0.26
	PRIN4*	0.01	1,54	0.92	0%	24%	-0.10	-0.01	0.97	-0.10	0.92	-0.38	0.95	49	-0.4	0.69
	*Class***	*2.45*	*5,49*	***0.047***	*15%*	*39%*										
B,C	Intercept						15.12		1.27			14.66	1.41	5		
	MAT1*	12.81	1,57	**0.0007**	16%	16%	-0.16	-0.42	0.04	-3.69	**0.0005**	-0.14	0.04	52	-3.38	**0.0014**
	Sex (F)	9.49	1,57	**0.0032**	12%	28%	-3.56	-0.35	1.15	-3.08	**0.0032**	-3.42	1.08	52	-3.16	**0.0026**
D	Intercept						13.1		1.15			12.7	1.34	5		
	MAT1	11.17	1,58	**0.0015**	16%	16%	-0.15	-0.40	0.046	-3.34	**0.0015**	-0.14	0.046	53	-3.02	**0.0039**

l§ Model A: Model comprised only by variables forced into the model (no step-wise selection procedure), B: Model established with a step-wise selection procedure with MAT1 forced into model, C: Model established with step-wise selection procedure (sequence of variables show the order by which predictors were selected, D: Similar to C, excepting that sex is not available for inclusion in model.

†Predictors eligible for model in step-wise selection procedures: MAT1, error, count all, count on, direct retrieval, derived fact, PRIN1, PRIN2, PRIN3, PRIN4 and sex.

*: Predictor forced into model

**: Type-1 effect of class identity if added to model A

Table 5. Statistical models to predict values and amount of explained variation of FRAC in year four as multiple regression functions of strategy use and mathematical achievement in early year one.

		Type 1 - effects					Partial (type-3) effects:					Partial effects (Class as random effect)				
Model	Variable†	F	df	p	ΔR^2	R^2	b	st,b	SE(b)	t	p	b	SE(b)	df	t	p
A	Intercept						7.97		1.49			8.06	1.55	2		
	MAT1*	2.04	1,25	0.17	6%	6%	-0.02	-0.06	0.06	-0.26	0.80	-0.02	0.06	23	-0.34	0.74
	PRIN1*	5.13	1,25	**0.03**	16%	22%	1.13	0.43	0.56	2.01	0.06	1.07	0.58	23	1.84	0.08
	PRIN2*	0	1,25	0.95	0%	22%	-0.01	0.00	0.72	-0.01	0.99	0.06	0.72	23	0.09	0.93
	PRIN3*	0.18	1,25	0.68	1%	23%	0.32	0.08	0.80	0.40	0.70	0.31	0.80	23	0.39	0.70
	PRIN4*	0.07	1,25	0.80	0%	23%	-0.32	-0.05	1.22	-0.26	0.80	-0.25	1.22	23	-0.21	0.84
	*Class***	1.31	2,23	0.29	8%	31%										
B	Intercept						6.15		1.59			6.15	1.59	2		
	MAT1*	2.41	1,28	0.13	6%	6%	-0.01	-0.05	0.05	-0.26	0.79	-0.01	0.05	26	-0.26	0.79
	Count all	7.94	1,28	**0.009**	21%	27%	-2.34	-0.50	0.83	-2.82	**0.009**	-2.34	0.83	26	-2.82	**0.009**
C	Intercept						5.79		0.80	7.25		5.99	0.77	2		
	Count all	9.32	1,32	**0.005**	0.23	0.23	-2.43	-0.52	0.75	-3.26	**0.003**	-2.14	0.70	30	-3.05	**0.005**
D	Intercept						3.88		1.30			3.89	1.30	2		
	Count all	11.18	1,30	**0.002**	23%	23%	-4.52	-1.00	1.05	-4.31	**0.0002**	-4.52	1.05	28	-4.31	**0.0002**
	Sex	0.00	1,30	0.98	0%	23%	2.91	0.36	1.55	1.88	0.07	2.91	1.55	28	1.89	0.07
	Count all*Sex	8.38	1,30	**0.007**	17%	39%	3.87	0.71	1.34	2.89	**0.007**	3.87	1.34	28	2.89	**0.007**
	(combined)	4.19	2,30	**0.02**	17%											

l§ Model A: Model comprised only by variables forced into the model (no step-wise selection procedure), B: Model established with a step-wise selection procedure with MAT1 forced into model, C: Model established with step-wise selection procedure, D: Modell C with statistically significant sex interaction added.

†Predictors eligible for model in step-wise selection procedures: MAT1, error, count all, count on, logit-direct retrieval, derived fact, PRIN1, PRIN2, PRIN3, PRIN4 and sex.

*: Predictor forced into model

**: Type-1 effect of class identity if added to model A

Figure 1. Distribution of the frequency by which different strategies were used to solve the 36 single-digit addition items, distributed on the year one's first (November) and second half (April) and on sex.

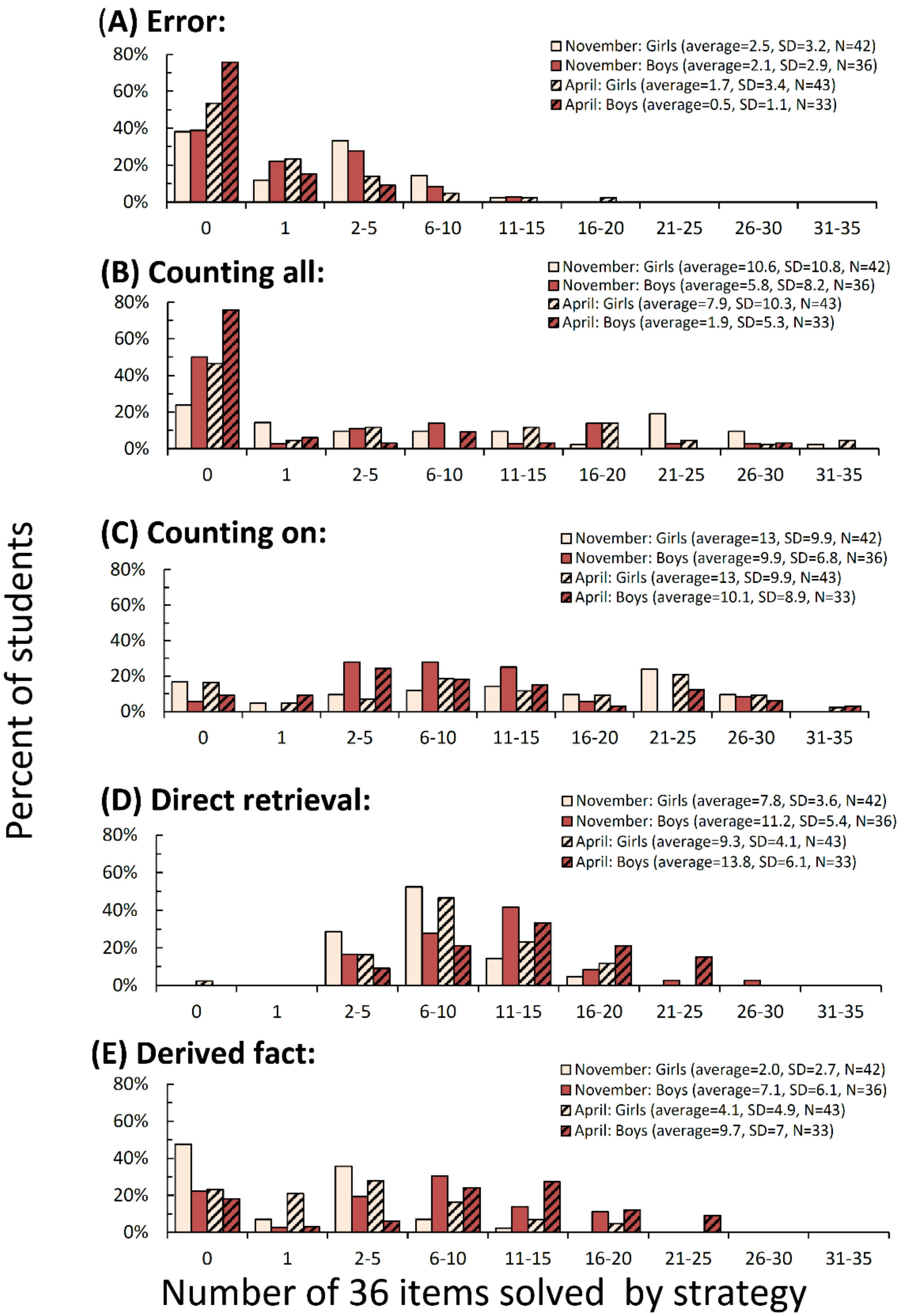


Figure 2. Test score achieved by students in year four within problem type WP plotted as a function of how much they used the arithmetic strategy use count all when tested in early year one. The thick line show the linear regression function (y = 6.9 [SE_{59}=0.68] – 2.26 [SE_{59}=0.58, P = 0.0003]x) with thin lines indicating 95% confidence zones. To separate data points with similar values in the graph all observations are provided with slight random scatter in vertical as well as horizontal direction.

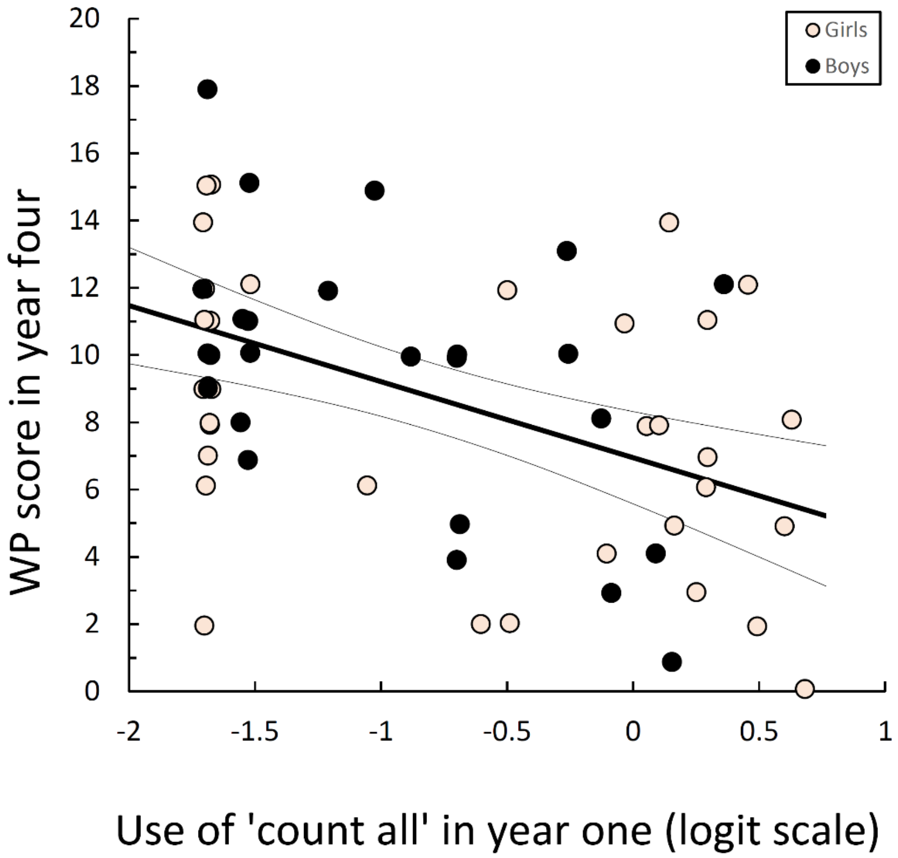


Figure 3 Test score achieved by students in year four within problem type NUMAR plotted as function of their score (number of errors) in the general mathematical achievement test MAT1 in early year one. The thick lines show the multiple linear regression function (model B and C) described in Table 4 that explained the mean score as main effects of sex and MAT1-score. Thin lines demarcates 95% confidence zones. To separate data points with similar values in the graph all observations are provided with slight random scatter in vertical as well as horizontal direction.

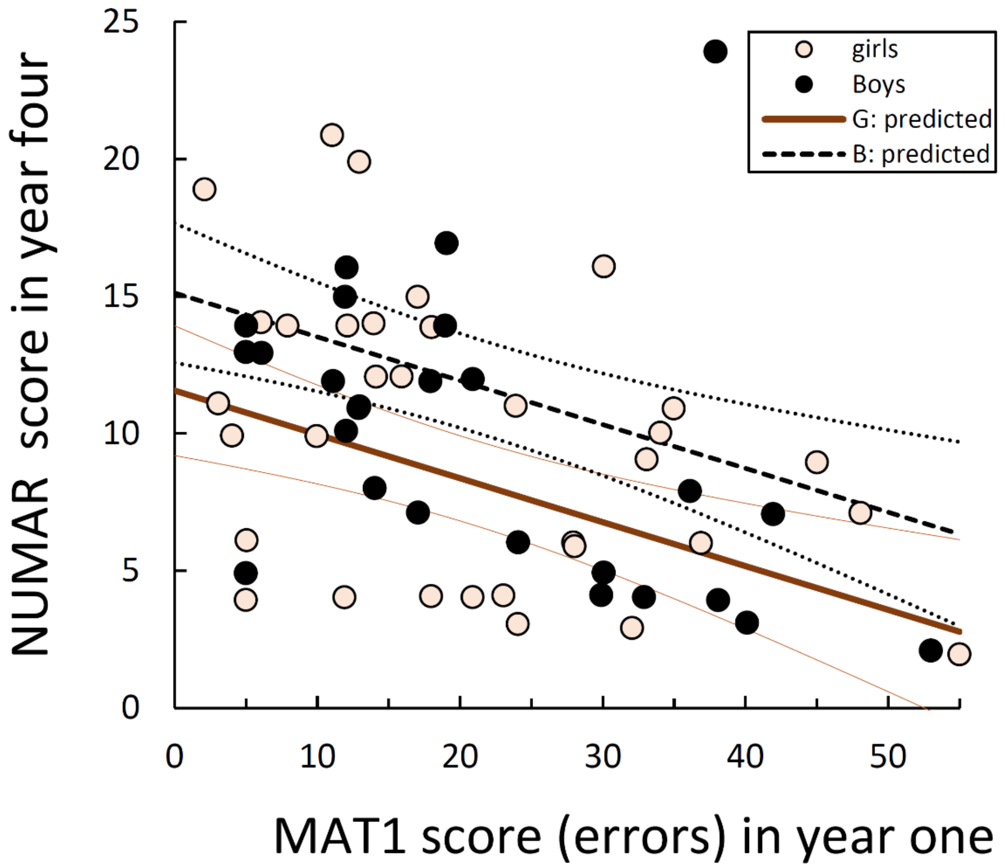
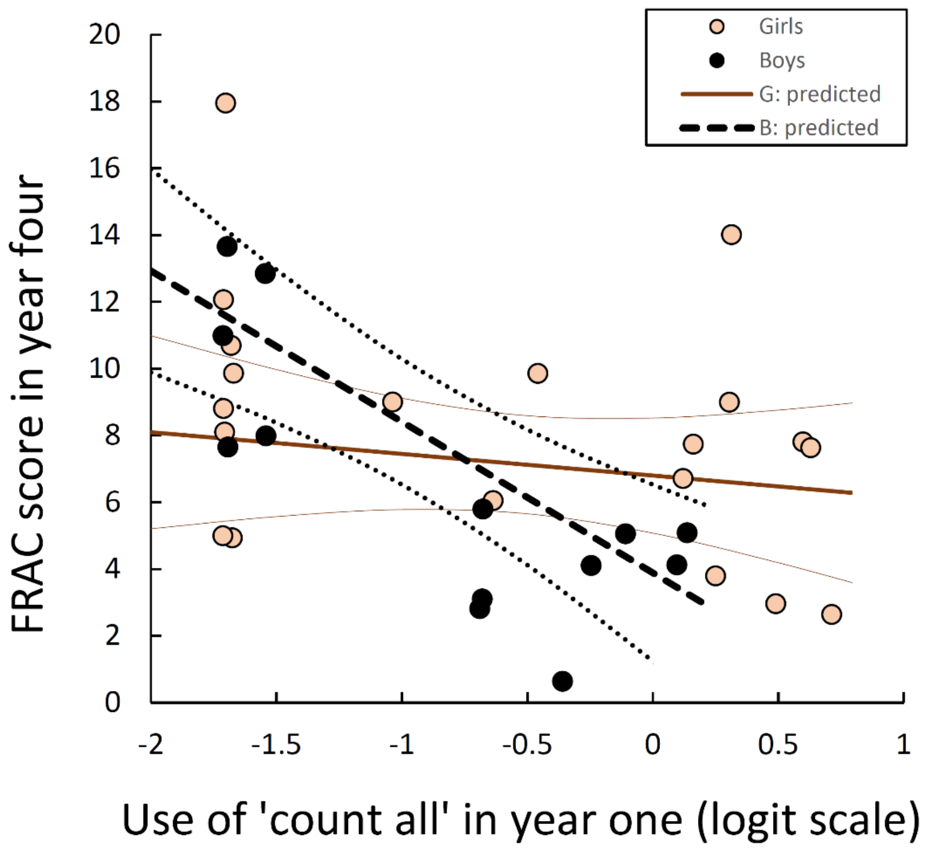


Figure 4. Test score achieved by students in year four within problem type FRAC plotted as function of how much they used the arithmetic strategy use count all when tested in early year one. The thick lines show the multiple linear regression function described in Table 4 (model D) that explained the mean score as an interactive function of sex and MAT1-score and strategy use in year one. Thin lines demarcates 95% confidence zones. To separate data points with similar values in the graph all observations are provided with slight random scatter in vertical as well as horizontal direction.



Paper IV

Danish teachers' expectations of year one pupils' additive competence

Pernille B. Sunde[1] and Judy M. Sayers[2]

[1]Aarhus University, Department of Education, Denmark; pebs@edu.au.dk

[2]Stockholm University, Mathematics and Science Education Department, Sweden; judy.sayers@mnd.su.se

In this paper we present an investigation of six Danish teachers' expectations for their first grade students within the subject addition. We asked the teachers what the students should know by the end of grade one. A constant comparison analysis revealed three themes: curriculum objectives, mathematical content, and differentiation. The teachers' expectations ranged from one teacher stating that most students should be able to do single-digit addition, but more importantly know the everyday language of addition, to a teacher requiring all students to be proficient in two-digit addition without bridging ten, and 20% of their pupils should be able to manipulate two-digit addition bridging ten operations. The results are discussed in relation to the common objectives of the Danish curriculum and the general findings of relationships between early proficiency and later achievements.

Keywords: Teacher expectations, addition, grade one objectives.

Introduction.

In this paper, we draw upon an interview study of six teachers' views and expectations of their grade one classes by the end of the academic year. In particular, this paper investigates what six Danish teachers consider as important learning objectives in early arithmetical learning of whole number addition. At present, mathematics education finds itself at the intersection of different trends in relation to curriculum objectives and curriculum programmes, brought about by international achievement comparisons (Remillard, 2005). Our paper investigates what do teachers expect their students to learn. The results of a constant comparison analysis revealed three key themes: curriculum objectives, mathematical content, and differentiation. We discuss these findings and their importance in relation to the requirements of the Danish curriculum and existing theory.

The Danish school system

The Danish public school, Folkeskolen (Danish Ministry of Education, 2014), consists of 10 years of compulsory education beginning with grade 0 (a preschool class) in august, the calendar year the child turns six. Thus, by the end of first grade, the students will have received two years of formal schooling. The Danish curriculum is constructed of Common Objectives which consist of: 1) the binding national objectives in the form of the subject aims, central knowledge and skill areas (end objectives) and form level (phase) objectives, and 2) guidelines for curricula and descriptions of the educational development designed for reaching the form level and end objectives. The Common Objectives are not specified on grade level as such, but operate with learning objectives for a three-year period specifying the knowledge and skills the student should acquire by the end of grade 3, 6 and by the end of compulsory education by grade 9. For each three-year period the learning objectives are divided in three (or two) phases or levels to specify or clarify the students' development in the specific topic. Furthermore, each 'phase learning objective' consist of a skill and a knowledge objective. The municipality can set additional objectives, and the individual school and teacher has a high degree of freedom in deciding how to implement and ensure the fulfilment of the goals. The Common Objectives build on a long and important Danish tradition of "metodefrihed", which suggest an individual teacher has high levels of autonomy (Klette, 2002) and is free to decide how and what to teach when. What is legislated is that the Common Objectives are met at the end of the three year intervals. The focus of this paper are the Common Objectives for Number and Algebra at the end of third grade. These objectives are broken down into broad phases: 1, 2 and 3, for grade 3, suggesting learning is constructed through instruction of skills and knowledge in relation to: number, calculation strategies and algebra. The implication is the existence of a learning pathway to guide teachers in their planning and expectations. Yet not to expect all children necessarily to achieve phase one in year 1, phase two in year 2 et cetera.

Theoretical background

Over many years, research has informed us of how the use of learning pathways (teaching-learning path) developed through intervention projects (inter alia: Carpenter, Fennema & Franke, 1996), have added to our knowledge and understanding of the development of number in early mathematics. Such pathways have been described as '*learning trajectories'*, '*instructional sequences'*, '*building blocks'*, (Sarama & Clements, 2004, 2009), '*mental structures'* (Gelman & Brenneman, 2004), '*growth points'* (Gervasoni, 2005), and '*learning frameworks'* (Wright, 2003). Such pathways can serve to connect research with practice and provide guidance for teachers in their instruction in a sequential way (Kühne, Lombrad & Moodley, 2013). Pathways can inform developmental progression in learning mathematics, offering teachers suggestions of rich environments that are developmental and effective (Bobis, Clarke, Clarke, Thomas, Wright & Gould, 2005). As an example, Cranfield, Kühne, van den Heuvel-Panhuizen, Ensor, Lombard, and Powell (2005), drew from Steffe's (1992) original framework to provide a 'trajectory for learning and how learner's learn, understand and solve number problems'(p274). Four stages are recognized: 1) Pre-school – grade R, progress through Emergent Numeracy. 2) Grade R-1, progress in learning to count-and-calculate and Integrated Mental and Written (operations up to 10). 3) grades 1-2, Calculation by structuring integrated mental and written operations up to 20 and beyond, and finally, 4) grades 3-4 progress through Formal Calculation Mental & Written (Operations up to 100 and beyond) (Cranfield et al., 2005, p. 274).

According to Rittle-Johnson and Schneider (2015), there appears to be a general consensus that mathematical competence rests on the development of knowledge of concepts and of procedures, and that the relationship between a conceptual and procedural knowledge is often bi-directional and iterative. In other words, with an increase of conceptual knowledge, leads to increased knowledge of procedural knowledge, and vice versa (Baroody, 2003). Positive correlations can be found in previous studies in early development of mathematics for example, counting (LeFevre, Clarke & Stringer, 2002), addition and subtraction (Jordan, Kaplan, Ramineni, & Locuniak, 2009), fractions and decimals (Hallett, Nunes, & Bryant, 2010), estimation (Dowker, 2003), and equation solving (Rittle-Johnson, Star, & Dunkin, 2012). Suggesting that the strength of the relationship between conceptual and procedural knowledge is fairly high. In sum, arithmetical competence relies on several related numerical components in order that young children up to grade 3, are able to solve arithmetical problems (Fuson & Burghardt, 2003). Desoete & Gregoire's, (2006) work focussed on five components underlying numerical competence: Logical knowledge, counting procedure, knowledge of the numerical system, representation of the number size and computation (p. 352) which illustrates the above notions of integration of content, concepts and procedures. In the following, we examine six Danish teachers' expectations of their grade one students and the extent to which this conforms with the above mentioned elements, that is the official curriculum, mathematical content, and procedural vs. conceptual knowledge.

Methodology.

The project draws on a Case study paradigm to address the research questions, as Case study allows us to explore in-depth 'what', 'how' and 'why' teachers teach in the ways they do in their natural setting (Yin, 2013). To this end, an exploratory case study was undertaken to examine six grade one teachers' perspectives on, and justifications for, what arithmetical concepts and strategies they expect children to learn about during this first year of formal schooling. We adopted a constant comparison analytical approach, drawn from grounded theory (Strauss & Corbin, 1998), which is commonly used in case study as it facilitates the thick description expected of case study analyses of complex educational settings (Yin, 2013). The study presented here is the initial results of a larger mixed methods case study, where classroom lesson observations, clinical interviews with grade one pupils were also undertaken, after the initial teacher interviews illustrated here. This first stage of the larger study identified three key themes in how the teachers thought and articulated their views on their pupils' additive competences. Their views related to: curriculum objectives, mathematical content and differentiation.

This qualitative study drew from teacher interviews from a purpose sampling, intended to avoid difficulties that random selection can produce. For example, the teachers had all been trained as primary teachers, 3 were mathematics specialists, 3 were male teachers, and a cross section of experience (2 – 24 yrs) and aged between 30-49 yrs old, was chosen to provide a representation of a broader selection, if that had been possible. Pseudonyms have been used throughout. The teacher participants were interviewed in October 2015 about their teaching of arithmetic in first grade, their choices of tasks, how they assess the students, the prerequisites for learning specific mathematical topics. As with most case studies, we recognise that the results presented here are in no way generalisable, but merely give an indication and representation of situations such as the actual case studied (Taylor & Bogdan, 1998).

In this paper we will analyse the teachers responses to the questions on what they considered their first grade students should know or be able to within the subject of addition by the end of first grade. The interviewer made no references to the Danish Common Objectives. All excerpts used below were translated into English, a process that included transforming

Danish idioms into forms that were recognisable to an English-speaker without losing the speaker's intended meaning (Kvale & Brinkmann, 2009).

Results

Three themes of the teachers' expectations for their first grade students emerged from the interview: Curriculum objectives, mathematical content, and differentiation. The theme c*urriculum objectives* consists of two categories: 1) references to the official Common Objectives and the use of goal related language like "objective", "goal" or "teaching plan" (DA årsplan). The *mathematical content* of the teachers' expectations include digit range, bridging ten, calculation methods and conceptual understanding. Digit range refers to addition with one- or two-digit numbers, which is closely linked to notions of bridging ten. Bridging ten is also an important part of different calculation methods. The last category is conceptual understanding and connections to real-life experiences. The last theme, *differentiation,* refers to the teachers' different expectations for different student groups, challenged students as well as high achievers.

In the following, we will present the teachers notions with examples from the interviews.

Allan (male, age 37, experience 2 yrs.) referred, as the only one, explicitly to official curriculum objectives: "The students should fulfil the common objectives". Furthermore, he had transferred some of the goals to his teaching plan. Regarding mathematical content, he used the expression doing addition with "numbers to hundred" but did not mention bridging ten operations or elaborate on calculation methods. Allan mentioned both high and low achieving student groups, but he focused on whether or not they will fulfil the goals: "some of them will have difficulties with fulfilling all of the things in this plan" and regarding high achievers he said "some of them has fulfilled the goals, and they just get some extra". He elaborates on giving the students extra work "fortunately it is those that are fast, they are also good at it, and often children like to do what they are good at".

Bettina (female, age 49, experience 24 yrs.) made an implicit reference to the curriculum objectives by saying that "there are rules for that I guess". She also referred to the students' own book with the learning goals (an interactive book on the students' ipads). Bettina specified that students should be able to operate with two-digit numbers, and 20 % of the students should be able to manage bridging ten operations". She implicitly referred to calculation methods by saying: "they should understand that we take the ones and then we take the tens and then we find out what that is all together and all that". Bettina expressed increased expectations for the high achievers and concern on how to challenge these students "they (referring to three specific boys and bridging ten) can already manage and they could easily have been challenged in class today".

Carl (male, age 40, experience 15 yrs.) used the words "my goal (for the students)" but he made no references to the official curriculum. Carl did not specify digit span directly, but implied two-digit addition through statements on working with base-ten and a written algorithm with bridging ten "the students should be proficient in some kind of bridging-ten-calculation-algorithm". Elaborating on different student groups, Carl described a conversation with a challenged student about her finger counting strategies and said, “it is about trying to push them on and push them out of the counting inflexibility”.

Dan (male, age 45, experience 20 yrs.) made no references to curriculum objectives. He specified that students should be able to operate with “two-digit plus one-digit bridging ten operations”. Elaborating on calculation methods he emphasized “mental calculation or by some kind of counting”. Dan elaborates on how he is aware if any students are ‘stuck’ on inadequate calculation strategies: “Then we talk about other ways of doing it … that there are other more smarter ways of doing it”. Dan, as the only one, mentioned bilingual students’ challenges.

Else (female, age 37, experience 9 yrs.) made no references to curriculum objectives or goal related language. Regarding mathematical content Else, did not know “how many digits they have to manage”. Conceptual understanding was only addressed by Else referring to the understanding of addition in an everyday context: “it is just as important that they know some of the words of everyday language. What does it mean, addition is a thing, that it is adding. That they know the symbol but also know the everyday language of what it means that we add, that we add things”. Else, referring only to the challenged students, said, “the wishful thinking would be that it was in first grade that all the students were proficient in single-digit addition, but I know that for many that will not be the case”.

Frida (female, age 30, experience 5 yrs.) did not mention curriculum objectives either. She did not specify digit range but said that the students should be able to do addition with “larger numbers”. Frida did not explicitly mention calculation methods but said, “I do not think it matters that they still use their fingers to count” but did not elaborate further. Frida did not relate expectations to different student groups.

In sum, three teachers, Bettina, Carl and Dan, articulated their expectations regarding the mathematical content very clearly. They also had the highest expectations regarding procedural knowledge compared to the other three teachers. However, the level of expectations of Allan and Frida was difficult to establish because of their unclear articulations. Else stands out from the others as being the only teacher referring to conceptual understanding and everyday language of addition, and at the same time having the lowest expectations regarding procedural knowledge, such as digit range.

A clear articulation on the expectations for specific mathematical content is apparently linked to notions on different student groups. The three teachers, Bettina, Carl and Dan, that had high expectations for their students also seemed to have focus on differences between student groups. They mentioned both challenged students and high achievers and talk about the students’ challenges and how to support their development. Especially Bettina and Carl had an awareness of the high achievers. Contrary, teachers with unclear or low expectations for the students seemed to be less aware of different student groups (Frida), focused on the challenged students (Else) or on how to keep students (high achievers) occupied (Allan).

Discussion

In this study, we have shown that the teachers’ expectations for their students are very different, from one teacher saying that not all students will be proficient in single-digit addition to another teacher saying that all students should be proficient in two-digit addition. This gives rise to two questions: what could cause such differences and what are the possible implications?

The causes to differences in expectations could be many, and it is beyond the possibilities for this paper to discuss in depth. However, we would like to point to the fact that the teachers in general do not make explicit references to the official Danish curriculum. On the other hand, the Danish Common Objectives do not specify explicitly when the students should be able to do addition with one or two digits or when they should be able to bridge ten. This is deliberately left to the individual teacher, which could explain the differences in the teachers' expectations. However, only one teacher explicitly mentions the Common Core Objectives, so we do not know to what extend the teachers in general look to the Common Objectives for guidance on goals for their students.

What are the possible implications for such differences in teachers' expectations? Early teacher expectations seem to have long lasting effect on individual students' later achievements and these effects are more pronounced for at risk students (Hinnant et al. 2009). This is especially pronounced for inaccurate expectations, that is the over- and underestimation of abilities, (Sorhagen, 2013). Thus, the existence of such a relationship makes the findings in this study important.

The teachers' expectations differ with what would be expected to be one or two years of learning according to different learning trajectories (e.g. Clements & Sarama, 2004). Thus, if teachers teach according to their expectations for the students, there is a risk that the first graders in this study will experience very different learning opportunities with different achievement levels by the end of grade one.

When considering the well-established link between early numerical skills and proficiency and later achievement in math (e.g. Jordan et al., 2009; Landerl, Bevan & Butterworth, 2004), the differences in expectations could cause some concerns both regarding the mathematical content, notions of different student groups and the focus on either procedural or conceptual knowledge.

Successful teaching of arithmetic should build on procedural as well as conceptual knowledge (Rittle-Johnson & Schneider, 2015) and be informed by knowledge from learning parthways (Kühne et al., 2013) in order to provide a rich learning environment (Bobis et al., 2005). As shown, the teachers in this study differed in their notions on these issues. However, the analysed excerpts in this paper cannot provide answers to how the teachers actually teach. We do not know whether they teach addition with procedural understanding (as rote learning) or with conceptual understanding (Rittle-Johnson & Schneider, 2015). As the project progresses, we hope that additional analysis of interviews and video observations from the classrooms will provide further insights into how these issues are manifested in, for example the teachers' choices of activities, a rational for their choices and the actual teaching of these activities.

References

Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: Constructing adaptive expertise* (pp.1-33).

Bobis, J., Clarke, B., Clarke, D., Thomas, G., Wright, B., Young-Loveridge, J., & Gould, P. (2005). Supporting teachers in the development of young children's mathematical thinking: Three large scale cases. *Mathematics Education Research Journal*, *16*(3), 27-57.

Carpenter, T. P., Fennema, E., & Franke, M. L. (1996). Cognitively guided instruction: A knowledge base for reform in primary mathematics instruction. *The Elementary School Journal*, *97*(1), 3-20.

Cranfield, C., Kühne, C., van den Heuvel-Panhuizen, M., Ensor, P., Lombard, A. P., & Powell, G. (2005). Exploring strategies used by grade 1 to 3 children through visual prompts, symbols and worded problems: a case for learning pathway for number. In Chick, H.L.& Vincent, J.L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education, 2*, 273-280.

Danish Ministry of Education, (2014). *Improving the Public School – overview of reform of standards in the Danish public school (primary and lower secondary education)*. Copenhagen: The Danish Ministry of Education. Accessed online 12Dec16: http://eng.uvm.dk/Education/Primary-and-lower-secondary-education/The-Folkeskole

Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and individual differences*, *16*(4), 351-367.

Dowker, A. (2003). Young children's estimates for addition: The zone of partial knowledge and understanding. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: constructive adaptive expertise* (pp. 243-266).

Fuson, K. C., & Burghardt, B. H. (2003). Multidigit addition and subtraction methods invented in small groups and teacher support of problem solving and reflection. In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise* (pp. 267-307).

Gelman, R., & Brenneman, K. (2004). Science learning pathways for young children. *Early Childhood Research Quarterly*, *19*(1), 150-158.

Gervasoni, A. (2005). The diverse learning needs of young children who were selected for an intervention program. In Chick, H. L. & Vincent, J. L. (Eds.). *Proceedings of the 29th Conference of the International Group for the Psychology of Mathematics Education, 3*, 57-64.

Hallett, D., Nunes, T., & Bryant, P. (2010). Individual differences in conceptual and procedural knowledge when learning fractions. *Journal of Educational Psychology*, *102*(2), 395.

Hinnant, J. B., O'Brien, M., & Ghazarian, S. R. (2009). The longitudinal relations of teacher expectations to achievement in the early school years. *Journal of educational psychology*, *101*(3), 662.

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology, 45*(3), 850-867. doi:10.1037/a0014939

Klette, K. (2002). Reform policy and teacher professionalism in four Nordic countries. *Journal of Educational Change, 3*(3-4), 265-282.

Kühne, C., Lombard, A. P., & Moodley, T. (2013). A learning pathway for whole numbers that informs mathematics teaching in the early years. *South African Journal of Childhood Education, 3*(2), 77-95.

Kvale, S. & Brinkmann, S. (2009). *Den kvalitativa forskningsintervjun*. Lund: Studentlitteratur.

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9-year-old students. *Cognition, 93*(2), 99-125. doi:http://dx.doi.org/10.1016/j.cognition.2003.11.004

LeFevre, J., Clarke, T., & Stringer, A. (2002). Influences of language and parental involvement on the development of counting skills: Comparisons of French- and English-speaking Canadian children. *Early Child Development and Care, 172*, 283–300

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211-246.

Rittle-Johnson, B., & Schneider, M. (2015). Developing Conceptual and Procedural Knowledge of Mathematics: In R. C. Kadosh & A. Dowker (Eds.), *Oxford Handbook of Numerical Cognition*, (pp. 1118-1134). Oxford: Oxford Uni Press.

Rittle-Johnson, B., Star, J. R., & Durkin, K. (2012). Developing procedural flexibility: Are novices prepared to learn from comparing procedures? *British Journal of Educational Psychology, 82*(3), 436-455.

Sarama, J., & Clements, D. H. (2004). Building blocks for early childhood mathematics. *Early Childhood Research Quarterly, 19*(1), 181-189.

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. Routledge.

Sorhagen, N. S. (2013). Early teacher expectations disproportionately affect poor children's high school performance. *Journal of Educational Psychology, 105*(2), 465.

Steffe, L. P. (1992). Schemes of action and operation involving composite units. *Learning and individual differences, 4*(3), 259-309.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Procedures and techniques for developing grounded theory*. Thousand Oaks: SAge

Taylor, S. J., & Bogdan, R. (1998). Working with data: data analysis in qualitative research. *Introduction to qualitative research methods, 3*, 134-163.

Wright, R. J. (2003). A mathematics recovery: Program of intervention in early number learning. *Australian Journal of Learning Difficulties, 8*(4), 6-11.

Yin, R. K. (2013). *Case study research: Design and methods*. Sage publications.

Paper V

Teaching addition: Teachers' perspectives on teaching and learning number and addition in year one

Pernille B. Sunde
VIA University College and Aarhus University, DPU

Abstract

This study investigates Danish teachers' perspectives on teaching and learning number and addition in year one through analyses of semi-structured interviews with six year one teachers. The teachers' perspectives is compared with theories on foundations for learning number and addition. The analysis indicates that important foundational factors, estimation and quantity discrimination, were either superficially discussed or not addressed at all by the teachers in their interviews. Furthermore, the individual teachers differed substantially with regard to their emphasis towards mental and algorithmic approach to computation.

Introduction

Learning and teaching of arithmetic has been a major topic for educational research for decades (Nunes et al., 2016) and the importance of students' arithmetic competence for mathematical development in general is widely acknowledged. Thus, numerical and arithmetic competence has been linked to development of mathematics achievement and difficulties (e.g. Feigenson, Libertus, & Halberda, 2013; Geary et al., 2013; Ostad, 1997a). Teaching of arithmetic in primary school relies on the implementation of key knowledge to set up strong foundations for a successful development of arithmetic competence, e.g. adaptive flexibility, strategies for mental calculation, and number knowledge. Students learn in many settings in and out of school for example through parents informal number talk in everyday activities (Chang et al., 2011) and when playing games (Bjorklund, Hubertz, & Reubens, 2004). However, the teacher plays a key role in providing opportunities for students to learn. Therefore, it is important to get insight into teachers' perspectives on the teaching of number and arithmetic in the early years of school, where number and basic arithmetic is the primary focus, and research suggests is a crucial stage in children's development of number competencies.

In this paper, I explore six Danish teachers' perspectives on the teaching and learning of number and addition and analyse the extent to which this is aligned with established knowledge on foundations of numerical and arithmetic competence.

Numerical and arithmetical competence

Development of arithmetical competence rely on several components of numerical competence or number sense (Desoete & Grégoire, 2006; Fuson & Burghardt, 2003), for example: symbolic knowledge and number words (e.g. Chu et al., 2015), mapping symbols to quantity (e.g. Geary, 2013), basic counting skills (e.g. Jordan et al., 2009), number comparison skills (e.g. De Smedt et al., 2013), estimation skills (e.g. Booth & Siegler, 2008; Gilmore et al., 2007), and knowledge of base-ten number structure (e.g. Laski et al., 2014).

There seem to be a consensus that a basic understanding of multidigit numbers and basic arithmetic is fundamental for success in multidigit arithmetic (e.g. Fuson, 1992;

Verschaffel et al., 2007). Hence Mazzocco, Murphy, Brown, Rinne, & Herold (2013) found that students's numbers of atypical errors in a written numbers task (write the smallest and largest one-, two- and thee-digit numbers) in year 2 and 3 predicted computational errors in year 8. Students with mathematics learning difficulties made significantly more infrequent computational errors at year 3 and the students' that in general made infrequent errors in year 3 (Students with mathematics learning difficulties as well as typical achievers and low achieving) made atypical place value errors (e.g. adding across tens and ones place) in year 8. Geary et al. (2013) found that number system knowledge measured as addition strategy choice, number sets test (determining wehter a set of objects or numerals equal a target number) and number line estimation beginning year one predicted functional numeracy (solving multistep word problems, whole number arithmetic, fraction arithmetic and comparing fractions) six years later.

The individual components of number sense are all important for further mathematical development, however, the links between them is essential for students' development (Gersten et al., 2005). Although all components have been shown to individually contribute to students' arithmetic ability, the combined contribution is complex and change with age (Lyons et al., 2014). In a longitudinal study, Lyons et al. (2014) found that the contribution of different basic numerical skills (matching number words and symbol, counting skill, numeral ordering, comparing dot arrays, symbolic comparison, matching sets of objects, estimation of dots in arrays, number-line estimation) to students arithmetic skills changed with age. They found that counting skill, numeral ordering, matching sets of objects, estimation of dots in arrays, and number-line estimation significantly interacted with grade, and number-line estimation was a strong unique predictor of arithmetic in year 1 to 2 whereas numeral ordering was a poor predictor of arithmetic ability in year 1. However, the predictive capacity of numeral ordering increased steadily until year 6 where it was the strongest predictor of arithmetic ability.

The quality and structure of the student's number knowledge and capabilities of reasoning with numbers has been found not only to influence arithmetic ability in general but also specific strategy use in mental arithmetic. For instance, number knowledge early in year one influence the use of derived fact strategies by the end of year one (Gaidoschik, 2012) and numerical (symbolic) magnitude processing skills are lower for students with persistent learning difficulties and these skills seem to be related to strategy use (Vanbinst et al., 2014).

The type of strategies used in solving mental addition and subtraction has been shown to influence development of mathematics competence (Carr et al., 2008; Cowan et al., 2011; Fennema et al., 1998) and be a valid predictor of later mathematical achievement and difficulties (Gersten et al., 2005; Ostad, 1997a; Price et al., 2013). Hence, Carr and Alexeev (2011) found that the use of cognitive strategies (calculations performed without manipulatives, including silent mental counting, derived fact, and direct retrieval) in year two was important for development of mathematics competency in year four, whereas excessive use of counting strategies using manipulatives resulted in poor outcomes on mathematics competences.

Mental and algorithmic computation

Proficiency in procedures for doing arithmetic in the four operations are the major focus of mathematics in elementary school curriculum (Kilpatrick et al., 2001). A key issue of teaching addition, and much debated, is what methods for computation students should master, the mental or the school-taught algorithmic approach (Kilpatrick et al., 2001; Verschaffel et al., 2007). The traditional algorithms consist of a series of well-defined and fixed steps. The steps operates on the digits and if the procedure is carried out correctly (and only then), a correct result is guaranteed. Algorithmic computation is typically performed with paper and pencil and tasks presented vertically. The mental computation is based on flexibly adapting the solution strategy to the specific numbers. That implies operating on numbers, not digits, and there is no single correct solution path. Therefore, correct result is not guaranteed. Mental computation is often done in the head, but written notations can occur. Tasks are often presented horizontally.

When learning multidigit arithmetic, students have been shown to experience difficulties in learning the multiple steps of the standard algorithm. This has been linked to lack of conceptual understanding (e.g. multidigit numbers and the 10-for-1 trading) as well as lack of procedural knowledge (Fuson, 1992; Verschaffel et al., 2007). The developmental relationship between conceptual understanding and procedural knowledge is iterative, that is bi-directional, an increase in procedural knowledge supports increase in conceptual knowledge and vice versa if procedural instruction is crafted to support the noticing of underlying concepts (B. Rittle-Johnson & Schneider, 2015).

In spite of the apparent difficulties in applying standard algorithms, several studies have shown that students seem to prefer the standard algorithm even when number based strategies would be easier. Reys and Yang (1998) found that students relied on the school-taught algorithms as their first choice of solution strategy and even though the students' were proficient in solving different tasks, interviews revealed that their number sense and understanding did not reflect their computational skills. Likewise, Torbeyns and Verschaffel (2016) found that even though the subtraction test items were intended to evoke mental computation, the students irrespective of achievement level applied the standard algorithm more frequently and efficiently than mental computation. High and above-average students based their strategy choices on their mastery of the different strategy, but in general, the students did not fit their choice of strategy to numerical characteristics of the items.

A prerequisite for using mental computation is an adequately developed number sense. For instance, Linsen, Verschaffel, Reynvoet, and De Smedt (2015) found that numerical magnitude processing skills was linked to both mental and algorithmic computation methods with respect to proficiency of doing subtraction in the number domain up to 100. However, the relationship was stronger for mental computation.
Although the beneficial effects of students learning and use of mental computation is often emphasized, as for example in Kilpatrick et al. (2001):

> Mental arithmetic, …, can provide opportunities for students to practice and use numbers and operations in ways that promote making sense of the mathematics and reveal further insights into the properties of numbers and operations. (p. 214)

The study of Torbeyns and Verschaffel (2016) show that it acquires an effort to maintain the students strategies and proficiency with mental computation once standard algorithms is introduced.

The Danish context

In a Danish context, focus in the curriculum has been on teaching strategies and mental computation since 2001. The current national curriculum comprises guiding learning objectives specified for skills as well as knowledge (Danish ministry of Education, 2017). As an example one of the objectives for years 1-3 of primary school in the topic Number and Algebra states that the students should be "*able to perform simple calculations with natural numbers"*, and have "*knowledge on strategies for simple calculations with natural numbers"* (Danish Ministry of Education, 2017, my translation).

The current study

This study investigates six Danish year one teachers' perspectives on the teaching and learning of number and addition in year one through analyses of semi-structured interviews. The study aims at analysing to what extent teachers' perspectives on the teaching and learning of number and addition in year one are aligned with established knowledge on foundations of numerical and arithmetic competence.

Methods

This study is based on semi-structured interviews with six Danish year one teachers. The teachers' utterances about number and addition were analysed using a constant comparison analytical approach drawn from grounded theory (Strauss & Corbin, 1998). In this context, constant comparison entailed a repeated reading of the data from each interview to identify categories of answers. As each category was identified, the data were reread to ensure the category had not been missed in the previously read interviews. Throughout the process, categorical definitions were constantly refined as incidents were compared and contrasted.

The selection of participants ensured an equal number of male and female teachers, and a cross section of professional experience (2-24 yrs.) and age (30-49 yrs.). The teachers were informed about the project both in writing and at an introductory meeting prior to the interviews took place. By the end of the project, the teachers were asked whether they wanted to read and approve the transcripts of the interview and whether they wanted to read the final analysis and paper. None of the teachers took advantage of this offer. Pseudonyms have been used throughout.

Teacher interviews

All interviews took place in October 2015. Each interview lasted between 45 and 60 minutes, and the interview focussed on the teachers' perspectives on teaching and learning of arithmetic in year one. The semi-structured interview was guided by questions related to the teacher's plans for and reflections on a specific observed lesson on number and arithmetic as well as general questions on the teaching and learning of number and arithmetic (addition and subtraction) in year one. The questions on the observed lesson was related to *planning of the lesson:* "Why did you choose these specific activities?", *carrying out the lesson*: "How did you experience the lesson? Did it proceed as you had expected?", *Progression:* "How will you follow up on this lesson? What will be the next step?", and c*haracteristics of a 'good* *activity'*: "what is a good activity and what makes it good? What does the students' learn in these activities?".

The general questions on the learning of number and addition was on *teaching addition*: "how do you introduce the students for addition" and "what aspects do you emphasize", *prerequisites for learning addition*: "What is the prerequisites for learning addition, how do you ensure the students has the prerequisites", and "Are there aspects of learning addition the students' find especially difficult", and *progression in teaching addition:* "How do you see the progression in teaching addition" and "What do you expect your students to know/be able to by the end of year one".

Throughout the interview, if the teacher primarily referred to practical aspects, e.g. "a good activity is easy to explain for the students" or "the lesson went well because many students participated in the activity", I specifically asked questions related to the mathematics of the activities and lesson, e.g. "what aspects of number and addition do you think the students learn through that activity". This was to ensure the teachers had opportunity to reflect on the aspects of mathematics of interest in this project.

The interviews were transcribed and analysed using NVivo, and excerpts presented here were all translated into English, a process that included transforming Danish idioms into equivalent English expressions without losing the speaker's intended meaning (Brinkmann & Kvale, 2015).

Results

The process of continual comparison and refinement of categories led to the identification of two themes: knowing numbers and doing addition. *Knowing numbers* is the term used by the teachers: "they (the students) should know the numbers". This theme includes categories related to number knowledge (e.g. knowing number names, symbols and quantity) and number competences, such as counting skills. The second theme *doing addition* comprises categories related to different calculation methods described by the teachers (e.g. strategies for mental addition and the standard written algorithm). It is worth to point out that all teachers mention the use of manipulatives in relation to both themes. However, this will not be detailed further in the analysis. In the following, the categories are described for each theme with examples from the interviews. Table 1 is a summary of the themes and categories and an overview of how the different categories are distributed over the teachers.

Knowing numbers

When the teachers were asked to elaborate on what they meant with "knowing numbers", they explicitly mentioned knowing 1) number symbol, 2) number name, and 3) quantity. Other categories assigned to this theme is 4) estimation, 5) counting skills (knowing the number sequence, the number before and after, skip counting or counting in times tables, and the use of number line), and 6) partitioning numbers (base ten number system knowledge, and decomposing and composing numbers).

Knowing the number symbols and number names was emphasised by all teachers but explicated very differently. The teachers expressed very different levels of necessary knowledge for the students. One teacher, Else, said that the students "need to know the numbers, their value and be able to write and read them, and recognise them" another, Allan, stated “the students need to learn the numbers to 100” and Dan explicated that the students do not have to know the number name as long as they “can write it” although he also emphasised knowing the names of the tens. Knowing the names of the tens was emphasised by for example Carl as a help to find the number names of two-digit numbers. Bettina, talking about the base ten number system, explained how she focused on “enhancing the students’ competencies of naming number”. Naming two-digit numbers are something that many students find difficult because of the Danish number names. Carl mentioned this and explains how he trained the students in activities where “they have to find the number 13 or 17 so they practise finding the right symbol for the correct number name”.

Quantity and the relation to number name and number symbol was addressed by four of the teachers but with different levels of articulation. Allan emphasised that “they need to recognise that quantity and number kind of go together”. Bettina said “the students need to understand the symbols and the naming of quantity”. Carl, elaborating on “the translation between number and quantity”, underlined that “they need to understand quantity; the symbol 4 equals four things”. Else emphasised the understanding of the relationship between number and quantity. “They need to have an idea of what value is and what is worth more (…) so many dots or centicubes, what is the size and quantity of that”.

Estimation was only mentioned by one teacher, Else. She very briefly referred to *estimation* of quantity by mentioning an activity of “how many in the jar”. This is an activity, where the students have different containers with an unknown number of items. The students then guess how many items are in the container and afterwards they count the exact number of items. However, she did not explicitly use the expression estimation.

Counting skills are addressed by all teachers. Carl stated “Early maths is mostly about counting” and Dan said that “they need to know the number sequence”. All the teachers provided many examples of counting procedures, often performed by the use of manipulatives or other representations e.g. a number line. The number line was mentioned by all the teachers in relation to activities of ordering numbers or “find the number” and when performing counting procedures. However, Carl and Else were the only teachers directly referring to knowledge of “the number before and after”. All teachers referred to skip counting, often by ten. Skip counting was paralleled with times tables and the teachers thus referred to “knowing the ten times table” when they taught the students to count in tens in order to find the name of a two-digit number.

Partitioning was addressed by all teachers. They either directly mentioned composing and decomposing numbers and quantities, mostly in groups of tens, or indirectly by referring to “ones and tens”. For example, Carl talked about being able to group in tens and Else described activities of grouping quantities and distinguishing tens and ones. The sub-category *base ten number system knowledge* was addressed explicitly by some of the teachers, but not all teachers articulated clearly. For example Frida described how she introduce two-digit numbers by explaining to the students that “when we reach 10 we put two numbers together”. Frida used the word ‘number’ and not ‘digit’. Distinguishing between number (in Danish ‘tal’) and digit (in Danish ‘ciffer’) is something students often find difficult. Dan, on the other hand, described how students should “know the system of numbers” and understand “that when we reach 39, what is the next number (…) that when you reach 9 in the ones you have to change in the tens (…) they should be able to handle the change in the tens”

In general the teachers’ level of articulation and vocabulary differed substantially. All teachers mentioned number name and number symbol. However, quantity is only mentioned by four of the teachers. Estimation stands out as the category only briefly mentioned by a single teacher. Knowing the number sequence and skip counting, especially with ten, is mentioned by all teachers, but only two specifically mentioned knowing the number before and after. Partitioning of numbers was explicated very differently, but addressed by six of the teachers. The last two categories, counting skills and partitioning numbers, are closely linked to the second theme *doing addition.*

Doing addition

This theme comprises the teachers different perspectives on how students should do addition in year one. The analysis revealed the following categories: 1) counting strategies, 2) direct retrieval (automatizing single-digit addition facts, knowing friends of ten), 3) derived fact strategies, and 4) standard written algorithms.

Counting strategies, as such, is mentioned by all teachers. However, they place different weight to the use of counting. Allan and Frida emphasized counting strategies and Allan differentiated between easy and difficult tasks: “Easy tasks are the ones you can count. When the tasks involves bridging ten and carrying, the counting is more complicated and then we use paper and pencil” and continued “The tools for doing calculations in first grade are primarily the fingers and then different number lines”. Frida referred to doing addition on the “counting line” and said she often asked her students to count on their fingers, even older students (year 9). The other teachers, Bettina, Carl, Dan, and Else, all explicated that counting was important but they emphasized that students should be able to use other strategies. Carl, for example said: “counting is often a bad habit” and he wanted to push the students forward and use other strategies. Else explained that although counting was important she found it essential to “work with friends of ten and automatize so we don’t have to ‘count on’ all the time”.

Direct retrieval is mentioned as knowing the single-digit sums “by heart” or some teachers only referred to “knowing friends of ten”. Allan stated that “I like friends of ten” and that the students should know the easy tasks by heart. Allan was the only teacher to emphasise speed of calculation and focus on right and wrong answers. Carl explicated that single-digit addition should not only be known by heart, but the students should also use the known facts for further calculation: “that you know the friends of ten and use them”. Dan emphasised knowing facts by heart and that knowing friends of tens “makes mental calculation so much easier”. Two teachers, Bettina and Frida, did not explicitly mention direct retrieval.

Derived fact strategies was a very explicit focus of Bettina, Carl, Dan, and Else. As an example, Bettina emphasised different calculation strategies like derived fact strategies as a focus point of her teaching and described how “the students are becoming aware that there are different ways of doing addition”. Dan described how he discussed different strategies with the students and present them for some “shortcuts” and reflects on the students own strategies: “Sometimes” he said, “the students present some very surprising ways of doing addition and it is fantastic when they can explain how they do it”. Allan and Frida did not talk about derived strategies. Besides counting and direct retrieval (Allan) they only mentioned doing addition by the standard written algorithm.

The standard written algorithm was mentioned by Carl and Dan as a long-term goal but they emphasized different calculation strategies. Whereas Allan and Frida described written algorithms as a part of their teaching in year one. As described above, Allan distinguished between easy and difficult tasks, the difficult tasks involving carrying is where "we use paper and pencil", a commonly used synonym of standard written algorithm. Frida described how she teaches the standard algorithm and that "they have to understand why you write one number over the other (…) that when you get 10 then you write the ones down here and the tens in the other row".

To sum up, the teachers perspectives on teaching and learning addition in year one fall in two distinct groups. Two teachers (Allan, and Frida) emphasised counting and standard written algorithms as their main focus in year one and they did not refer to derived fact strategies, whereas the other four teachers (Bettina, Carl, Dan and Else) in addition to counting emphasised derived fact strategies and described the standard written algorithm as a long-term goal.

Table 1: Overview of the two themes: knowing numbers and doing addition. X indicates that the teacher's utterances in the interview is assigned to that sub-category. A: Allan, B: Bettina, C: Carl, D: Dan, E: Else and F: Frida.

Knowing numbers							
Category	**Sub-category**	**A**	**B**	**C**	**D**	**E**	**F**
Number symbol		X	X	X	X	X	X
Number name		X	X	X	X	X	X
Quantity	Link between number and quantity	X	X	X		X	
Estimation						X	
Counting skills	Number sequence	X	X	X	X	X	X
	Number before and after			X		X	
	Use number line	X	X	X	X	X	X
	Skip counting/times tables	X	X	X	X	X	X
Partitioning numbers	Base ten knowledge		X	X	X	X	X
	Decomposing/composing numbers		X	X	X	X	X
Doing addition							
Category	**Sub-category**	**A**	**B**	**C**	**D**	**E**	**F**
Counting strategies		X	X	X	X	X	X
Direct retrieval	Automatisation of single-digit sums	X		X		X	
	Knowing friends of ten	X		X	X	X	
Derived fact strategies			X	X	X	X	
Written algorithms		X		X	X		X

Discussion

The teaching and learning of number is an important part of mathematics in year one. Providing students with opportunities to develop a profound number sense is a prerequisite for students' further development of arithmetic competence. The analysis of the interviews revealed that many components of number sense was addressed explicitly or explicitly by all teachers (i.e. number name, number symbol, quantity, counting skills, and partitioning of number) but two of the components found to be of importance in the early years, estimation and number comparison, was only addressed implicitly by a single teacher (estimation) or not at all (number comparison).

Although all teachers explicitly emphasized number symbols and number names (Chu et al., 2015) as the most important aspects of “knowing number” not all teachers mentioned quantity and creating the link between symbols and quantity (Geary, 2013). However, all teachers mentioned the importance and relevance of using manipulatives, which implicitly provides learning opportunities for students to create the link between number and quantity. Counting skills (Jordan et al., 2009) was explicitly mentioned by all teachers. Thus, some of the basic components of number sense and prerequisites for doing addition is explicitly or implicitly part of the teachers perspectives on the teaching and learning of number in year one. Some of the important aspects of number sense is number comparison (De Smedt et al., 2013) and estimation (Booth & Siegler, 2008; Gilmore et al., 2007). Of these skills, estimation of quantity was addressed by a single teacher, and comparison of numbers was not addressed by any of the teachers in the interviews. These findings resemble a cross-cultural study by Sayers and Andrews (2015) on the opportunities to learn different aspects of foundational number sense in different activities observed in six different European classrooms. Across countries, they found no episodes where teachers encouraged students to estimate and only 2 of 18 episodes where a single teacher introduced quantity discrimination.

Partitioning and knowledge of base-ten number structure is a prerequisite of developing flexible strategies for mental computation (Fuson, 1992; Laski et al., 2014; Verschaffel et al., 2007). All teachers addressed base ten knowledge explicitly or implicitly. However, the extent to which the teachers emphasised the students’ further development of flexible strategies and mental computation differed. One group of teachers, Bettina, Carl, Dan and Else, emphasized the development of derived fact strategies as well as automatisation, and standard written algorithms were seen as a long-term goal, thus a focus on mental computation as the main objective in year one. Another group, Allan and Frida, demonstrated an emphasis towards the teaching and learning of counting, automatisation and standard written algorithms.

In conclusion, the findings indicate that teachers are cognisant of a wide range of the important foundations for developing arithmetic competence. However, it is also apparent that two crucial aspects, quantity discrimination and estimation skills, were either not addressed or only mentioned implicitly by a single teacher in the interviews.

This study cannot provide insight in how teachers actually teach. The analysis can only give an indication of how teachers view what they emphasise in class. Further research on video observations and student development will show to what extend the findings of emphasising mental and/or algorithmic computation and the lack of awareness on quantity discrimination and estimation skills are accentuated in the actual teaching and learning in the classroom.

References

Bjorklund, D. F., Hubertz, M. J., & Reubens, A. C. (2004). Young children’s arithmetic strategies in social context: How parents contribute to children’s strategy development while playing games. *International Journal of Behavioral Development*, *28*(4), 347–357. https://doi.org/10.1080/01650250444000027

Booth, J. L., & Siegler, R. S. (2008). Numerical Magnitude Representations Influence Arithmetic Learning. *Child Development*, *79*(4), 1016–1031.

https://doi.org/10.1111/j.1467-8624.2008.01173.x

Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed.). Thousand Oaks, CA: SAGE.

Carr, M., & Alexeev, N. (2011). Fluency, accuracy, and gender predict developmental trajectories of arithmetic strategies. *Journal of Educational Psychology, 103*(3), 617–631. https://doi.org/10.1037/a0023864

Carr, M., Steiner, H. H., Kyser, B., & Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learning and Individual Differences*, *18*(1), 61–75. https://doi.org/10.1016/j.lindif.2007.04.005

Chang, A., Sandhofer, C. M., & Brown, C. S. (2011). Gender biases in early number exposure to preschool-aged children. *Journal of Language and Social Psychology*, *30*(4), 440–450. https://doi.org/10.1177/0261927X11416207

Chu, F. W., VanMarle, K., & Geary, D. C. (2015). Early numerical foundations of young children's mathematical development. *Journal of Experimental Child Psychology*, *132*, 205–212. https://doi.org/10.1016/j.jecp.2015.01.006

Cowan, R., Donlan, C., Shepherd, D.-L., Cole-Fletcher, R., Saxton, M., & Hurry, J. (2011). Basic calculation proficiency and mathematics achievement in elementary school children. *Journal of Educational Psychology*, *103*(4), 786–803. https://doi.org/10.1037/a0024556

De Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55. https://doi.org/10.1016/j.tine.2013.06.001

Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences*, *16*(4), 351–367. https://doi.org/10.1016/j.lindif.2006.12.006

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Development Perspectives*, *7*(2), 74–79. https://doi.org/10.1111/cdep.12019

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M. L., & Levi, L. W. (1998). A Longitudinal Study of Gender Differences in Young Children's Mathematical Thinking. *Educational Researcher*, *27*(5), 6–11. https://doi.org/10.3102/0013189X027005006

Fuson, K. C. (1992). Research on whole number addition and subtraction. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning: A project of the National Council of Teachers of Mathematics.* (pp. 243–275). New York, NY, England: Macmillan Publishing Co, Inc.

Fuson, K. C., & Burghardt, B. H. (2003). Multidigit addition and subtraction methods invented in small groups and teacher support of problem solving and reflection. In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise* (pp. 267–307). https://doi.org/10.4324/9781410607218

Gaidoschik, M. (2012). First-Graders' Development of Calculation Strategies: How Deriving Facts Helps Automatize Facts. *Journal Für Mathematik-Didaktik*, *33*(2), 287–315. https://doi.org/10.1007/s13138-012-0038-6

Geary, D. C. (2013). Early Foundations for Mathematics Learning and Their Relations to Learning Disabilities. *Current Directions in Psychological Science*, *22*(1), 23–27. https://doi.org/10.1177/0963721412469398

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLOS ONE*, *8*(1), e54651. Retrieved from https://doi.org/10.1371/journal.pone.0054651

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities*, *38*(4), 293–304. https://doi.org/10.1177/00222194050380040301

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*, 589–591. https://doi.org/10.1038/nature05850

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850–867. https://doi.org/10.1037/a0014939

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: helping children learn mathematics*. Washington: National Academy Press.

Laski, E. V., Ermakova, A., & Vasilyeva, M. (2014). Early use of decomposition for addition and its relation to base-10 knowledge. *Journal of Applied Developmental Psychology*, *35*(5), 444–454. https://doi.org/10.1016/j.appdev.2014.07.002

Linsen, S., Verschaffel, L., Reynvoet, B., & De Smedt, B. (2015). The association between numerical magnitude processing and mental versus algorithmic multi-digit subtraction in children. *Learning and Instruction*, *35*, 42–50. https://doi.org/10.1016/j.learninstruc.2014.09.003

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, *17*(5), 714–726. https://doi.org/10.1111/desc.12152

Mazzocco, M., Murphy, M., Brown, E., Rinne, L., & Herold, K. (2013). Persistent consequences of atypical early number concepts. *Frontiers in Psychology*, *4*, 486. https://doi.org/10.3389/fpsyg.2013.00486

Nunes, T., Dorneles, B. V., Lin, P.-J., & Rathgeb-Schnierer, E. (2016). *Teaching and Learning About Whole Numbers in Primary School*. https://doi.org/10.1007/978-3-319-45113-8_1

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology*, *67*(3), 345–357. https://doi.org/10.1111/j.2044-8279.1997.tb01249.x

Price, G. R., Mazzocco, M. M. M., & Ansari, D. (2013). Why Mental Arithmetic Counts: Brain Activation during Single Digit Arithmetic Predicts High School Math Scores. *Journal of Neuroscience*, *33*(1), 156–163. https://doi.org/10.1523/JNEUROSCI.2936-12.2013

Reys, R. E., & Yang, D.-C. (1998). Relationship between Computational Performance and Number Sense among Sixth- and Eighth-Grade Students in Taiwan. *Journal for Research in Mathematics Education*, *29*(2), 225–237. https://doi.org/10.2307/749900

Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge of mathematics. In R. C. Kadosh & A. Dowker (Eds.), *The oxford handbook of numerical cognition; the oxford handbook of numerical cognition* (pp. 1118–1134). New York, NY.: Oxford University Press.

Sayers, J., & Andrews, P. (2015). Foundational Number Sense : Summarising the Development of an Analytical framework. In N. V. Konrad Krainer & N. V. Konrad Krainer (Eds.), *CERME 9 - Ninth Congress of the European Society for Research in Mathematics Education, Prague, Czech Republic, February 2015* (pp. 361–367). Retrieved from https://hal.archives-ouvertes.fr/hal-01286842

Torbeyns, J., & Verschaffel, L. (2016). Mental computation or standard algorithm? Children's strategy choices on multi-digit subtractions. *European Journal of Psychology of Education*, *31*(2), 99–116. https://doi.org/10.1007/s10212-015-0255-8

Vanbinst, K., Ghesquière, P., & De Smedt, B. (2014). Arithmetic strategy development and its domain-specific and domain-general cognitive correlates: A longitudinal study in children with persistent mathematical learning difficulties. *Research in Developmental Disabilities*, *35*, 3001–3013. https://doi.org/10.1016/j.ridd.2014.06.023

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning.* (pp. 557–628). Greenwich, CT: Information Age Publishing.

Appendix A

Appendix A

Interview guide for teacher interviews (in Danish).

Interviewguide

Lærerne interviewes to gange i forløbet, så ikke alle interview-dele gennemføres hver gang.

Interviewet indeholder flere delelementer. Første del er et tidslinje interview med det formål at afdække lærerens baggrund og udvikling til det faglige ståsted de befinder sig på nu. Dette er inspireret af Adriansen (1999). Anden del omhandler den observerede og videofilmede undervisning. Tredje del er bygget op om case-eksempler om undervisningssituationer. Gennem diskussion og refleksion over disse er målet at få indblik i lærerenes viden om læring og undervisning samt faglige indsigt. Dette er inspireret af Ball (1988) og Ma (1999).

Tidslinjeinterview

Dette tidslinje interview strækker sig over to gange.

Din personlige historie:

> Første interviewgang:
> *Jeg vil gerne starte med at spørge ind til de begivenheder som har bragt dig derhen, hvor du er i dag rent fagligt. Vi skal sammen konstruere en tidslinje for de begivenheder der har betydet noget for de matematik-faglige valg du træffer i din undervisning i 1. klasse. Vi begynder ved din uddannelse som lærer og fortsætte op til i dag. Vi kan bevæge os frem og tilbage og skrive til undervejs.*

På et A3 tegnes en linje. Eksterne faktorer (uddannelse, kurser etc.) skrives på venstre side, Personlige oplevelser og erkendelser skrives på højre.

> Anden interviewgang:
> *Første gang vi talte sammen lavede vi denne tidslinje sammen over de begivenheder og ting der har haft betydning for hvordan du underviser i 1. klasse - hvad det er for nogle faglige valg du træffer. Hvis vi skal prøve at fortsætte tidslinjen frem til i dag hvor der er gået ca. ½ år, er der så noget vi kan skrive på? Hvis du har yderligere kommentarer til den første tidslinje kan vi også vende tilbage til den undervejs.*

Respondenten får det oprindelige A3 med tidslinjen og der tilføjes evt. et nyt ark.

Den observerede undervisning

Denne del af interviewet gennemføres efter første og sidste observation. Første gang tales blot om optagelserne. Anden gang har jeg udvalgt et kort klip som vi taler ud fra.

Tidsforbrug: ca. 20 min

Valget af undervisningsaktiviteter (gennemføres ved begge interview)

Lad os tale om den lektion jeg har videofilmet. Jeg vil gerne høre lidt om de undervisningsaktiviteter du havde valgt. Vi starter med at tale om dine overvejelser inden undervisningen og bagefter taler vi om selve gennemførelsen og selve undervisningssituationen.

1) *Hvorfor havde du valgt netop disse aktiviteter?*
   a) *Hvad ville du gerne have eleverne skulle have med sig fra aktiviteterne?*
   b) *Er der noget du synes er særlig vigtigt i disse aktiviteter?*
   c) *Hvad gjorde du dig af overvejelser inden selve undervisningen?*
   d) *Er der nogle særlige udfordringer i forhold til denne elevgruppe og de valgte aktiviteter?*
2) *Hvordan oplevede du aktiviteterne i selve undervisningen?*
   a) *Forløb de som du havde forventet?*
   b) *Hvad oplevede du eleverne fik ud af aktiviteterne?*
3) *Hvordan tænker du at følge op på de aktiviteter/ hvad er det næste du vil lave med eleverne?*
   a) *Hvordan tænker du det hænger sammen med de foregående aktiviteter?*
4) *Hvilke aktiviteter oplever du som særligt gode?*
   a) *Hvad er det der gør dem gode? Hvordan er de gode?*
   b) *Hvad er det eleverne lærer lige præcis ved de aktiviteter?*

Udvalgte episoder fra de to forrige observationer (kun sidste interview)

Jeg har udvalgt et par sekvenser fra de to forrige observationsgange. Dem får du lov til at se igennem. Bagefter taler vi om hvad og hvordan du tænker aktiviteterne understøtter arbejdet med tal og regning.

1) *Hvilke områder inden for tal og regning tænker du aktiviteten understøtter?*
   a) *Hvad er det i aktiviteten/situationer som gør at du tænker det?*
   b) *Hvor vigtigt er det i forhold til at lære at regne?*
   c) *Hvad vil det sige "at lære at regne"?*
2) *Er der noget andet du vil sige om disse klip?*

Tanker om undervisning og case-eksempler: første interview

Tidsforbrug ca. 20 min

Addition

Et af de emner der fylder meget i 1. klasse er addition og jeg vil gerne høre lidt om hvordan du griber det an og hvilke ting du lægger særligt vægt på i din undervisning.

1) *Hvordan introducerer du eleverne for addition?*
   a) *Er der noget du lægger særlig vægt på?*
2) *Hvilke forudsætninger tænker du skal være på plads før eleverne kan lære addition?*
   a) *Er der noget som du oplever er særlig svært for eleverne?*
   b) *Hvordan sikrer du dig at elevernes forudsætninger er i orden?*
3) *Hvordan tænker du progressionen i undervisningen?*
   a) *Hvad tænker du eleverne skal kunne ved udgangen af 1. klasse?*
   b) *Hvad ser du efter eller hvad holder du særligt øje med at eleverne får lært?*

Subtraktion med 10'er overgang

Jeg vil gerne høre lidt om dine tanker om og erfaringer med at undervise i subtraktion med 10-er overgang. Vi kan tage udgangspunkt i opgaver som f.eks. 52-25, 91-79 (respondenten præsenteres for et ark med forskellige opgaver).

1) *Hvordan vil du undervise elever i 1.-2. klasse i denne type opgaver?*
   a) *Hvorfor vil du gøre det på den måde?*
   b) *Hvordan er du kommet på den ide?*
   c) *Hvad mener du med __________?*
   d) *Kan du give nogle eksempler på hvad du mener?*
   e) *Kan du forestille dig at gøre det på andre måder?*
2) *Hvordan kan du se at dine elever kan finde ud af subtraktion med 10'er overgang?*
   a) *Hvad ser du efter eller hvad er du særligt opmærksom på?*
3) *Hvad tænker du en elev skal forstå eller kunne før de begynder at lære subtraktion med 10-er overgang?*
   a) *Hvorfor tænker du _____ er vigtigt?*
   b) *Er der noget her du tænker er særlig vanskeligt for eleverne?*
4) *Hvad gør du hvis du opdager at en af dine elever regner på denne måde: 40-22=22?* (Respondenten vises et ark med opgaver fra før i lodret opstilling og med diverse regnefejl).
   a) *Hvad tænker du eleven forstår? Hvorfor?*
   b) *Hvad tænker du eleven ikke forstår? Hvorfor?*
   c) *Hvordan tænker du eleven kan hjælpes videre?*

Tanker om undervisning og case-eksempler: andet interview

Tidsforbrug ca. 20 min

Addition

Sidste gang talte vi om hvordan du ville arbejde med addition i 1. klasse, hvordan du introducerer eleverne for det og hvad du ville lægge særlig vægt på. Nu er der jo gået noget tid og du har været med til at evaluere eleverne både med en MAT-test og strategi-interview

1) *Er der noget du ville gøre anderledes med den viden du har om eleverne i dag?*
2) *Hvad er det der har ændret/ikke ændret dit syn på undervisningen og eleverne?*

Evaluering med MAT og strategi-test

I dette projekt er eleverne blevet testet med en MAT-test og strategi-interview. Hvordan oplever du disse test i forhold til dit arbejde som lærer, f.eks. i forhold til planlægningsopgaven?

1) *Har du brugt resultaterne?*
   a) *Hvordan?*
2) *Hvordan plejer du at arbejde med evaluering i forbindelse med undervisningen?*
   a) *Hvilke fordele eller ulemper ser du i forhold til det at anvende henholdsvis en MAT test og en strategi-test?*

Afrunding

Efter hver interviewrunde afsluttes med følgende:

Nu har vi været igennem mine spørgsmål, så her til sidst vil jeg spørge dig, om du har noget du gerne vil slutte af med at sige eller tilføje i forhold til de ting vi har talt om?

Tak…

Referencer

Adriansen, H. K. (2012). Timeline interviews: A tool for conducting life history research. *Qualitative studies*, *3*(1), 40-55.

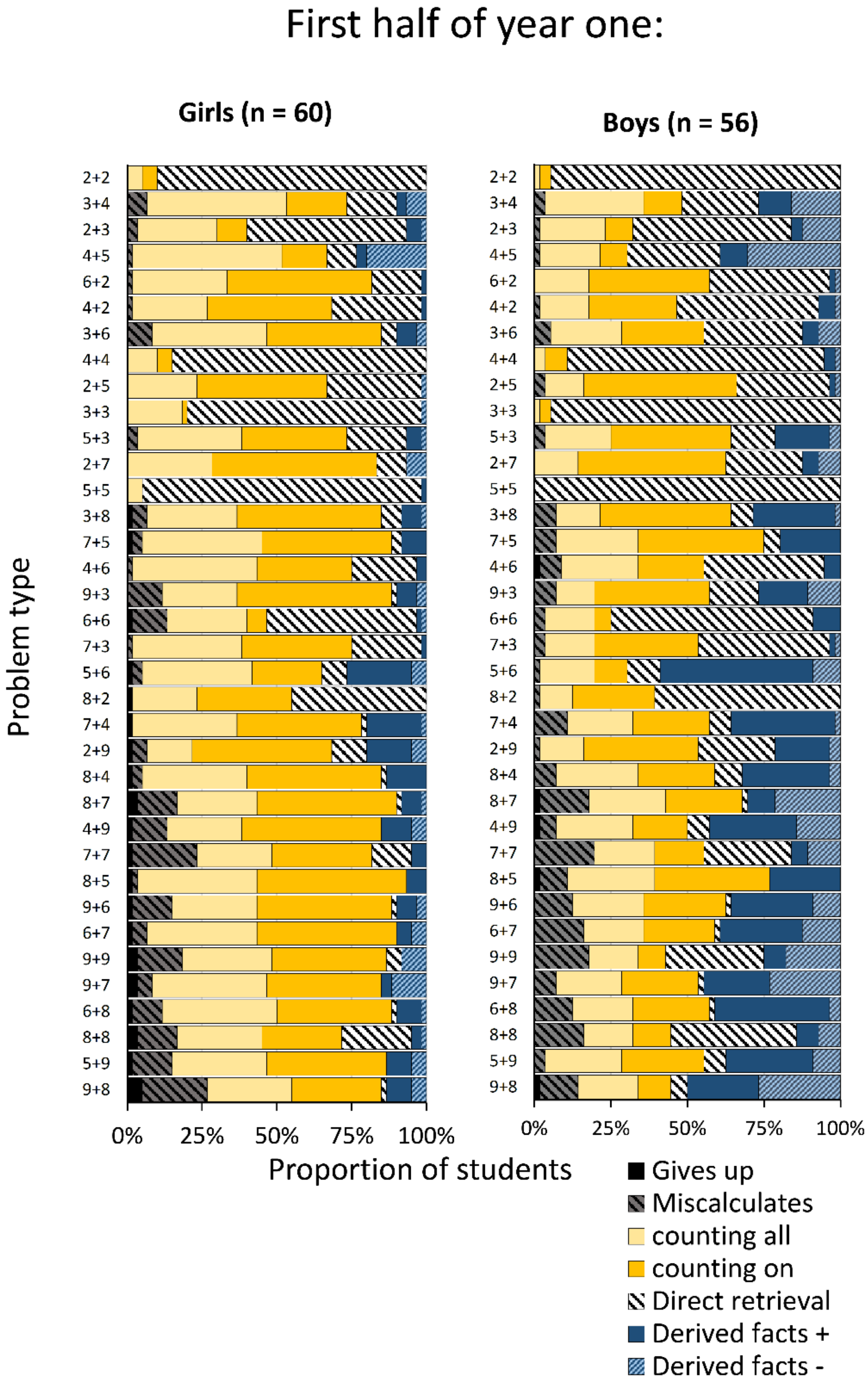
Ball, D. L. (1988). *Knowledge and reasoning in mathematical pedagogy: Examining what prospective teachers bring to teacher education* (Doctoral dissertation, Michigan State University).

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Lawrence Erlbaum Associates.

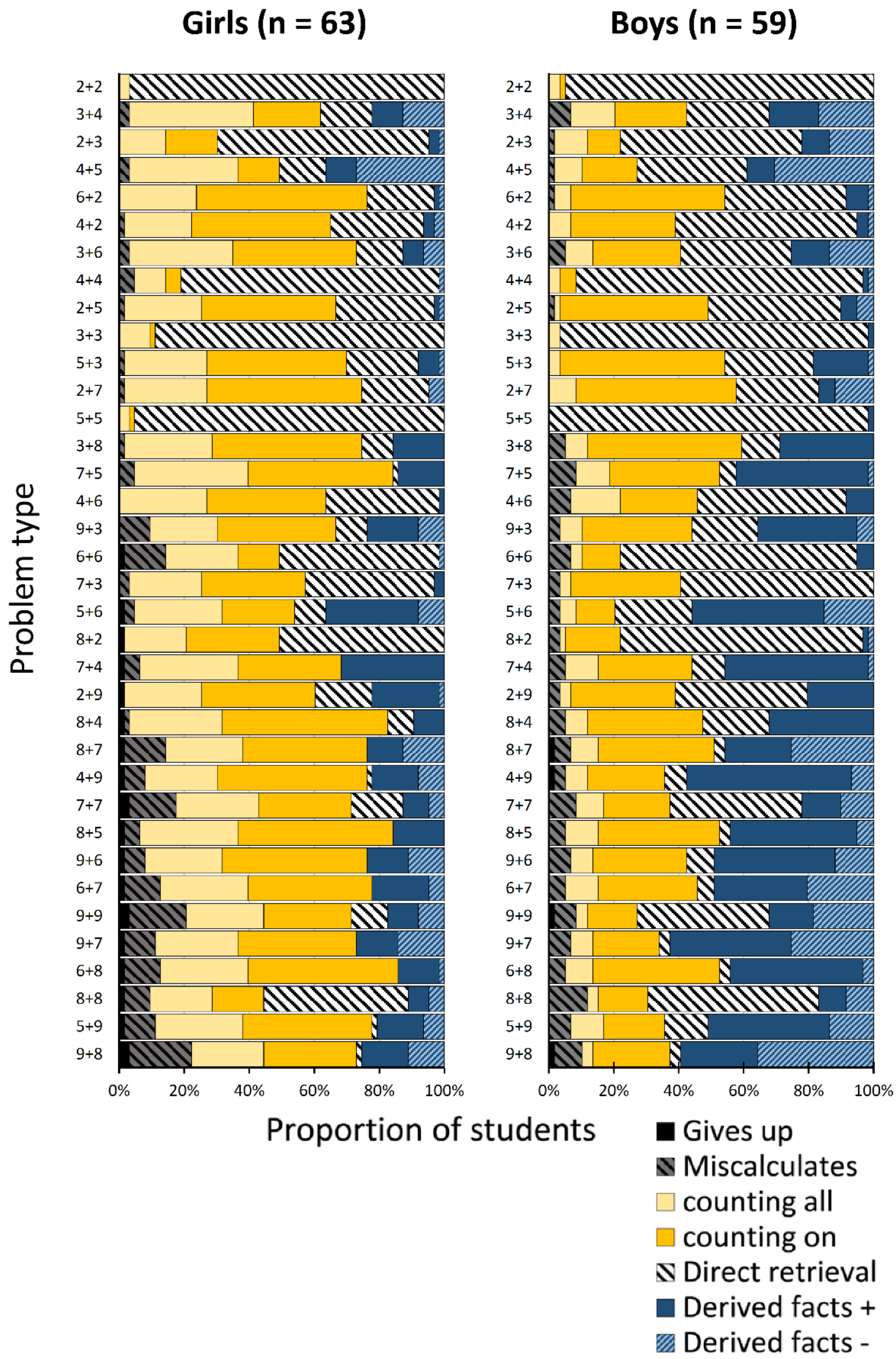
Appendix B

Appendix B

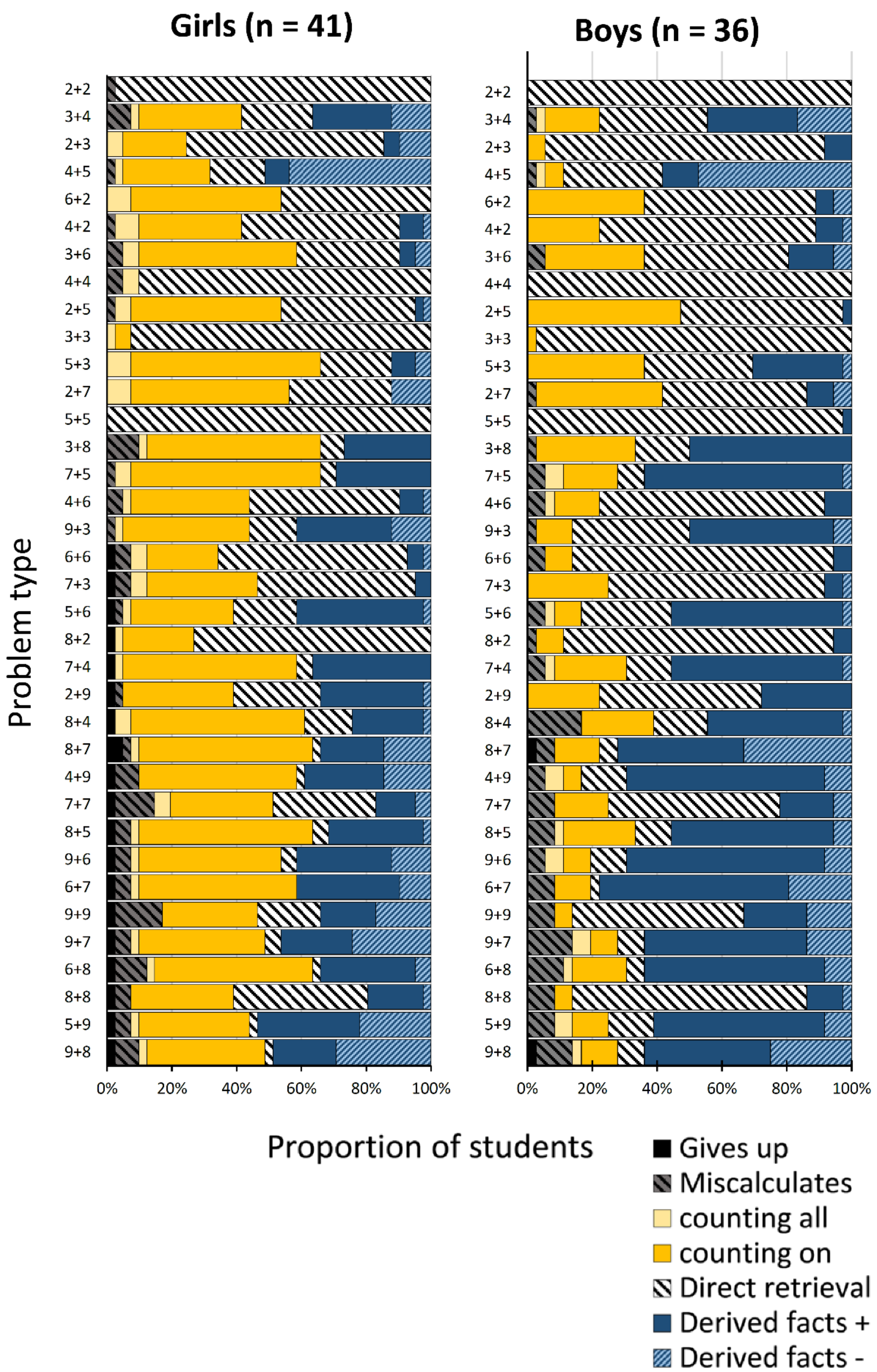
Distribution of strategy use categories amongst students divided on addition items, sex and school age (Study A and Study B combined).



Second half of year one:

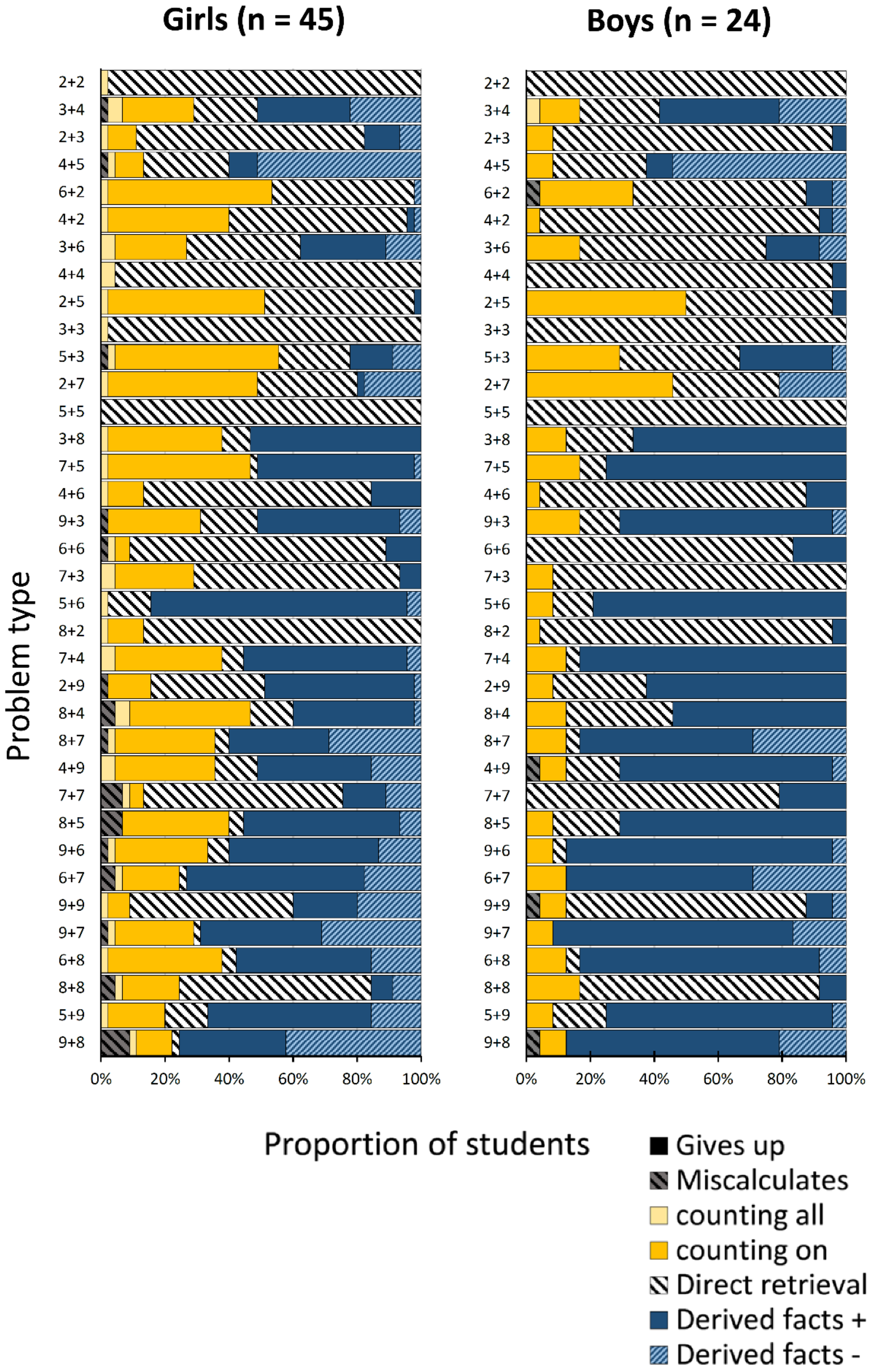


Year two:

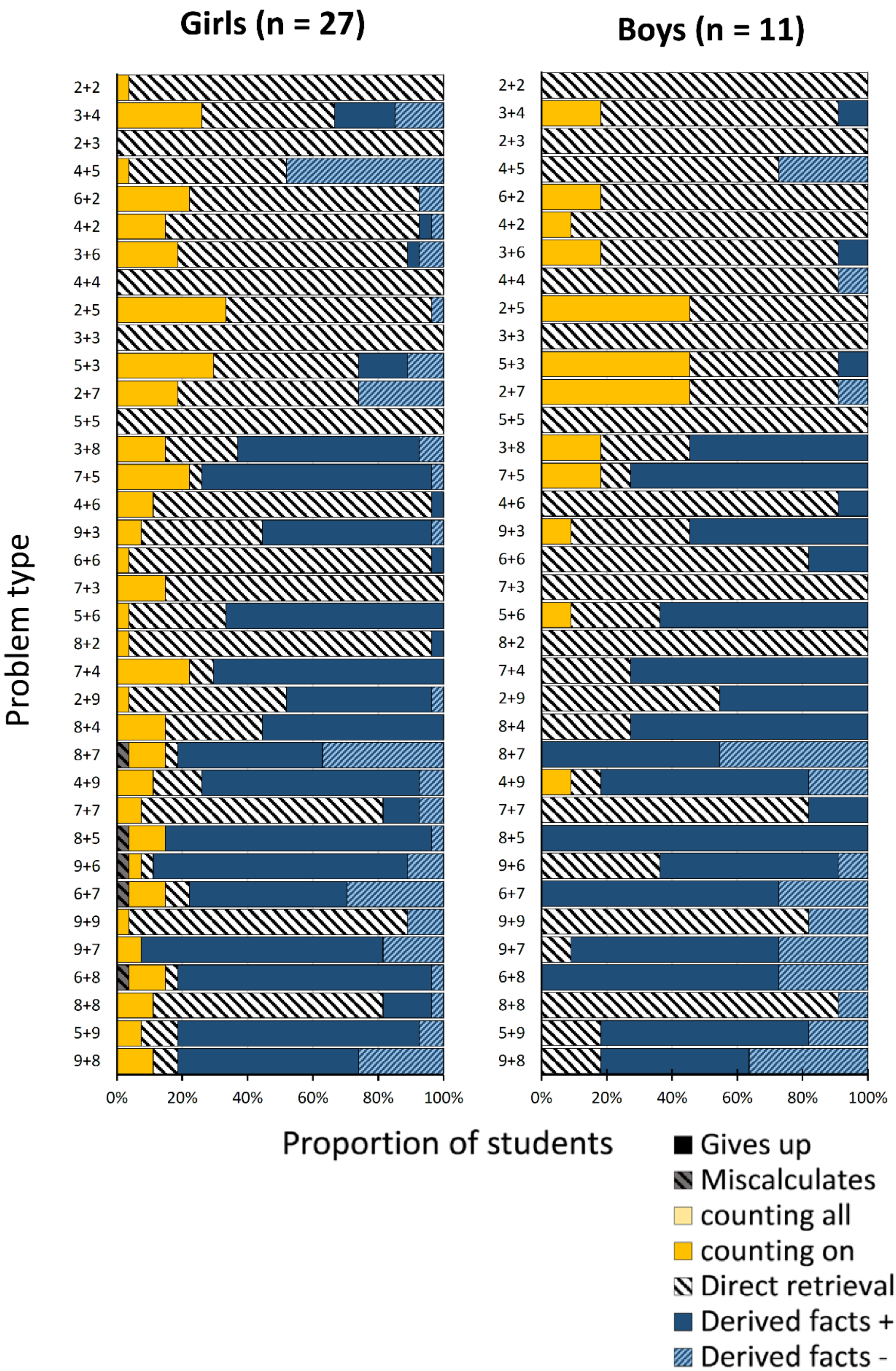


Girls (n = 41)
Boys (n = 36)
Problem type
2+2
3+4
2+3
4+5
6+2
4+2
3+6
4+4
2+5
3+3
5+3
2+7
5+5
3+8
7+5
4+6
9+3
6+6
7+3
5+6
8+2
7+4
2+9
8+4
8+7
4+9
7+7
8+5
9+6
6+7
9+9
9+7
6+8
8+8
5+9
9+8
0%
20%
40%
60%
80%
100%
Proportion of students
Gives up
Miscalculates
counting all
counting on
Direct retrieval
Derived facts +
Derived facts -

Year 3:


Girls (n = 45)
Boys (n = 24)
Problem type
2+2
3+4
2+3
4+5
6+2
4+2
3+6
4+4
2+5
3+3
5+3
2+7
5+5
3+8
7+5
4+6
9+3
6+6
7+3
5+6
8+2
7+4
2+9
8+4
8+7
4+9
7+7
8+5
9+6
6+7
9+9
9+7
6+8
8+8
5+9
9+8
0%
20%
40%
60%
80%
100%
Proportion of students
Gives up
Miscalculates
counting all
counting on
Direct retrieval
Derived facts +
Derived facts -

Year 4:



Girls (n = 27)
Boys (n = 11)
Problem type
2+2
3+4
2+3
4+5
6+2
4+2
3+6
4+4
2+5
3+3
5+3
2+7
5+5
3+8
7+5
4+6
9+3
6+6
7+3
5+6
8+2
7+4
2+9
8+4
8+7
4+9
7+7
8+5
9+6
6+7
9+9
9+7
6+8
8+8
5+9
9+8
0%
20%
40%
60%
80%
100%
Proportion of students
Gives up
Miscalculates
counting all
counting on
Direct retrieval
Derived facts +
Derived facts -

This dissertation addresses Danish students' use and development of mental strategies in single-digit addition in the first years of school and relates it to teaching practices, later mathematical achievement and teachers' perspectives on teaching and learning of number and arithmetic.

The work builds on data from a study of 147 students' development of strategy use from year one to four, and a study of six year one classes (83 students, six teachers) assessed twice (October/November, April/May) in year one. The latter study also included data from teacher interviews, classroom observations, and students' achievement in arithmetic, fraction knowledge and word problem solving in year four.

From year one to year four, students' use of counting strategies decreased, and their use of fact-based strategies increased, but with substantial individual variation. On average boys, were two years ahead of girls in strategy use development. Strategy use patterns varied little between classes and did not develop differently across classes in year one. It follows that differences in teaching practice did not result in any traceable differences in the pace of development of strategy use during year one. Measures of strategy use in year one explained variation in mathematical achievement in year four that could not be explained by other year-one variables, achievement test measures included.

The results indicate that habits of strategy use are deeply rooted within the individual child and seem to establish either before or at the outset of formal schooling, after which it develops slowly over time. The results highlight the relevance of students' early understanding of number and arithmetic (i.e. strategy use) both as indicators of later achievement and as a focus for intervention and targeted teaching to support student development in all cases.

DPU – Danmarks institut for Pædagogik og Uddannelse
Aarhus Universitet
Tuborgvej 164
2400 København NV

Tlf.: 8716 1300
E-mail: dpu@au.dk
edu.au.dk