Universidade Federal de Minas Gerais

Faculdade de Filosofia e Ciências Humanas

Programa de Pós-graduação em Psicologia: Cognição e Comportamento

Fernanda Rocha de Freitas

To calculate in your mind, don't miss, take five:
The effect of split-five error in Brazilian children

O efeito do erro split-five em crianças brasileiras

Belo Horizonte

2021

Fernanda Rocha de Freitas

To calculate in your mind, don't miss, take five:
The effect of split-five error in Brazilian children

O efeito do erro split-five em crianças brasileiras

Versão Final

Dissertação de mestrado apresentada à Faculdade de Filosofia e Ciência Humanas da Universidade Federal de Minas Gerais como parte dos requisitos de obtenção do grau de Mestre em Psicologia: Cognição e Comportamento.
Área de concentração: Neuropsicologia do Desenvolvimento

Orientador: Vitor Geraldi Haase
Coorientadora: Júlia Beatriz Lopes Silva

Belo Horizonte

2021

153.4
F866t
2021

Freitas, Fernanda Rocha de.

To calculate in your mind, don't miss, take five [manuscrito] : the effect of split-five error in Brazilian children / Fernanda Rocha de Freitas. - 2021.

80 f. : il.

Orientador: Vitor Geraldi Haase.
Coorientadora: Júlia Beatriz Lopes Silva.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Faculdade de Filosofia e Ciências Humanas.

Inclui bibliografia.

1.Psicologia – Teses. 2. Psicologia cognitiva – Teses. 3.Aprendizagem – Teses. 4. Aritmética – Teses. I .Haase, Vitor Geraldi. II. Silva, Júlia Beatriz Lopes. III. Universidade Federal de Minas Gerais. Faculdade de Filosofia e Ciências Humanas. IV.Título.

Ficha catalográfica elaborada por Vilma Carvalho de Souza - Bibliotecária - CRB-6/1390

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA: COGNIÇÃO E COMPORTAMENTO

UFMG

FOLHA DE APROVAÇÃO

O efeito do erro split-five em crianças brasileiras

FERNANDA ROCHA DE FREITAS

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em PSICOLOGIA: COGNIÇÃO E COMPORTAMENTO, como requisito para obtenção do grau de Mestre em PSICOLOGIA: COGNIÇÃO E COMPORTAMENTO, área de concentração PSICOLOGIA: COGNIÇÃO E COMPORTAMENTO, linha de pesquisa Neuropsicologia do Desenvolvimento.

Aprovada em 30 de julho de 2021, pela banca constituída pelos membros:

Prof(a). Vitor Geraldi Haase - Orientador
UFMG

Prof(a). Julia Beatriz Lopes Silva
UFMG

Prof(a). Julia Bahnmueller
Loughborough University

Antonio Jaeger:89249763034
Assinado de forma digital por Antonio Jaeger:89249763034
Dados: 2021.08.05 16:37:23 -03'00'

Prof(a). Antônio Jaeger
UFMG

Belo Horizonte, 30 de julho de 2021.

Acknowledgement

To my supervisor, Professor Vitor Haase, whom I deeply admire as a researcher, teacher, and human being. Your profound belief in my potential pushed me through all this journey. It is difficult to summarize in words the gratitude I feel for the broad teachings and important life advice that you gave to me.

To my co-supervisor, Professor Júlia Lopes Silva, not only for being a valuable advisor, but also a role model of someone who knows how to balance professional and personal life with grace and joy. You always encourage me to trust my potential and face the pathway challenges.

To CAPES for the financial support of my work all these years since undergraduate times.

To the schools, families and children who accepted participating in our research projects. In special, to the teachers whose contributions are inestimable through the whole data collection.

To the Master's thesis committee, namely Professors Antônio Jaeger, Julia Bahnmueller, Maycoln Teodoro and Moritz Herzog for kindly accepting my invitation. You are great examples of researchers for me, who already have contributed a lot during my trajectory. I am honored to receive your feedbacks on my work.

To Professor Korbinian Moeller, for the valuable advice on the study and for kindly teaching me how to write a simple, but great story. I have always admired your work, and I feel honoured to work in collaboration with you.

To Professor Annemarie Fritz, for being such a role model. I do not have words to express my gratitude for your warm welcoming and for always encouraging me to go forward, supporting my difficulties and acknowledging my achievements. You and Moritz have waited patiently for the conclusion of this Master cycle, and it is finally the time to work hard together on our project.

To all formers and recent LND members, who had immensely contributed over these 11 years of research in numerical cognition. The present study is also the result of your efforts. To Andressa Antunes, Annelise Júlio Costa and Isabela Starling for supporting my first steps in research. Also, to my work team Marcela Rodrigues and Higna Ester, it is a pleasure to orientate such a competent team, your efforts were especially important for the conclusion of this study.

To all my LND friends – that I will not cite fearing to forget someone – for making this lab my second home. I will care all our experiences and your friendship to my whole life. My special thanks to Malu Gomides and Giulia Paiva for guiding my academic path, always encouraging me to overcome my difficulties.

To my duo Luana Teixeira, who has been throughout the entire journey since undergraduate times, always growing together and supporting each other. Sharing the long learning journey with you made it even more special.

To my master’s colleagues, I feel grateful for sharing the challenges and the great moments of this experience with you. Specially, to my dear friend Ângela Ribeiro, our friendship was one of the greatest gifts of the master’s time and I am sure this is just the first adventure we will share in life. To my friends Daniel Ezequiel and Larissa Coutens for the meetings full of laughs that made the journey lighter.

To my friend Charlotte Steinke, not only for teaching me German, but also for being a great supporter during this difficult pandemic times. Our escapes to breathe fresh air helped me to focus and complete my master's thesis.

To Layla Alvarenga, for being the best friend I could have and holding my hand on the most difficult times, including the conclusion of the thesis.

To my families – Rocha, Freitas, and Hamdan – for being my foundation and the cosiness that I know I can count on any circumstance. In special to my mother, father, and my dear brother for the unconditional love and all the support.

Resumo

Os dedos desempenham um papel importante no início da aprendizagem aritmética. A sub-base cinco derivada da utilização dos dedos para construção do raciocínio aritmético continua a influenciar a performance mesmo quando já não há mais a utilização de estratégias explícitas de contagem nos dedos. O efeito do erro *split-five* (S5), um desvio de exatamente cinco entre a resposta da criança e o resultado correto, é uma evidência da influência da sub-base cinco. O principal objetivo da presente dissertação é compreender a performance e os padrões de erros das crianças brasileiras visando a investigação do erro S5 em cálculos aritméticos básicos de adição, subtração e multiplicação. Foram avaliadas 1160 crianças brasileiras do 1º ao 5º ano do ensino fundamental. Os resultados apoiam as evidências já encontradas na literatura de que o erro S5 é mais frequente em crianças mais novas, possivelmente associado ao uso de estratégias que utilizam a estrutura das representações numéricas baseadas nos dedos. Não foram observadas diferenças de sexo na frequência de crianças que cometeram erros S5. Um efeito de distância (relação inversa entre a frequência de erros e a distância entre o erro e o resultado correto) também foi observado para todas as operações, exceto multiplicação simples. As crianças que cometeram erros S5 na adição complexa, subtração complexa e multiplicação simples tiveram um desempenho geral inferior em matemática. Em suma, os resultados da presente dissertação demonstram a importância da análise qualitativa dos erros cometidos nos cálculos aritméticos básicos, e a influência persistente das representações numéricas baseadas nos dedos.

Palavras-chave: Operações aritméticas básicas, Erro Split-five, Representações numéricas baseadas nos dedos, Estratégia de contagem nos dedos.

Abstract

Fingers play an important role in early arithmetic learning. Subbase five derived from the use of fingers to build arithmetic reasoning continues to influence performance even when there is no longer the explicit use of finger-based counting strategies. The split-five error effect (S5), a deviation of exactly five between the child's response and the correct result, is evidence of the influence of subbase five. The main goal of this master thesis is to understand the performance and error patterns of Brazilian children to investigate the split-five error in basic arithmetic operations of addition, subtraction, and multiplication. 1160 children from the 1st to the 5th grade of elementary school were selected. The results support the evidence that the S5 error is more frequent in younger children, possibly associated with the use of strategies that use the structure of finger-pattern. No sex differences were observed in the frequency of children who made S5 errors. A Split effect (inverse relationship between the frequency of errors and the distance between the error and the correct result) was also observed for all operations, except simple multiplication. In general, children who commit S5 errors in complex addition, complex subtraction, and simple multiplication had lower overall math performance. In summary, the results of the present study demonstrate the importance of qualitative analysis of errors in basic arithmetic calculations, and the persistent influence of finger-based numerical representations.

Keywords: Basic arithmetic operations, Split-five error, Finger-based numerical representation, Finger counting strategy.

Abbreviations and Acronyms

S	Split
S5	Split-five
ANS	Approximate Number System
FBRs	Finger-Based Numerical Representations
FCS	Finger-Counting Strategies
RCPM	Raven's Coloured Progressive Matrices
BAOT	Basic Arithmetic Operations Task
TDE	*Teste de Desempenho Escolar* (Brazilian School Performance Test)
GS5	Groups of children who committed at least one Split-five error
GNoS5	Groups of children who did not commit Split-five error

Summary of Figures

Summary of Tables

Table of Contents

Introduction

Mathematical reasoning is involved in many everyday situations like shopping, measuring distances, cooking or even for deciding about activating or not the snooze function before getting up in the morning. Math education has been related to employability, income, and mental health (Parsons & Bynner, 2005). Although humans have incorporated and elaborated the use of quantities in numerical symbolic systems over thousands of years, this was only possible due to an adaptive skill, naturally selected, named approximate number system (ANS). The ANS is an innate cognitive system shared between humans and other animals, wherein quantities are processed in a non-symbolic, analogical and amodal way. In the wild, this mechanism is important for animals to survive, helping them to optimize in situations where choices based on quantities are required, for example to catch or reproduce (Dehaene, 1992, 2011).

Initially, ANS is useful for infants to perform estimations about the quantity of objects in a collection and to solve simple arithmetic problems by approximation. Also, the acquisition of the cardinal principle (i.e., the ability to recognize that the final number said, when counting a group, is equivalent to the total of items in a set) may be facilitated by increased ANS acuity (Shusterman et al., 2006). ANS might be the foundation for symbolic arithmetic and although it continues to be a useful system for later arithmetic performance, for example in online error detection (Park & Brannon, 2013, 2014; Szkudlarek & Brannon, 2017), when it comes to the resolution of more complex problems, symbolic processing is highly required.

Different from ANS, symbolic arithmetic is a cultural artifact that evolved about five thousand years ago, which allows the performance of simple arithmetic operations in a precise way. The abstract nature of symbolic representations of numbers makes the process of acquiring symbolic arithmetic difficult. The acquisition of symbolic arithmetic is an arduous and time-consuming process, which requires years of effort by the child (Lehtinen et al., 2017).

Finger-based representations (FBRs) play an important role mapping between symbolic and nonsymbolic magnitude representations in early development. Wasner and colleagues (2015) differentiated three types of FBRs: ordinality, cardinality and correspondence one-to-one. While ordinality is related to finger counting habits, cardinality is the most common way for montring the whole number. On the other hand, correspondence one-to-one is the association between a specific number and the finger used to represent it. There is evidence that these abilities are related but are recruited independently from each other, according to the numerical information to be transmitted (Lafay et al., 2013; Wasner et al., 2015).

Children of all investigated cultures use Finger-Counting Strategies (FCS) during some point of their development for calculating (Butterworth, 1999). Fingers are natural tools and provide a multi-sensory input that transmits concrete information of the cardinal and ordinal numbers and help to understand the principles of counting and the base ten number system (Beller & Bender, 2011; Moeller et al., 2011). Evidence suggests that FCS are more important when children are learning new operations and solving those that were not automated (Crollen & Noel, 2015; Jordan et al, 2008; Reeve & Humberstone, 2011). In these situations, children need to allocate scarce processing resources between coordinating the verbal-numerical correspondence and keeping track of the procedure. FCS provides concrete representations with one-to-one correspondence, reducing working memory load of a complex task, in terms of an offloading support. (Alibali & DiRusso, 1999; Costa et al., 2011).

Geary and Hoard (2005) proposed a model that summarizes the trajectory of strategies used by children to solve simple arithmetic problems. According to the authors, children initially use counting-based strategies that are gradually replaced by memory-based strategies. Strategies are becoming increasingly more effective, as they require few cognitive resources and response time. Children start using their fingers to perform counting-based strategies. First, they count all the elements of the sets (counting all), then start the counting from the cardinality of the first set (counting on) and next, they learn to count starting from the biggest set (counting min; Geary & Hoard, 2005; Siegler, 1999).

As children acquired certain proficiency in finger counting, they gradually progress to verbal counting, using the three previous cited counting strategies (counting all, counting on and counting min). Finally, memory-based strategies emerge, first children rely in mental manipulation of the FBRs, progressing to decomposition strategy (i.e. use a previously known fact to solve a new problem (e.g. 14+7, if 13+7=20, so (13+1)+7=21) and at last, they advance for the direct retrieval of the correct response from long-term memory. Although the expected progression across strategies, this development happens like overlapping waves, as more than one strategy are often used at the same period for solving different problems and learning a new strategy does not mean completely stop using the previous one (Siegler, 1999).

The influence of FBRs in calculations can be observed in different levels of strategies. Initially, children use more procedural counting-based strategies, often externally representing the operands with their hands with one-to-one correspondence (Geary & Hoard, 2005). Later, when children are learning to implement memory-based strategies, they may continue relying on the FBRs, using the five-pattern of the full hand for supporting the mental manipulation (Geary & Hoard, 2005). An example of the influence of FBRs in mental calculation was observed by Thompson (1999), in a report about the strategy used by a child, for correctly solving the problem 6 + 7: "13…I took 5 out of the 6 and 5 out of the 7 and I was left with 3…".

Although some children use their fingers to solve operations with different types of complexity, the strategy need to be adapted depending on the demands of the arithmetical problem. Initially for solving simple addition and subtraction problems with results below or equal 10, children take advantage from the correspondence one-to-one propriety of fingers, concretely manipulating the operands externally with their fingers. However, for more complex problems, with results or operands above 10, children need to adapt their strategies, by reusing a full hand

and keeping track of how many hands were reused to represent the two-digits results. Some children develop coping strategies for using their fingers in complex problems, for example touching a surface with one finger for each reused full hand (Fuson and Kwon, 1992). Other children face difficulties by keeping track of the number of reused full hands and commit errors deviating from the correct by five (split-five errors, Domahs et al., 2008).

Typically, the transition from counting-based strategies to more mature memory-based strategies, such as retrieving the correct answer from long term memory, starts at the end of first grade or beginning of second grade (for a review see Soylu et al., 2018). The ability of implementing the strategies of counting on and decomposition mark two important steps on children's development of quantitative reasoning. The use of counting on strategy is only possible when children acquire the numerical concept of cardinality, understanding the additive composition (Fritz et al, 2013). On the other hand, the use of decomposition strategy requires the understanding of the reversibility between addition and subtraction operations and the Part-Whole concept (Langhorst et al., 2012).

Children with mathematical learning difficulties usually starts using FCS later and use this strategy during more time in comparison with their pairs (Butterworth, 2019). Although the trajectory of numerical development being qualitative the same for all children, those with math difficulties have a delayed trajectory, requiring more time and support to progress in math learning (Balt et al., 2020). Even when these children acquire conceptual knowledge, they face difficulties in integrating it into their strategies, persistently relying on counting all strategy (Balt et al., 2020).

The role of FBRs in calculations go beyond learning a new operation and using FCS, being also observed in mental calculation (Klein et al., 2011; Thompson, 1999). Klein and coworkers (2011) found that educated adults, when solving simple addition problems, took more time to solve problems that involve a break 5 (both operands lower than 5 and the sum of them resulting in more than 5, e.g. 4+2; 3+4), than would be expected. This shows evidence of an implicit sub-base-five, that is the exact number of fingers that we have one hand, being part of our numerical base-ten system (Klein et al., 2011).

There is also evidence of sub-base-five in children. Domahs and colleagues (2008) demonstrated the occurrence of a specific type of error by the difference of five between the child answer and the correct response, named split-five (S5). The results showed that this error happens mainly in complex operations of addition and subtraction, when children cannot represent all operands in their hands and need to keep track of how many hands were already counted. Additionally, an unexpected prevalence of S5 errors was observed in simple addition by the end of second grade, when children were learning multiplication. Besides this, a split-ten error effect was also observed in complex operations as a combination of two S5 errors.

Domahs and colleagues (2008) observed that girls committed significantly more S5 errors than boys. However, the authors did not explore the meaning of this result. Although some studies found a difference between boys and girls in math performance (Dwyer, 1975), it is not a consistent result (for a review see Hyde, 2016). A study investigating implicit attitudes demonstrated that an association between men and better performance in mathematics is at an unconscious level, and the same author observed that this stereotype is present in different cultures (Nosek et al. 2002; 2009).

Domahs and colleagues (2008) considered that different stages of abstractness for solving the problems can be influenced by FBR. First, when fingers are used for performing the procedure, there is a mental storage of how many hands (or five-patterns) were reused. Second, when mental calculation is performed with internal procedure based on previous finger manipulation, there is also a storage of how many five-patterns were reused. And third, when children can retrieve the response from long-term memory, the mental representations carry some characteristics of finger representations, in particular the five-pattern. The FBRs influence on the different stages could be observed by a disproportionate percentage of errors with the exact difference of five or multiples of five from the correct result.

Besides Domahs and colleagues (2008), to our knowledge Di Nuovo (2018) was the only other study reporting the S5 effect. Di Nuovo (2018) integrates as the artificial intelligence of the humanoid robot named *iCub*, a modern architecture named Long-Short Term Memory (LSTM) which effectively model the working memory for investigating the execution of simple arithmetic operations of addition. When a finger-counting based system was introduced on the model, not only the accuracy in the single-digit addition increased, but also a significant S5 effect was observed.

Present Study

To our knowledge, until now, Domahs and colleagues (2008) was the only study investigating if the FBRs influenced the later mental calculation in children. Domahs and colleagues investigated the S5 effect in a longitudinal study that followed one hundred and fifty children from the end of first grade, until the middle of third grade. This study aims to continue the specific investigation of the S5 effect. At first, we aim to replicate the cross-sectional findings regarding the occurrence S5 effect. Moreover, previous findings about the association between FBRs and calculation will be further expanded through the analysis of whether this effect occurs for multiplication operations and by exploring possible associations of this error with learning difficulties in mathematics.

For this purpose, first the performance of Brazilian children from 1st to 5th grade in basic arithmetic operations of addition, subtraction and multiplication will be explored. Gomides and colleagues (2021) demonstrated the difficulty of Brazilian children from 3rd to 5th grade to master fluency in the basic arithmetic operations. Results in this study will be expanded for 1st and 2nd graders. In addition, we will explore the Brazilian children's performance on arithmetic calculation considering the accuracy and not only the fluency, hypothesizing that the difficulties of Brazilian children are more related to a delay in the progression from procedural strategies to retrieval strategies than to the accuracy.

Domahs and colleagues (2008) investigated the S5 effect in addition and subtraction, grouping the results for simple and complex problems. Since children can use different strategies at the same developmental period for different operations, it is important to examine the S5 effect separately for addition and subtraction. Additionally, we aim to go further and investigate if there is a S5 effect in multiplication. Although it is known that the most common strategy used for solving multiplication problems is fact retrieval (Dehaene, 1992), Brazilian children start learning multiplication in 3rd grade and most of them have still not mastered it in 5th grade (Gomides et al., 2021), so they may produce S5 errors due to the use of procedural strategies based on finger-patterns.

Another specific goal of this study is to evaluate the developmental pattern in S5 effect. As the S5 effect is most expected when children are implementing strategies that rely on FBRs, and children use these strategies more often in early years, it is hypothesized that S5 effect will be more produced by younger children. However, as children progress in strategies at different times and continue to use procedural strategies even when they have progressed to using more retrieval strategies (Siegler, 1999), S5 effect may also occur in older children.

Additionally, another association that we aim to explore sex differences regarding the S5 effect. Even though Domahs and colleagues (2008) demonstrated that girls committed significantly more S5 errors than boys, the authors did not discuss these results. As sex differences in math performance are not consistently observed in the literature (Hyde et al. 2016), sex differences are not expected regarding S5 errors.

Beyond S5 effect, the well-known Split effect will be also investigated. Split effect is usually demonstrated in verification tasks, in which participants are asked to verify if the response presented for arithmetical operation is correct or not. The Split effect is observed by a significant decrease of response time, the greater the distance between the presented incorrect and the correct results (Ashcraft & Stazyk, 1978) and could be interpreted as a distance effect, reflecting the imprecision of numerical magnitude representations in the mental number line (Ashcraft & Fierman, 1982). In the present study, the Split effect will be investigated in a fluency task, it is expected higher frequencies as smaller the distance between child's answers and the correct result. It is expected that children who do not commit S5 errors present a more significant Split effect. As demonstrated by Domahs and colleagues (2008), children who commit S5 errors are less prone to commit small splits errors (differentiating by 1 or 2 from the correct answer). Moreover, Split effect is more expected for addition and subtraction because multiplication may be solved by retrieval strategies.

As mentioned before, children with math learning difficulties are more prone to persist in procedural strategies presenting difficulties to progress for retrieval strategies. As the effect of the S5 error seems to be associated to the use of strategies that rely on FBRs both by externally counting or internally mental manipulation, probably children that commit this type of error are more likely to present learning difficulties in mathematics. At last, this study also aims to investigate whether there is a difference in math achievement between children who commit and those who do not commit S5 error.

Methods

Participants

The initial sample included 1288 pupils attending 1st to 5th grade in either Belo Horizonte or Porto Alegre, Brazil. Of these 128 children were excluded from analyses because of a neuropsychiatric diagnosis (3 children with autism, 1 child with cerebral palsy), history of grade retention (25 children), and low non-verbal reasoning (99 children with performance below percentile rank 10 in Raven's Coloured Progressive Matrices, henceforth RCPM, Raven et al., 2018). The final sample was composed of 1160 children (Age range: 5-11 years; Mean age=8.20 years, DP=1.53 years; Girls=52.70%; Belo Horizonte=87.2%). For the analyses, we only considered children that responded to at least one item and committed at least one error in the specific block analyzed (Descriptive analyses of participants' characteristics are available in Supplementary Material A).

Procedures

Data were collected over 11 years from different research projects conducted at the Laboratório de Neuropsicologia do Desenvolvimento (LND-UFMG). Children had to complete different task batteries, assessing i) general cognitive abilities (e.g. non-verbal reasoning, working memory) as well as ii) numerical-arithmetical abilities (e.g., numerical magnitude comparison, basic arithmetic operations) and iii) math achievement. All tests were administered in quiet rooms in the participants' schools. For the purpose of the present study, only data of non-verbal reasoning (measured by RCPM), math achievement (assessed by the arithmetic subtest of Brazilian School Achievement Test, henceforth TDE, Stein, 1994) and a test on basic arithmetic operations (i.e., addition, subtraction and multiplication: Basic Arithmetic Operations Task, henceforth BAOT, Gomides et al., *in press*) were considered. The RCPM and TDE were assessed in groups of up to five children whereas the BAOT was administered individually. None of 1st graders were not submitted to TDE, and 2 children from 3rd grade did not complete the TDE, furthermore, data from these children will not be included in math achievement analysis.

The projects were conducted in accordance with the principles described in the Declaration of Helsinki for research involving human participants and were approved by the Research Ethics Committees of the Universidade Federal de Minas Gerais (ETIC 42/08, CAAE 15070013.1.0000.5149, and CAAE 21361213.0.0000.5149) and Universidade Federal do Rio Grande do Sul (Protocol number 1.023.371). Participation was voluntarily and required written informed consent from parents or caregivers and oral assent from children prior to testing.

Instruments

Raven's Coloured Progressive Matrices (RCPM): The RCPM was used to assess general cognitive abilities in terms of non-verbal reasoning ability. In this task, children have to complete 36 patterns by a missing part, which they have to pick from six response options. The Brazilian validated version was used (Raven et al., 2018). Analyses were based on z-scores calculated from the manual's norms for the respective age group.

Brazilian School Achievement Test (Teste de Desempenho Escolar - TDE; Ferreira et al., 2012, Stein, 1994): The TDE is the most widely used standardized test of school achievement with norms for the Brazilian population. It comprises three subtests: arithmetic, single-word spelling and single-word reading. In the present study, only data from the arithmetic subtest will be considered. The arithmetic subtest is composed of three simple verbally presented word problems (e.g., Which number is larger, 28 or 42?) and 45 written arithmetic calculations of increasing complexity (e.g., very easy: 4-1; easy: 1230+150+1620; intermediate: 823*96; hard: 3/4+2/8). Children are instructed to work on the problems to the best of their capacity but without a time limit. One point is awarded for each problem solved correctly. Analyses were based on z-scores calculated from the sample's norms.

***Basic* Arithmetic Operations Task (BAOT):** The BAOT is part of the PRONÚMERO battery of tests (Costa et al., 2011, Gomides et al., *in press*). The test aims to evaluate children's fluency in solving addition, subtraction, and multiplication problems. The BAOT is subdivided into six blocks: i) Simple Addition including 12 problems with results smaller than 10 (e.g. 4+3); ii) Complex Addition including 15 problems involving carry over with the result between 11 and 17 (e.g. 5+8); iii) Simple Subtraction including 12 problems with both operands smaller than 11 (e.g. 10-8); iv) Complex Subtraction including 15 problems involving borrow operation with the minuend ranging from 11 to 17 and the subtrahend ranging from 3 to 9 (e.g. 13-8); v) Simple Multiplication including 15 problems with the result below 25 or with 5 being one of the operands (e.g. 2x7, 6x5); vi)

Complex Multiplication including problems with the result ranging from 24 to 72 (e.g. 8x7). No negative results were not used in subtraction problems and no tie problems were included in the BAOT overall.

Problems of each operation were presented horizontally on a sheet of A4 size paper in fixed order. Children were instructed to work as fast and as accurately as possible, writing the answer on the paper in the Arabic digit format. Different protocols of BAOT were applied, depending on grade level: i) 1st graders were administered the Simple Addition and Simple Subtraction subtasks only with a time limit of 3 minutes each; ii) 2nd graders were administered both (simple and complex) Addition and Subtraction blocks again with a time limit of 3 minutes each; iii) 3rd to 5th graders were administered all six blocks (i.e., Simple Addition, Complex Addition, Simple Subtraction, Complex Subtraction, Simple Multiplication and Complex Multiplication) with a time limit of 1 minute per block.

The validity's evidence of BAOT were analyzed and presented on PRONÚMERO (Gomides et al., *in press*), separating the blocks per operation, leading to three subtests: Addition, Subtraction and Multiplication. Regarding the dimensionality of BAOT, Exploratory Factor Analysis indicates two factors for Addition and Subtraction subtests, explaining respectively 58% and 61%, and only one factor for Multiplication, explaining 60% of the total variance. In all three subtests, some items were redundant, presenting high polycolic correlation (>0.97), but because of their theoretical importance on BAOT, they were not excluded from the final version of the task. Also, the internal consistency, assessed using the Kuder-Richardson coefficient (Kurder & Richardson, 1937, KR-20), showed that all subtests had satisfactory reliability index (Addition=0.94; Subtraction=0.93; Multiplication=0.95).

Statistical analyses

Statistical analyses were performed separately for each of the six BAOT blocks (Simple Addition, Complex Addition, Simple Subtraction, Complex Subtraction, Simple Multiplication, Complex Multiplication), divided in five sections:

i) Basic arithmetic performance

First, percentage of children that did not respond to any item were described. Data from these children were excluded from further analyses. Then, the percentage of unresponded items (for sample level), were described according to Formula 1 (percentage of correct and incorrect answers were also calculated and are available in Supplementary material B). Next, the rate of children that solved all responded items incorrectly and children that solved all responded items correctly were described (illustrative representations of these results are available in Supplementary material B). It is important to emphasize that as BAOT is a fluency task, only the responded items of each block were considered for calculating children that responded to all items incorrectly or correctly, not the block's total of items (for example, if the child answered 4 items of the Simple Addition block, it was considered that they responded 100% correct when they these 4 items correctly).

Formula 1. Calculation of unresponded items percentage in each BAOT block.

$$\text{Percentage of unresponded items in the BAOT block} = \frac{\text{N of unresponded items in the BAOT block}}{\text{N of items in the BAOT block}^{*} \times \text{N of children who responded the BAOT block}} \times 100$$

Next, to control for influences of the differing time limits of BAOT for the grades (1st and 2nd graders had a time limit of 3 minutes, while 3rd to 5th graders had a time limit of 1 minute) the percentage of incorrect answers was calculated for each child considering only the responded items according to Formula 2. Finally, differences in incorrect answers across grades were analyzed using one-way ANOVA and sex differences were analyzed by t-test.

Formula 2. Calculation of errors percentages in responded items for each child in each BAOT block.

$$\text{Percentage of errors in responded items in the BAOT block} = \frac{\text{N of errors in the BAOT block}}{\text{N of responded items in the BAOT block}} \times 100$$

Next sections will present the Split Error Analyses. For this purpose, the percentages (relative frequency) of split errors from S1 to S>11 were calculated. At first, the absolute split error of incorrect answers in each responded item was calculated (Formula 3: |child answer - correct result|). Absolute split error can be defined as the distance between the child answer and the correct result, regardless of whether the erroneous result was smaller or larger than the correct result (e.g., in the problem 3 + 4 = 7, answers of 6 or 8 are considered split-one; answers of 5 or 9 are considered split-two, and so on). Secondly, the frequency of occurrence for each split error was calculated (n of errors for each split category) for each child. Split errors ranging from 1 to 11 were considered as separate categories (S1 to S11) and split errors above 11 were considered as a single category (S>11), as done by Domahs and colleagues (2008). At last, split errors per category were used to calculate the percentage (relative frequency) of split errors from S1 to S>11 by dividing the number of split errors in each category by the total of incorrect answers for each child according to the Formula 4. It needs to be noted that the Simple Multiplication block had 7 items excluded from the Split Error Analyses (sections ii, iii, iv and v) because they had 5 as operands, so S5 errors were ambiguous as it also reflects operand errors (i.e. error by any multiple of one of the operands).

Formula 4. Calculation of split error percentages (S1 to S>11) for each child in each BAOT block.

$$\text{Percentage of each Split Error}_{1,2,3\ldots,>11} = \frac{\text{Split Error}_{1,2,3\ldots,>11}}{\text{N of incorrect answers in the BAOT block}} \times 100$$

The following analyses were performed considering only children that responded to at least one item and committed at least one error in the specific BAOT block.

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

The percentage of children who committed at least one S5 error (classified as GS5) were described and analyzed by Chi-square, to investigate whether there were differences in the distribution across grades and between sexes. As the number of children in GS5 in some grades is small, Chi-square analyses were performed with Monte Carlo correction. Further analyses were performed separately for the GS5 and for the group of children that did not commit S5 errors (GNoS5; descriptive characteristics analyses for both groups are presented in Supplementary material A.

iii) Investigation of Split effect in the group that did not commit S5 error (GNoS5)

To investigate the Split effect (i.e., the inverse relationship between the frequency of errors and the distance/split between the incorrect answer and the correct result), for GNoS5 repeated-measures ANOVAs were performed comparing the frequency of S1 to S4 errors. Only S1 to S4 were tested because these are the split errors with the highest expected frequency. S5 effect analysis was not included because GNoS5 did not commit any S5 error (Means/standard deviations of each Split error percentage ranging from S1 to S>11 for GNoS5 are presented in Supplementary material C).

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

To investigate the Split effect for GS5 repeated-measures ANOVAs were performed comparing the frequency of S1 to S4 errors. Next, to investigate the S5 effect, repeated-measures ANOVAs were performed comparing the percentage of occurrence of S5 errors to the percentage of its adjacent errors, S4 and S6 (Mean/standard deviation of Splits ranging from S1 to S>11 for GS5 are presented in Supplementary material C).

v) Comparison of math achievement between GS5 and GNoS5.

For comparing the math achievement between GS5 and GNoS5 a subsample of children who completed the TDE was considered. Potential differences in math achievement between children who commit and did not commit an S5 error were compared using an independent sample t-test on the z-score of TDE arithmetic subtest calculated by grade, for controlling the difference of performance across grade levels.

Results

Each one of six BAOT's blocks (i.e., Simple Addition, Complex Addition, Simple Subtraction, Complex Subtraction, Simple Multiplication and Complex Multiplication) will be considered separately in five steps of analyses: i) Basic arithmetic performance, ii) Frequency of children who committed at least one S5 error comparing across grades and sexes, iii) Investigation of Split effect in the group that did not commit S5 errors (GNoS5), iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5), v) Comparison of math achievement between GS5 and GNoS5.

Simple Addition block

i) Basic arithmetic performance

The Simple Addition block was the easiest one for children in all grades. The percentage of 1st graders who did not respond to any item was 17.21%, and this rate was lower than 6% in 2nd to 5th graders. In 1st and 3rd grades, around 20% of items remained unresponded, while in 2nd, 4th and 5th grade this rate is lower than 10%. Another evidence of the ease of this block is the low percentage (<3%) of children who responded to all items incorrectly in all grades. Notably the highest rate was found in 4th grade (2.53%) and none of the 5th graders responded to all items incorrectly.

Next, the percentage of incorrect answers based on responded items for Simple Addition was analyzed comparing grade levels and sex. One-way ANOVA resulted in significantly incorrect response rates across grades, $F(4, 1088)=63.08$, $p<.001$, $\eta_p^2=.20$. Bonferroni corrected pairwise comparisons indicated higher incorrect response rates for 1st graders in comparison with all other grades. Also, 2nd graders had a higher percentage of incorrect response, in comparison with 4th and 5th grade, but not in comparison with 3rd grade. Additionally, there was no significant difference in the percentage of incorrect responses between 3rd, 4th, and 5th graders. The median error percentage from 2nd to 5th grade is zero (see Figure 1.1). No sex differences were observed in the rate of incorrect response, $t(1093)=1.00$, $p>.05$, Cohen's $d=0.06$.

Considering the accuracy in the responded items, 20.22% of 1st graders responded to all items correctly, while more than 50% of 2nd to 5th graders had 100% of accuracy, answering all responded items correctly. These children's data will not be considered, resulting in a total of 405 children (37.05%) for further analyses.

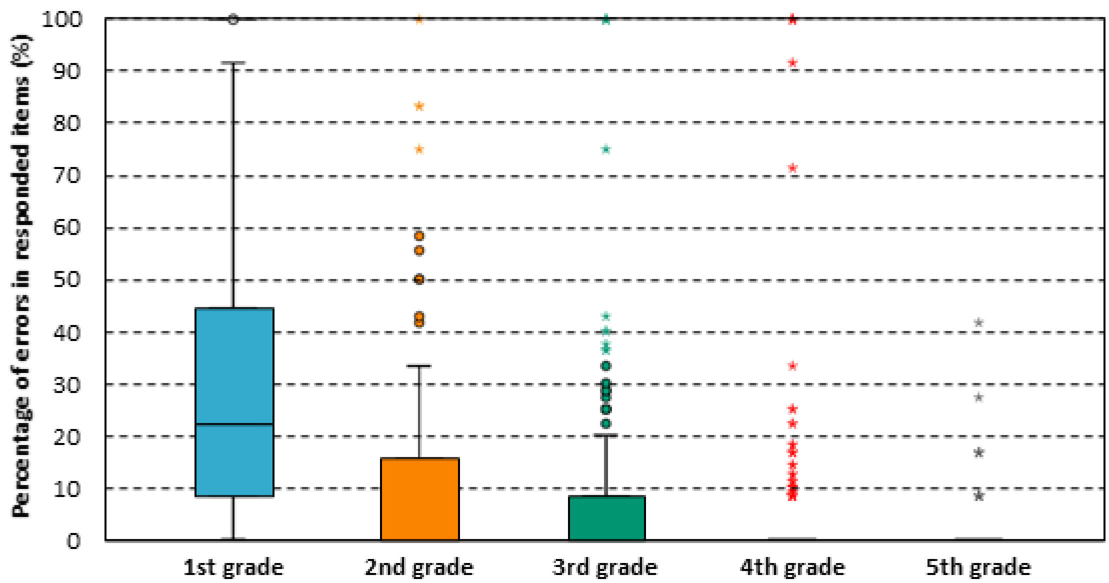

Figure 1.1. Distribution of incorrect responses in the Simple Addition block, stratified by grade.

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

From 405 children selected for the Split Error Analysis in the Simple Addition, 13.09% (n=53) made at least one S5 error. The percentages of children from GS5 decreased across grades, except for 4th graders: 1st graders (n=29, 20.42%), 2nd graders (n=6, 9.84%), 3rd graders (n=6, 7.32%), 4th graders (n=10, 11.63%), 5th graders (n=2, 5.88%). A Chi-square test was conducted to evaluate the distribution of GS5 across grades. There was a significant difference in the distribution of children who commit S5 errors across grades, $\chi^2(4)=11.40$, $p<.05$. Adjusted residuals analysis indicated that 1st graders were 2.35 times more probable to commit S5 errors than 2nd graders, 3.25 times more than 3rd graders, 1.95 times more than 4th graders, and 4.11 times more than 5th graders.

A Chi-square test of independence was performed to examine the distribution of GS5 between sexes in the Simple Addition block. There were no significant differences in the distribution of children who commit S5 errors between sexes, $\chi^2(1)=0.35$, $p>.05$.

iii) Investigation of Split effect in the group that did not commit S5 error (GNoS5)

As shown in Figure 1.2, GNoS5 presented an overall tendency decrease in the frequency of errors, as the distance from the correct response increases (i.e., Split effect). A repeated-measures ANOVA of error frequency by category of split error from S1 to S4 was significant, $F(2.01, 703.91)=255.86$, $p<.001$, $\eta_p^2=.42$. Bonferroni pairwise comparisons indicate that S1 frequency was the highest and S4 frequency was the lowest one. However, there was no difference between S2 and S3 frequencies.

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

The Split effect was also observed in GS5, but with lower effect size $F(2.26, 117.74)=3.32$, $p<.05$, $\eta_p^2=.06$. Bonferroni pairwise comparisons indicate that S2 frequency was higher than S4 frequency. However, there was no difference among S1 and S3 frequencies compared to each other and to S2 and S4 frequencies. In GS5, the most notable tendency was the S5 effect. A repeated-measures ANOVA comparing S5 frequency with its adjacent errors, S4 and S6, was significant, $F(1.21, 62.72)=57.08$, $p<.001$, $\eta_p^2=0.52$. Bonferroni pairwise comparisons indicated that the frequency of S5 errors was significantly higher than the frequency of S4 and S6 errors whereas the latter two did not differ in frequency (see Figure 1.2).

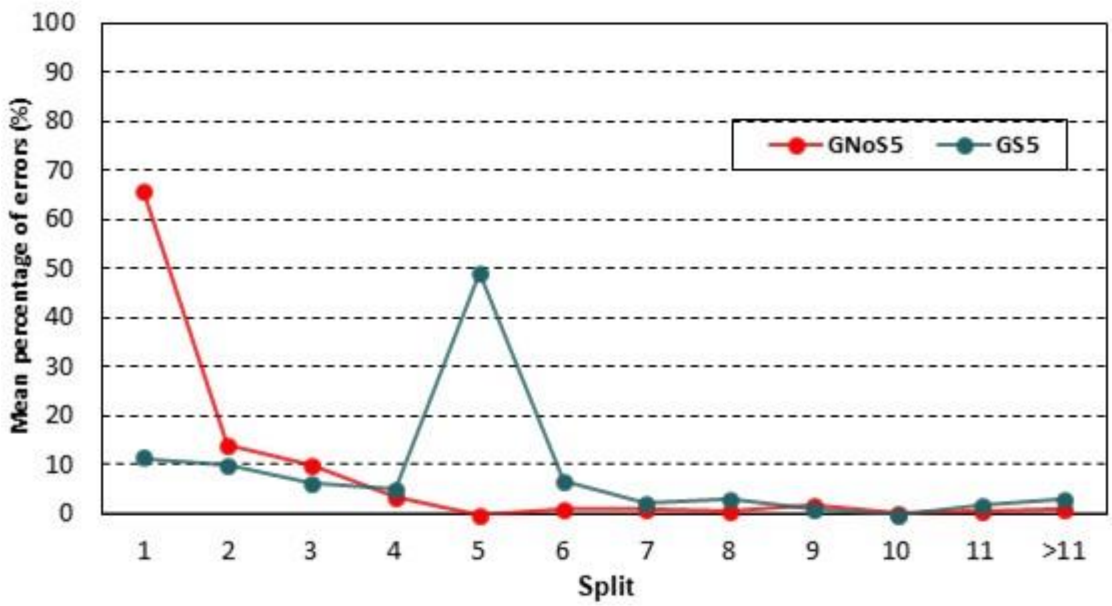

Figure 1.2. Mean percentage of Splits errors ranging from S1 to S>11 in the Simple Addition block for the GNoS5 and GS5.

v) Comparison of math achievement between GS5 and GNoS5

In a subsample, data were available that allowed the comparison of math achievement between groups that committed and did not commit any S5 error in the Simple Addition block. A t-test was conducted to investigate if there is a difference in TDE subtest of arithmetic between GNoS5 (n=239; mean=-0.36; sd=0.87) and GS5 (n=24; mean=-0.26; sd=0.97). The result showed no significant difference between groups for math achievement, $t(261)=-0.53$, $p>.05$, Cohen's d=0.11.

Complex Addition

i) Basic arithmetic performance

In the Complex Addition block, the percentage of 2nd to 5th graders who did not respond to any item was lower than 6.50%. In 2nd grade and 5th grade, less than 25% of items remained unresponded, while in 3rd grade and 4th grade this rate was higher, with respectively 56.74% and 34.40% of unresponded items. However, the percentage of children who responded to all items incorrectly is low (<5%) for all grades.

Next, the percentage of incorrect answers based on responded items for the Complex Addition block was analyzed comparing grade levels and sex. One-way ANOVA resulted in significantly differing rates of incorrect response across grades, $F(3, 907)=30.59$, $p<.001$, $\eta_p^2=.09$. Bonferroni corrected pairwise comparisons indicated higher incorrect response rates for 2nd graders in comparison with 3rd, 4th, and 5th graders. Also, 3rd graders had a higher percentage of incorrect responses, in comparison with 5th graders. There was no difference in the percentage of incorrect responses between 4th and 5th graders. The median error percentage from 3rd to 5th grade is zero (see Figure 2.1). No sex differences were observed in the rate of incorrect response, $t(821.844)=1.59$, $p>.05$, Cohen's $d= 0.11$.

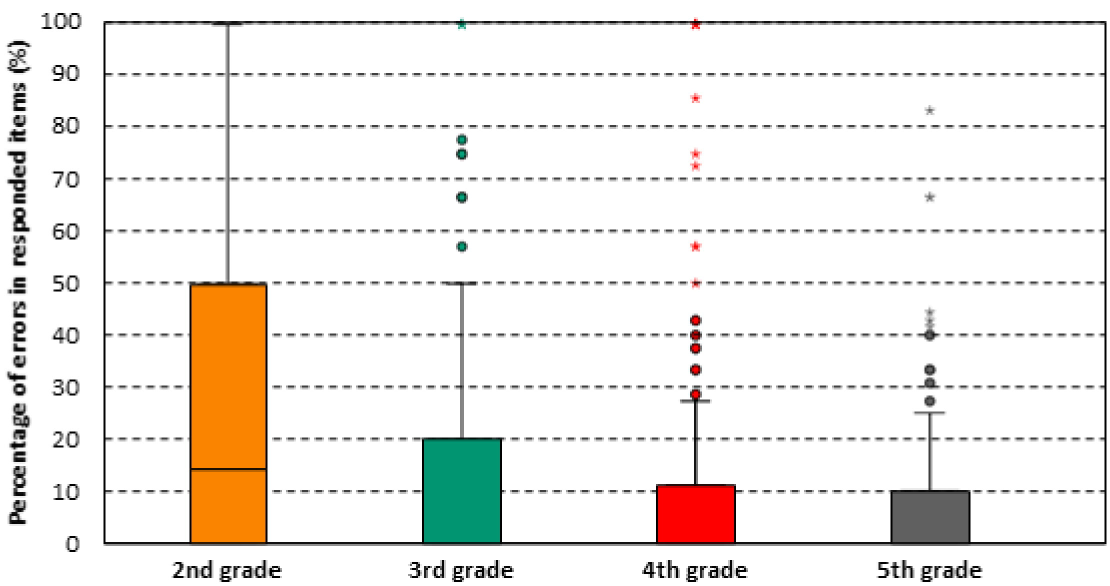

Figure 2.1. Distribution of incorrect responses in the Complex Addition block, stratified by grade.

Considering the accuracy on the responded items, 28.35% of 2nd graders responded to all items correctly. More than 50% of 3rd to 5th graders had 100% of accuracy, answering all responded items correctly. These children's data will not be considered, resulting in a total of 393 children (43.14%) for further analyses.

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

From the 393 children selected for the Split Error Analysis in the Complex Addition, 16.54% (n=65) made at least one S5 error in the Complex Addition block. The percentages of children from GS5 decreased across grades: 2nd graders (n=29, 31.87%), 3rd graders (n=18, 16.82%), 4th graders (n=15, 11.45%), 5th graders (n=3, 4.69%). A Chi-square test was conducted to evaluate the distribution of GS5 across grades. There was a significant difference in the distribution of children who commit S5 error across grades, $\chi^2(3)=24.47$, $p<.001$. Adjusted residuals analysis indicated that 2nd graders were 2.31 times more probable to commit S5 errors than 3rd graders, 3.62 times more than 4th graders and 9.51 times more than 5th graders. Otherwise, 5th graders were 4.11 times less probable to commit S5 errors than 3rd graders and 2.63 times less than 4th graders.

A Chi-square test of independence was performed to examine the distribution of GS5 between sexes in the Complex Addition block. There were no significant differences in the distribution of children who commit S5 error between sexes, $\chi^2(1)=0.06$, $p>.05$.

iii) Investigation of Split effect in the group that did not commit S5 (GNoS5)

Further analyses will be performed separately for GS5 and GNoS5. As shown in Figure 2.2, GNoS5 presented an overall tendency of decrease in the frequency of errors, as the distance from the correct answer (or split) increases i. e. a Split effect. A repeated measure ANOVA of error frequency by category of split error from S1 to S4 was significant, $F(1.91, 624.23)=300.56$, $p<.001$, $\eta_p^2=.48$. Bonferroni pairwise comparisons indicate that S1 frequency was higher than S2, S3 and S4 frequencies. S2 frequency was higher than S3 and S4 frequencies. However, there was no difference between S3 and S4 frequencies.

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

The Split effect was also observed in GS5, but with lower effect size $F(2.20, 141.05)=8.80$, $p<.001$, $\eta_p^2=.12$. Bonferroni pairwise comparisons indicate that S1 frequency was the highest one. However, there was no difference among S2, S3 and S4. In GS5, the most notable tendency is the S5 effect. A repeated measure ANOVA of error frequency by category comparing S5 frequency with its adjacent errors, S4 and S6, was significant $F(1.20, 77.03)=106.81$, $p<.001$, $\eta_p^2=0.63$. Bonferroni pairwise comparisons indicate that S5 frequency was higher than S4 and S6 frequency. However, there was no difference between S4 and S6 frequencies (see Figure 2.2).

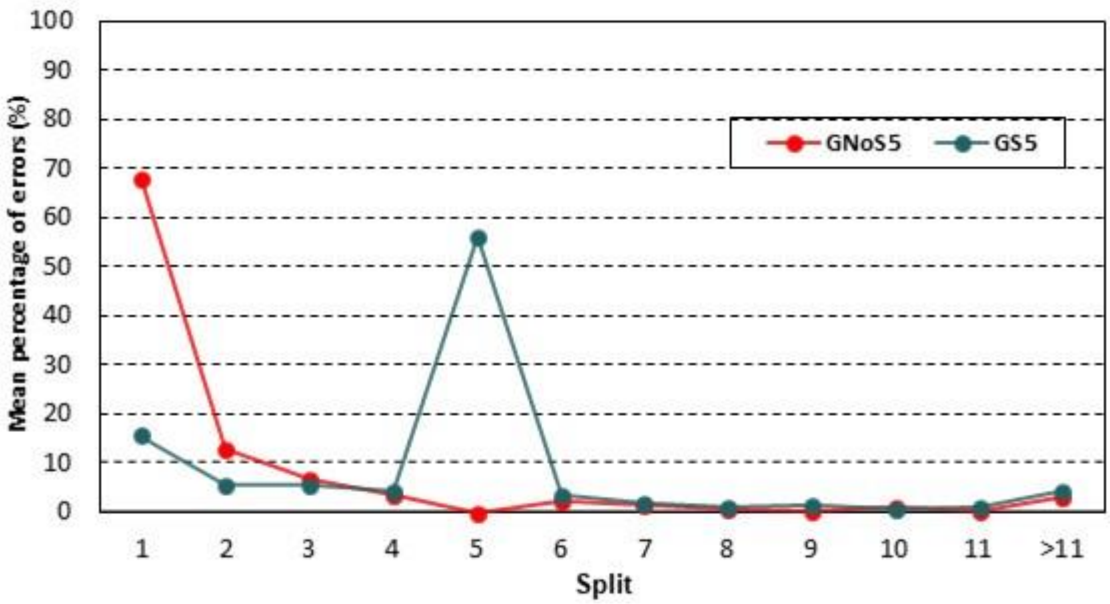

Figure 2.2. Mean percentage of Splits errors ranging from S1 to S>11 in the Complex Addition block for the GNoS5 and GS5.

v) Comparison of math achievement between GS5 and GNoS5

In a subsample, data were available that allowed the comparison of math achievement between groups that committed and did not commit any S5 error in the Complex Addition block. A t-test was conducted to investigate if there is a difference in TDE subtest of arithmetic between GNoS5 (n=328; mean=-0.15; sd=0.95) and GS5 (n=65; mean=-0.50; sd=0.75). The result showed that GS5 had a significant lower score in math achievement in comparison with GNoS5, t(391)=2.81, p<.01, Cohen's d=0.41.

Simple Subtraction

i) Basic arithmetic performance

In the Simple Subtraction block, a high rate of 75.81% of 1st graders did not respond to any item, and this rate was lower than 10% for 2nd to 5th graders. In 1st and 4th grade, around 20% of items remained unresponded, while in 2nd and 5th grade this rate was around 10% and in 3rd grade this rate was 40.61%. In 1st grade 13.46% of children responded to all items incorrectly, while in 2nd to 5th grade this rate is lower than 10%.

Next, the percentage of incorrect answers based on responded items for the Simple Subtraction block was analyzed comparing grade levels and sex. One-way ANOVA revealed a significant effect of grade level on the rate of incorrect responses, F(4, 953)=29.63, p<.001, η_p^2=.11. Bonferroni corrected pairwise comparisons indicated higher rates of incorrect responses for 1st and 2nd graders in comparison with 3rd, 4th and 5th graders. Also, 1st graders committed more incorrect responses than 2nd graders. Moreover, 3rd graders had a lower percentage of errors as compared to 4th graders, but not in comparison to 5th graders. There was no difference in the percentage of errors between 4th and 5th graders. The median error percentage from 3rd to 5th grade is zero (see Figure 3.1). No sex differences were observed in the rate of incorrect response, t(875.202)=1.60, p>.05, Cohen's d= 0.10.

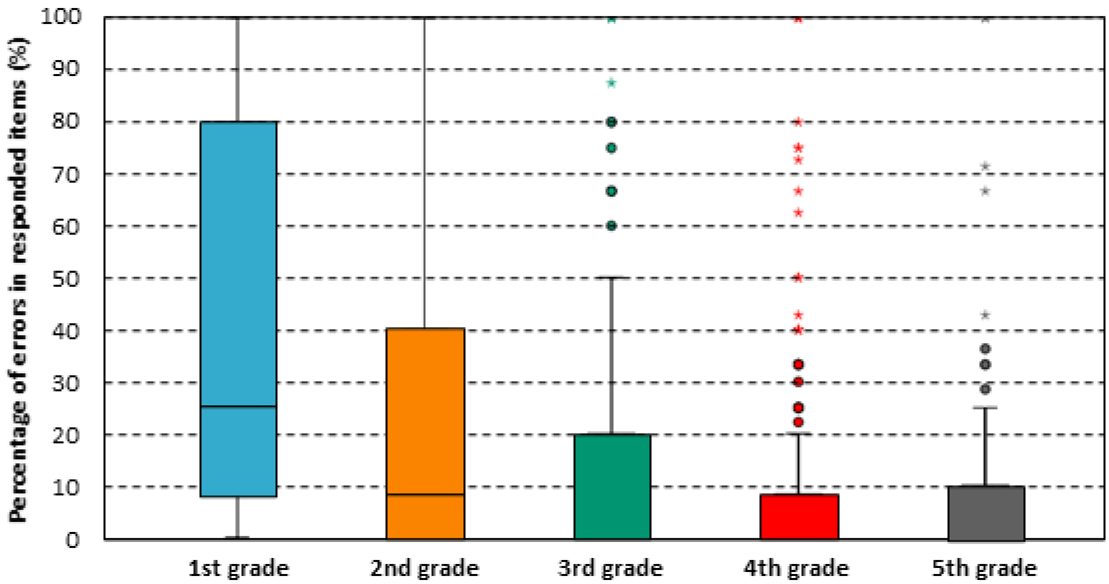

Figure 3.1. Distribution of incorrect responses in the Simple Subtraction block, stratified by grade.

Considering the accuracy on the responded items, 23.08% of 1st graders and 38.02% of 2nd graders responded to all items correctly. More than 50% of 3rd to 5th graders had 100% of accuracy, answering all responded items correctly. These children's data will not be considered, resulting in a total of 420 children (43.84%) for further analyses.

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

From 420 children selected for the Split Error Analysis in the Simple Subtraction block, 14.52% (n=65) made at least one S5 error. The percentages of children from GS5 decreased across grades: 1st graders (n=9, 22.50%) 2nd graders (n=17, 22.67%), 3rd graders (n=15, 12.93%), 4th graders (n=14, 11.38%), 5th graders (n=6, 9.09%). A Chi-square test was conducted to evaluate the distribution of GS5 across grades. There was a marginally significant difference in the distribution of children who commit S5 error across grades, $\chi^2(4)=8.84$, $p=0.06$). Adjusted residuals analysis suggested that 2nd graders were the same probable to commit S5 error than

1st graders, 1.97 times more than 3rd graders, 2.28 times more than 4th graders and 2.93 times more than 5th graders.

A Chi-square test of independence was performed to examine the distribution of GS5 between sexes in the Simple Subtraction block. There were no significant differences in the distribution of children who commit S5 error between sexes, $\chi^2(1)=0.05$, $p>.05$.

iii) Investigation of Split effect in the group that did not commit S5 (GNoS5)

As shown in Figure 3.2, GNoS5 presented an overall tendency of decrease in the frequency of errors, as the distance from the correct answer (or split) increases i. e. a Split effect. A repeated measure ANOVA of error frequency by category of split error from S1 to S4 was significant, $F(1.97, 704.64)=298.92$, $p<.001$, $\eta_p^2=.46$. Bonferroni pairwise comparisons indicate that S1 frequency was higher than S2, S3 and S4 frequencies. S2 frequency was higher than S3 and S4 frequency. However, there was no difference between S3 and S4 frequencies.

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

The Split effect was also observed in GS5, but with lower effect size $F(2.32, 138.91)=3.19$, $p<.05$, $\eta_p^2=.05$. Bonferroni pairwise comparisons indicate that S1 frequency was the highest one. However, there was no difference among S2, S3 and S4. In GS5, the most notable tendency is the S5 effect. A repeated measure ANOVA of error frequency by category comparing S5 frequency with its adjacent errors, S4 and S6, was significant $F(1.35, 80.95)=58.94$, $p<.001$, $\eta_p^2=0.50$. Bonferroni pairwise comparisons indicate that S5 frequency was higher than S4 and S6 errors frequency. However, there was no difference between S4 and S6 frequencies (see Figure 3.2).

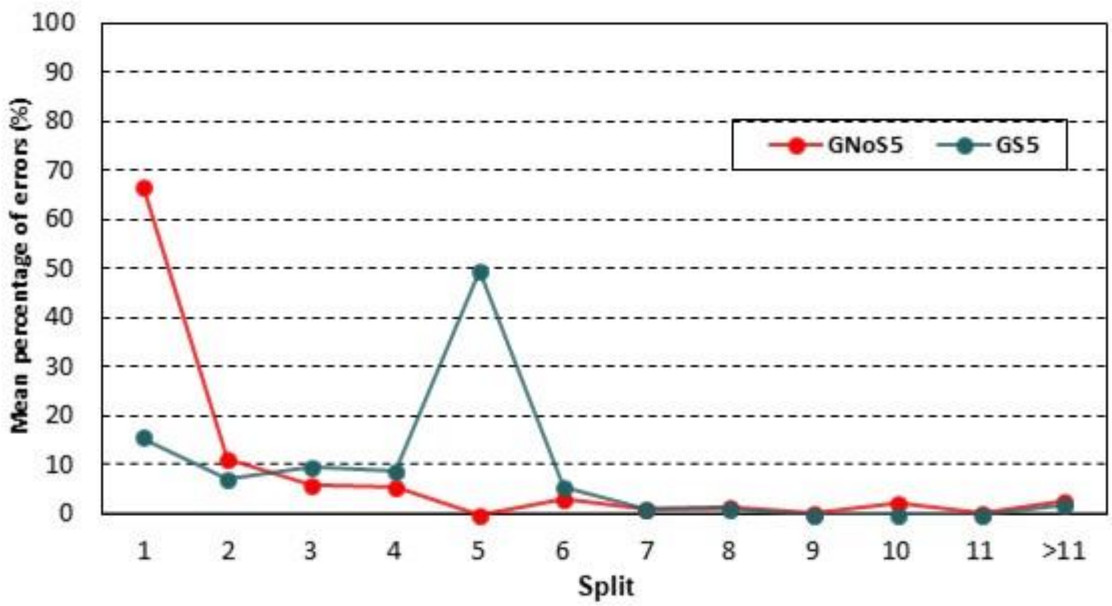

Figure 3.2. Mean percentage of Splits errors ranging from S1 to S>11 in the Simple Subtraction block for the GNoS5 and GS5.

v) Comparison of math achievement between GS5 and GNoS5.

In a subsample, data were available that allowed the comparison of math achievement between groups that committed and did not commit any S5 error in the Simple Subtraction block. A t-test was conducted to investigate if there is a difference in TDE subtest of arithmetic between GNoS5 (n=327; mean=-0.24; sd=0.91) and GS5 (n=52; mean=-0.37; sd=0.78). The result showed no significant difference between the groups in math achievement, t(377)=0.97, p>.05, Cohen's d=-0.15.

Complex Subtraction

i) Basic arithmetic performance

In the Complex Subtraction block, the percentage of children who did not respond to any item was 14.93% for 2nd graders and 10.62% for 3rd graders. This rate is low (<5%) for 4th and 5th graders. In 2nd grade and 5th grade, around 45% of items remained unresponded, while in 3rd grade and 4th grade this rate was higher, with 70.88% and 58.15% of unresponded items, respectively. In 2nd and 3rd grade around 20% of children responded to all items incorrectly, while in 4th to 5th grade this rate is lower than 10%.

Next, the percentage of incorrect answers based on responded items for the Complex Subtraction block was analyzed comparing grade levels and sex. One-way ANOVA resulted in significantly incorrect response rates across grades ($F(3, 872)=27.00$, $p<.001$, $\eta_p^2=.09$). Bonferroni corrected pairwise comparisons resulted in higher incorrect response rates of 2nd and 3rd graders in comparison with 4th and 5th graders. Also, 2nd graders had a higher percentage of incorrect answers in comparison with 3rd graders. There was no difference in the errors between 4th and 5th grades (see Figure 4.1). No sex differences were observed in the rate of incorrect response, $t(874)=0.03$, $p>.05$, Cohen's $d=0.00$.

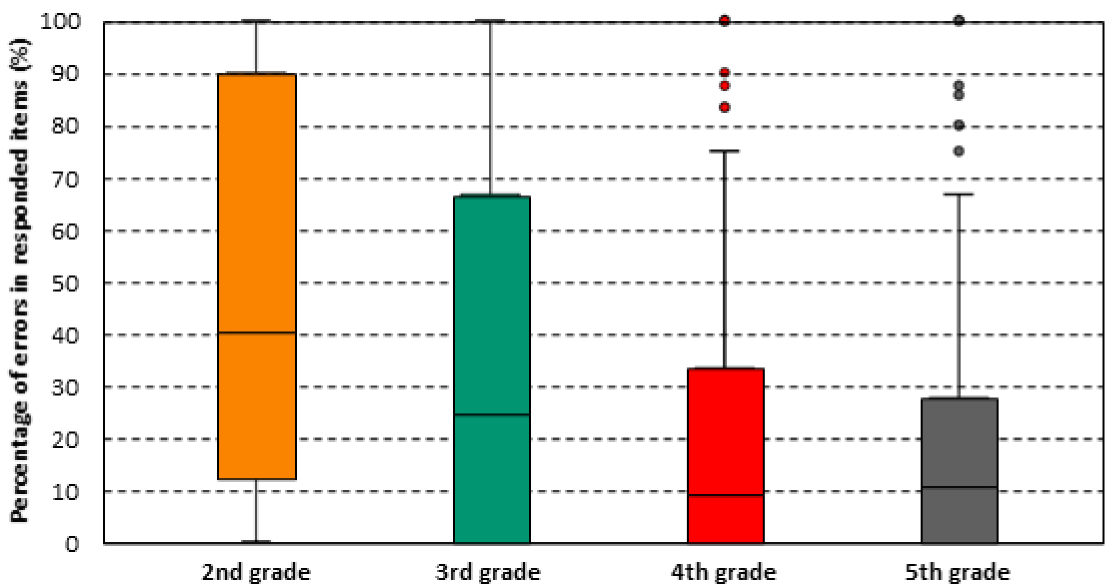

Figure 4.1. Distribution of incorrect responses in the Complex Subtraction block, stratified by grade.

Considering the accuracy on the responded items, 20.18% of 2nd graders responded to all items correctly, while in 3rd and 5th grades this rate was around 40% and 4th graders had the highest rate of 46.99%. These children's data will not be considered, resulting in a total of 527 children (60.16%) for further analyses.

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

From the 393 children selected for the Split Error Analysis in the Complex Subtraction, 16.13% (n=85) made at least one S5 error in the Complex Subtraction block. The percentages of children from GS5 decreased across grades: 2nd graders (n=24, 26.37%), 3rd graders (n=22, 13.66%), 4th graders (n=30, 16.22%), 5th graders (n=9, 10.00%). A Chi-square test was conducted to evaluate the distribution of GS5 across grades. There was a significant difference in the distribution of children who commit S5 error across grades, $\chi^2(3)=10.28$, $p<.05$. Adjusted residuals analysis indicated that 2nd were 2.26 times more probable to commit S5 errors than 3rd graders, 1.85 times more than 4th graders and 3.22 times more than 5th graders.

A Chi-square test of independence was performed to examine the distribution of GS5 between sexes in the Complex Subtraction block. There were no significant differences in the distribution of children who commit S5 error between sexes, $\chi^2(1)=1.82$, $p>.05$.

iii) Investigation of Split effect in the group that did not commit S5 error (GNoS5)

As shown in Figure 4.2, GNoS5 presented an overall tendency of decrease in the frequency of errors, as the distance from the correct answer (or split) increases i. E. a Split effect. A repeated measure ANOVA of error frequency by category of split error from S1 to S4 was significant, $F(2.01, 885.78)=179.95$, $p<.001$, $\eta_p^2=.29$. Bonferroni pairwise comparisons indicate that S1 frequency was higher than S2, S3 and S4 frequencies. S2 frequency was higher than S3 and S4 frequencies. However, there was no difference between S3 and S4 frequencies.

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

The Split effect was also observed in GS5, but with lower effect size $F(2.49, 209.16)=7.07$, $p<.001$, $\eta_p^2=.08$. Bonferroni pairwise comparisons indicate that S1 frequency was higher than S3 and S4 frequencies. However, there was no difference between S1 and S2 frequencies, and also there were no difference among S2, S3 and S4 frequencies. In GS5, the most notable tendency is the S5 effect. A repeated measure ANOVA of error frequency by category comparing S5 frequency with its adjacent errors, S4 and S6, was significant, $F(1.13, 94.96)=93.85$, $p<.001$, $\eta_p^2=0.53$. Bonferroni pairwise comparisons indicate that S5 frequency was higher than S4 and S6 frequency. However, there was no difference between S4 and S6 frequencies (see Figure 4.2).

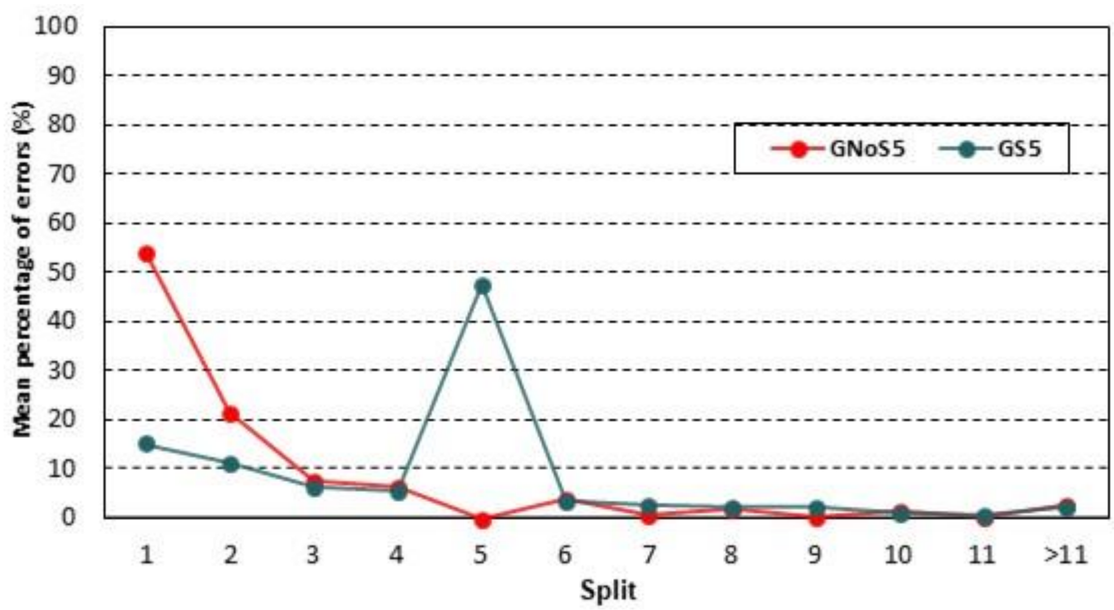

Figure 4.2. Mean percentage of Splits errors ranging from S1 to S>11 in the Complex Subtraction block for the GNoS5 and GS5.

v) Comparison of math achievement between GS5 and GNoS5

In a subsample, data were available that allowed the comparison of math achievement between groups that committed and did not commit any S5 error in the Complex Subtraction block. A t-test was conducted to investigate if there is a difference in TDE subtest of arithmetic between GNoS5 (n=441; mean=-0.06; sd=0.92) and GS5 (n=85; mean=-0.35; sd=0.96). The result showed that GS5 had a significant lower score in the TDE subtest of arithmetic in comparison with GNoS5, $t(524)=2.66$, $p<.01$, Cohen's $d=0.31$.

Simple Multiplication

i) Basic arithmetic performance

In the Simple Multiplication block, the percentage of children who did not respond to any item was 23.63% for 3rd graders. This rate is low (<5%) for 4th and 5th graders. In 3rd grade 69.84% of items remained unresponded, while in 4th and 5th grade this rate was lower with around 40% of unresponded items. In 3rd grade 21.97% of children responded to all items incorrectly, while in 4th to 5th grade this rate is lower than 5%.

Next, the percentage of incorrect answers based on responded items for the Simple Multiplication block was analyzed comparing grade levels and sex. One-way ANOVA resulted in significantly incorrect response rates across grades ($F(2, 724)=55.31$, $p<.001$, $\eta_p^2=.133$). Bonferroni corrected pairwise comparisons resulted in higher incorrect response rate of 3rd graders in comparison with 4th and 5th graders. There was no difference between 4th and 5th graders. The median error percentage of 4th and 5th grade is zero (see Figure 5.1). No sex differences were observed in the rate of incorrect response, $t(725)=-0.65$, $p>.05$, Cohen's $d=-0.05$.

Considering the accuracy on the responded items, 41.70% of 3rd graders responded to all items correctly, while in 4th and 5th grades this rate was higher, with around 60% of children. These children's data will not be considered in further analyses. Data from children who answered all 8 items not involving operand 5 correctly (N=91) will be also excluded, resulting in a total of 248 children (34.11%) for further analyses.

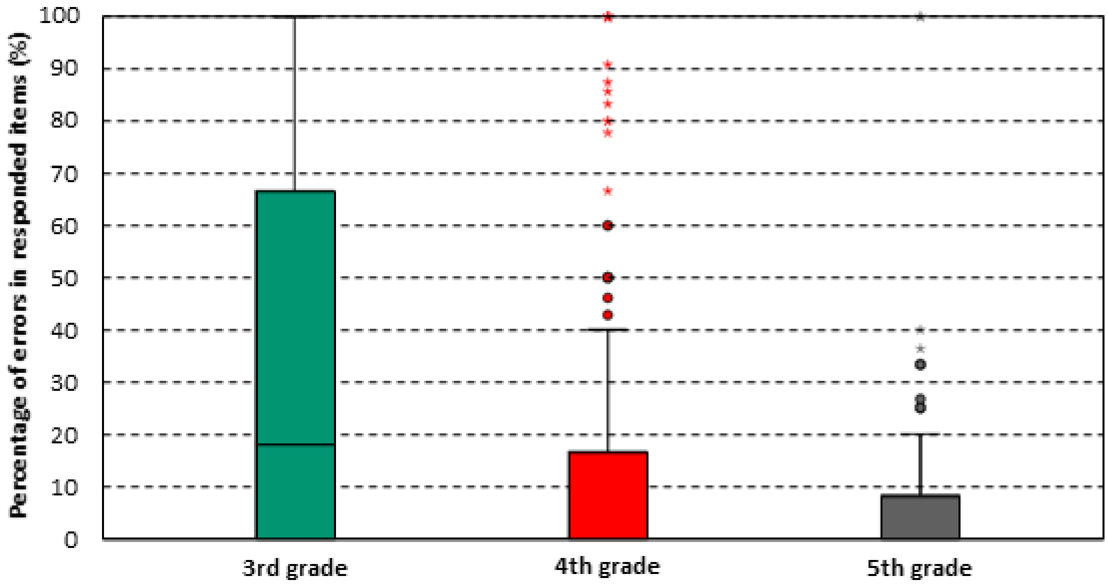

Figure 5.1. Distribution of incorrect responses in the Simple Multiplication block, stratified by grade

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

From the 248 children selected for the Split Error Analysis in the Simple Multiplication block, 10.89% (n=27) made at least one S5 error. The percentages of children from GS5 decreased across grades: 3rd graders (n=15, 14.71%), 4th graders (n=9, 9.00%), 5th graders (n=3, 6.52%). A Chi-square test was conducted to evaluate the distribution of GS5 across grades. There was no significant differences in the distribution of children who commit S5 error across grades, $\chi^2(2)=2.80$, $p>.05$.

A Chi-square test of independence was performed to examine the distribution of GS5 between sexes in the Simple Multiplication block. There were no significant differences in the distribution of children who commit S5 error between sexes, $\chi^2(1)=1.00$, $p>.05$.

iii) Investigation of Split effect in the group that did not commit S5 error (GNoS5)

As shown in Figure 5.2, GNoS5 presented no tendency of Split effect in small splits errors (S1 to S4). A repeated measure ANOVA of error frequency by category of split error from S1 to S4 was not significant, $F(3, 660)=0.44$, $p>.05$, $\eta_p^2=.00$. However, unlike the other operations, the frequencies of other splits, like S6, S9 and S>11 were raised.

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

In GS5, the Split effect was also not observed, $F(3,78)=0.34$, $p>.05$, $\eta_p^2=.01$. In GS5, the most notable tendency is the S5 effect. A repeated measure ANOVA of error frequency by category comparing S5 frequency with its adjacent errors, S4 and S6, was significant, $F(1.21, 32.52)=$, $p<.001$, $\eta_p^2=0.64$. Bonferroni pairwise comparisons indicate that S5 frequency was higher than S4 and S6 frequency. However, there was no difference between S4 and S6 frequencies (see Figure 5.2).

v) ***Comparison of math achievement between GS5 and GNoS5.***

In a subsample, data were available that allowed the comparison of math achievement between groups that committed and did not commit any S5 error in the Simple Multiplication block. A t-test was conducted to investigate if there is a difference in TDE subtest of arithmetic between GNoS5 (n=220; mean=-0.29; sd=0.88) and GS5 (n=27; mean=-0.73; sd=0.81). The result showed that GS5 had a significant lower score in math achievement in comparison with GNoS5, $t(245)=2.45$, $p<.05$, Cohen's $d=0.52$.

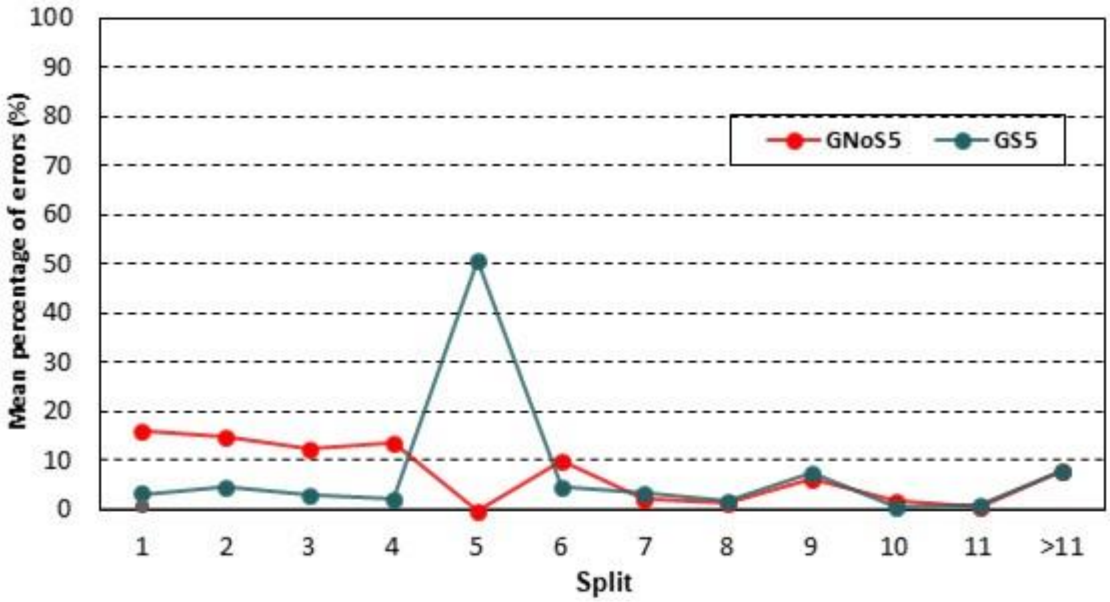

Figure 5.2. Mean percentage of Splits errors ranging from S1 to S>11 in the Simple Multiplication block for the GNoS5 and GS5.

Complex Multiplication

i) Basic arithmetic performance

In the Complex Multiplication block, the percentage of children who did not respond to any item was 45.55% for 3rd graders. This rate is low (<10%) for 4th and 5th graders. In 3rd grade 75.86% of items remained unresponded, while in 4th grade this rate was 67.14% and in 5th grade this rate was lower with 52.97% of unresponded items, respectively. In 3rd grade 54.72% of children responded to all items incorrectly, while in 4th to 5th grade this rate is lower than 20%.

Next, the percentage of incorrect answers based on responded items for the Simple Multiplication block was analyzed comparing grade levels and sex. The percentage of errors on the Complex Multiplication block was significantly different over grades ($F(2, 633)=34.47$, $p<.001$, $\eta_p^2=.098$). Bonferroni corrected pairwise comparisons demonstrated that 3rd graders had a higher percentage of errors in comparison with 4th and 5th graders. There was no difference between

4th and 5th graders (see Figure 6.1). No sex differences were observed in the rate of incorrect response, $t(634)=-0.58$, $p>.05$, Cohen's $d=-0.05$.

Considering the accuracy on the responded items, 25.16% of 3rd graders responded to all items correctly, while in 4th and 5th grades this rate was higher, with respectively 39.63% and 33.56% of children who answered 100% of items accurately. These children's data will not be considered, resulting in a total of 220 children (34.59%) for further analyses.

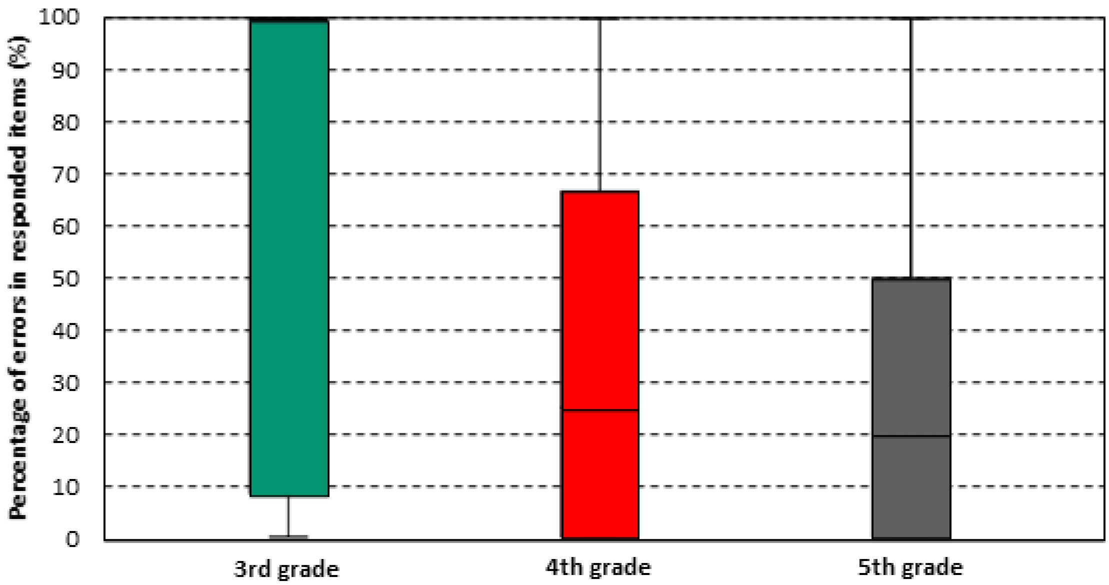

Figure 6.1. Distribution of incorrect responses in the Complex Multiplication block, stratified by grade.

ii) Frequency of children who committed at least one S5 error comparing across grades and sexes

From the 416 children selected for the Split Error Analysis in the Complex Multiplication block, 5.53% (n=23) made at least one S5 error. The percentages of children from GS5 per grade: 3rd graders (n=4, 3.36%), 4th graders (n=13, 6.57%), 5th graders (n=6, 6.06%). A Chi-square test was conducted to evaluate the distribution of GS5 across grades. There was no significant differences in the distribution of children who commit S5 error across grades, $\chi^2(2)=1.53$, $p>.05$.

A Chi-square test of independence was performed to examine the distribution of GS5 between sexes in the Complex Addition block. There were no significant differences in the distribution of children who commit S5 error between sexes, $\chi^2(1)=1.49$, $p>.05$.

iii) Investigation of Split effect in the group that did not commit S5 error (GNoS5)

As shown in Figure 6.2, GNoS5 presented a slight tendency of Split effect in small splits errors (S1 to S4). A repeated measure ANOVA of error frequency by category of split error from S1 to S4 was significant, $F(2.71, 1064.05)=8.89$, $p<.001$, $\eta_p^2=.02$. Bonferroni pairwise comparisons indicate that S1 frequency was higher than S3 and S4 frequencies, but there were no differences between the frequency of S1 compared to S2 and among frequencies of S2, S3 and S4. However, unlike the other operations, the frequencies of other splits, like S6, S8 and mainly S>11 were raised.

iv) Investigation of Split effect and S5 effect in the group that committed at least one S5 error (GS5)

In GS5, the Split effect was not observed, $F(2.18, 47.92)=0.90$, $p>.05$, $\eta_p^2=.04$. However, in GS5, other two tendencies were observed. Similar to GNoS5 the frequency of S>11 was raised and also the S5 effect is notable. A repeated measure ANOVA of error frequency by category comparing S5 frequency with its adjacent errors, S4 and S6, was significant, $F(1.14, 25.13)=25.61$, $p<.001$, $\eta_p^2=0.54$. Bonferroni pairwise comparisons indicate that S5 frequency was higher than S4 and S6 frequency. However, there was no difference between S4 and S6 frequencies (see Figure 6.2).

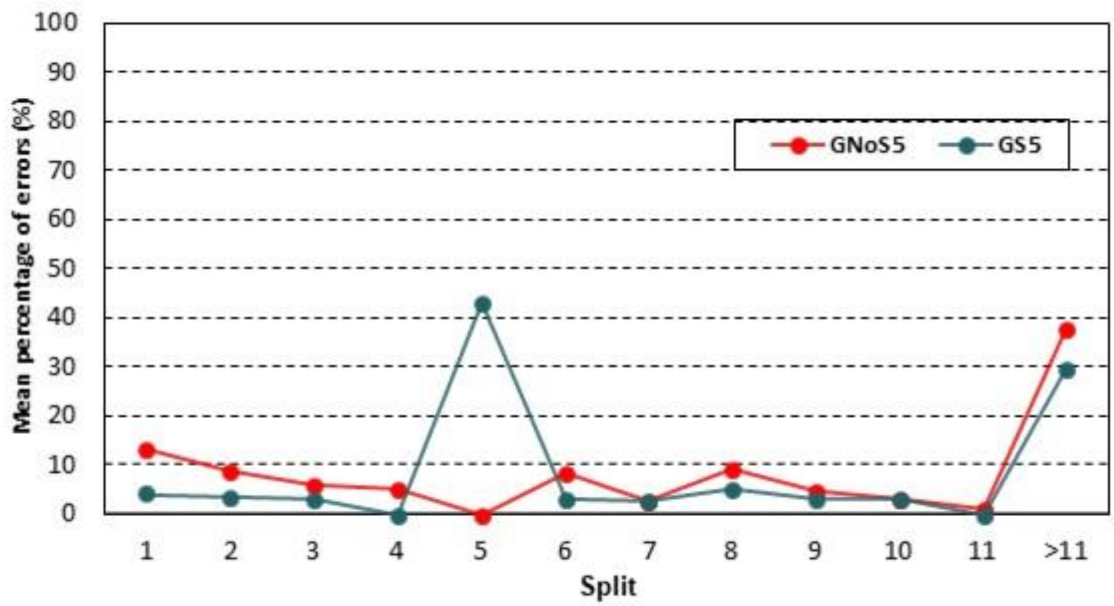

Figure 6.2. Mean percentage of Splits errors ranging from S1 to S>11 in the Complex Subtraction block for the GNoS5 and GS5.

v) Comparison of math achievement between GS5 and GNoS5.

In a subsample, data were available that allowed the comparison of math achievement between groups that committed and did not commit any S5 error in the Complex Multiplication block. A t-test was conducted to investigate if there is a difference in TDE subtest of arithmetic between GNoS5 (n=393; mean=-0.19; sd=0.91) and GS5 (n=23; mean=-0.41; sd=0.95). The result showed no significant difference between the groups in math achievement, t(414)=1.16, p>.05, Cohen’s d=-0.24.

Summary of results:

Table 1. Summary of Split analyses results (sections ii to v).

Section/ BAOT block	ii) Frequency of children who committed S5 errors across grades	ii) Sex differences in the frequency of who committed S5 errors	iii) Split effect in GNoS5	iv) Split effect in GS5	iv) S5 effect in GS5	v) Difference in math achievement between GNoS5 and GS5
Simple Addition	**1st grade>other grades**	$p>.05$	**$p<.001$** **$\eta p^2 = .42$**	**$p<.05$** **$\eta p^2=.06$**	**$p<.001$** **$\eta p^2=.52$**	$p>.05$
Complex Addition	**2nd grade>other grades** **5th grade<other grades**	$p>.05$	**$p<.001$** **$\eta p^2=.48$**	**$p<.001$** **$\eta p^2 =.12$**	**$p<.001$** **$\eta p^2 =.63$**	**$p<.01$** **Cohen's d=.41**
Simple Subtraction	**2nd grade>other grades except from 1st grade**	$p>.05$	**$p<.001$** **$\eta p^2=.46$**	**$p<.05$** **$\eta p^2=.05$**	**$p<.001$** **$\eta p^2=.50$**	$p>.05$
Complex Subtraction	**2nd grade>other grades**	$p>.05$	**$p<.001$** **$\eta p^2=.29$**	**$p<.001$** **$\eta p^2 =.08$**	**$p>.001$** **$\eta p^2=.53$**	**$p<.01$** **Cohen's d=.31**
Simple Multiplication	$p>.05$	$p>.05$	$p>.05$	$p>.05$	**$p<.001$** **$\eta p^2=.64$**	**$p<.01$** **Cohen's d=.52**
Complex Multiplication	$p>.05$	$p>.05$	**$p<.001$** **$\eta p^2=.02$**	$p>.05$	**$p<.001$** **$\eta p^2=.54$**	$p>.05$

Significant results are in bold letter

Discussion

Following Domahs and coworkers' (2008) lead, the present master thesis investigated the possible role of S5 effect in learning addition, subtraction, and multiplication by 1st to 5th graders. Results supported the hypothesis that younger children commit S5 errors more frequently. No grade difference in the frequency of children who commit S5 errors were observed from multiplication blocks. S5 effect was observed for all operations both at the relatively more simple and more complex levels. No sex differences were observed in the frequency of children who committed S5 errors. A Split effect was also observed for all operations, except simple multiplication, being of higher effect size for children who did not commit S5 errors. Math achievement was lower for children who committed S5 in complex but not in simple addition and subtraction. Otherwise, math achievement was lower for children who committed S5 in simple but not in complex multiplication. The Discussion section will be subdivided into the following topics: i) Basic arithmetic performance; ii) Frequency of children who committed S5 errors across grades; iii) Sex differences in the frequency of children who committed S5 errors; iv) Split effect in basic arithmetic; v) Split-five effect in basic arithmetic; vi) Association between Split-five errors and math achievement; vii) Weaknesses and Strengths of the study.

i) Basic arithmetic performance

A partial analysis of the data on performance in basic arithmetic operations was published by Gomides et al. (2021), showing that despite general improvement of arithmetic abilities from 3th to 5th grade, a considerable percentage of Brazilian children struggle in mastering addition, subtraction and multiplication. In the present study, we add information about Brazilian children's performance in 1st and 2nd grades, showing that most part of children start learning how to solve addition problems in 1st grade and subtraction problems in 2nd grade. Most part 1st graders did not know how to perform simple subtraction problems. Among children who know how to solve simple addition and subtraction, 1st graders presented a considerable rate of incorrect answers and unresponded items even with a longer limit time on the task. The same pattern of performance was observed for 2nd graders in complex addition and for both subtraction blocks. These findings are in line with Brazilian curricular guidelines (MEC, 2018) of gradually teaching the basic operations, starting with addition in 1st grade and subtraction in 2nd grade.

Moreover, in addition to Gomides and colleagues' (2021) study, we demonstrated that the difficulty of children from 3rd to 5th grade in mastering addition and subtraction is more related to fluency than to accuracy. In fact, children do not achieve 100% of accuracy on the tasks mostly because they leave items unresponded and not because they are committing too many errors. In both addition blocks and in simple subtraction, more than 50% of children answered all responded items correctly, while in complex subtraction blocks more than 40% of children answered all items correctly.

Results of the multiplication blocks are in line with Gomides and coworkers (2021) study. Although multiplication starts to be taught in 3rd grade, children start mastering the resolution of multiplication problems just in 4th grade. A possible explanation for the difficulty of Brazilian children in performing the basic operations with fluency is that Brazilian curricular guidelines emphasize the learning of different procedural strategies for solving the arithmetic problems, without actively promoting the automatization of arithmetic facts (MEC, 2018).

Results on simple arithmetic performance are in line with Domahs and colleagues' (2008) findings, who demonstrated a decrease in accuracy in the simple addition block when children are learning multiplication at the end of 2nd grade. However, in our sample, this effect was found in 4th grade, in which children presented the highest percentage of incorrect answers and the highest rate of children who answered all items incorrectly in the Simple Addition block.

ii) Frequency of children who committed S5 errors across grades

To the best of our knowledge, Domahs and colleagues (2008) is the only study investigating S5 effect in children. Nuovo (2018) observed the same effect in a robot simulation when a finger counting system was introduced. Present results are in line with those of Domahs and colleagues' study. The number of children who committed S5 errors decreases for addition and subtraction across grades, being more frequent in young children. This result is in consonance with an increase in accuracy across grades, demonstrating that children are progressively more able to implement effective strategies for solving the problems. Specifically, in the Simple Subtraction block, 1st and 2nd graders had the same chance of committing S5 errors. A plausible explanation is that children start learning subtraction in 2nd grade, so the small part of 1st graders (less than 25%) who were able to solve this block were in a similar stage of development as 2nd graders.

An inconsistent profile of S5 effect in simple addition problems was observed by Domahs and colleagues (2008) in the 2nd grade. In the present study, we observed an analogous increased percentage of 4th graders committing S5 errors. In Brazil, there seems to be a delay in this process, probably related to a general pattern of slower increase in arithmetical performance.

Domahs et al. (2008) assume that the increase in the number of children who committed S5 errors in the 2nd grade could be related to interference from learning multiplication and the transition from using procedural strategies to retrieval strategies. The items used by Domahs et al. (2008) in simple addition were purposely those for which the children had already learned the multiplication result. This may have induced the children to confuse the operation of addition with multiplication. An alternative explanation is that older children commit S5 errors in simple addition, only because they are responding to the items using the multiplication reasoning that they are learning.

In the present study, the frequency of children committing S5 errors across grades was also investigated for multiplication, which was not analyzed by Domahs and colleagues (2008). A nonsignificant tendency of decrease across grades was observed for the frequency of children who committed S5 errors for simple but not for complex multiplication. This could be related to the fact that multiplication operations are rather difficult for the children in the sample regardless of their grade and they often rely on ineffectiveness strategies for solving the problems, such as procedural strategies. This is corroborated by the higher rate of unresponded items in multiplication problems because procedural strategies demand more time to be implemented than long-term retrieval strategies (Geary et al., 2012).

iii) Sex differences in the frequency of children who committed S5 errors

Apart from Domahs and collegues (2008) results, which observed that girls committed significantly more S5 errors than boys, the present study did not find any sex difference in both basic arithmetical performance and distribution of children who committed S5 errors. Domahs and collegues did not explore which reason would explain the sex difference for S5 errors, and this result could be associated with several factors, such as math anxiety that is more prevalent in girls (Dowker et al., 2016; Orbach et al., 2019). In general, sex differences are not consistently found in math abilities (Hyde, 2016), and it can explain the present results. Due to the inconsistency in the results regarding sex differences in S5 errors, further studies should continue the investigation for clarifying its association.

iv) Split effect in simple arithmetic

The Split effect corresponds to a decrease in the frequency of errors, as the distance from the correct response (split magnitude) increases and are mostly observed in verification tasks by a decrease of response time, the greater the distance between the presented incorrect and the correct results (Ashcraft & Battaglia, 1978). Domahs and colleagues (2008) observed the Split effect in a production task. However, they did not analyze its significance. In the present study, the Split effect was also observed. In children who did not commit S5 errors, a significant Split effect in simple arithmetic was observed for all BAOT blocks, except for simple multiplication. In children who committed S5 errors, the Split effect had a lower effect size (eta^2 approximately equal 0.1) than in children that did not commit S5 errors (eta^2 approximately equal 0.4) and it was observed only for addition and subtraction. These results are in line with Domahs et al. results showing a small effect size of Split effect for children who committed S5 errors, and a general increase of S1 and S2 errors proportions followed by a decrease in S5 errors.

One possible interpretation is that the Split effect is associated with the approximate nature of numerical magnitude representations in the mental number line (Ashcraft & Fierman, 1982). As children do not master efficient retrieval strategies, and still rely on procedural strategies, a broader field of numerical magnitudes is activated, resulting in higher frequency of small split magnitude errors.

The Split effect was not observed for the Simple Multiplication block. Although the Split effect was significant for the Complex Multiplication block, the effect size is low ($eta^2=0.02$). These results could indicate the variability of strategies used by the children. Some children may be using retrieval strategies while others are using procedural strategies.

A marked rate of S>11 errors was observed in both the Simple (GNoS5=7.97%; GS5=7.95%) and Complex (GNoS5=37.70%; GS5=29.35%) Multiplication blocks. Presence of this error indicates the use of the retrieval strategy, as using this strategy, children may retrieve from long-term memory the fact of another multiplication problem, i.e. operand and table errors (Butterworth et al., 2003).

v) Split-five effect in basic arithmetic

Domahs and colleagues (2008) observed that the S5 effect occurred only for more complex addition and subtraction operations, except for the occurrence of S5 effect in simple addition by 2nd graders, when children are learning multiplication. In the present study, the S5 effect was observed for both levels of complexity (simple and complex), in the three types of operations (addition, subtraction, and multiplication).

In complex addition and subtraction blocks, the results are in consonance with Domahs and coworkers' (2008) findings. The S5 effect in complex addition and subtraction can be explained by the use of procedural strategies wherein children struggle keeping track of how many hands they reused, regardless of if the procedure being performed externally on fingers (FCS) or mentally based on FBRs.

However, the S5 effect found for simple addition and subtraction problems were not expected. In the simple addition and subtraction problems, the S5 effect cannot be well explained by the explicit use of FCS, as the magnitude of the operands and results are equal or less than ten, so they could be manipulated with one-to-one correspondence by fingers. A possible explanation for these S5 errors in simple problems is the implementation of mental strategies using finger-pattern, in which children lose track of the quantities in the mental procedure.

Another unexpected S5 effect was observed for simple and complex multiplication problems. Despite the frequency of children committing S5 errors being lower in multiplication in comparison with addition and subtraction, this type of error was not expected at all for multiplication, as this operation is known to be resolved by fact retrieval strategies (Dehaene, 1992). S5 errors in multiplication could then be related to the fact that some children in the present study are still struggling to learn multiplication. Moreover, as already mentioned, Brazilian curricular guidelines do not encourage the consistent use of fact retrieval strategies (MEC, 2018).

In contrast with Domahs and colleagues (2008) findings, which indicates an increase of S10 errors over time in complex operations, the present study failed to find this effect in all operations and complexity levels tested. In the Domahs and colleagues’ study, the increase of S10 errors was paralleled by a decrease of S5 errors, indicating a shift to procedures embedded into a pure base-10 representation. A possible explanation could be that Brazilian children are still not implementing base-10 strategies. Clearing this question would require investigating conceptual knowledge of the base-10 system.

vi) Association between Split-five errors and math achievement

The current study is the first to investigate the association between S5 errors and math achievement. The t-tests comparing children who committed and those who did not commit S5 errors revealed a significant lower math achievement of children who committed S5 errors in the Complex Addition, Complex Subtraction, and Simple Multiplication blocks. However, the same effect was not observed for the Simple Addition, Simple Subtraction, and Complex Multiplication blocks.

In Figure 7, a summary of the results regarding the association of S5 errors and math achievement is presented with possible interpretations about the meaning of S5 in each operation and level of complexity. Addition and subtraction were considered together, as the same pattern of association were observed. Inferences about the presumed strategy used by children who commit S5 errors were made considering the type of operation (addition/subtraction or multiplication), level of complexity (simple or complex), the evidence of presence or absence of difference in math achievement between children who committed and did not committed S5 errors. For multiplication, follow up analysis were performed to check the hypothesized rationale.

On the one hand, simple addition and subtraction problems have results and operands equal or less than ten, so it can be manipulated externally by the fingers with one-to-one correspondence. Children who commit S5 errors in the simple addition and subtraction problems are likely to be the ones who are progressing in strategies, trying to make use of mental calculation with manipulation of FBRs. Progressing in calculation strategies is not an indicative of learning difficulties in mathematics. In fact, it is the opposite, children who have difficulty learning math are more likely not to progress in strategies, depending on the external use of the fingers (Geary, 2004; Butterworth, 2019).

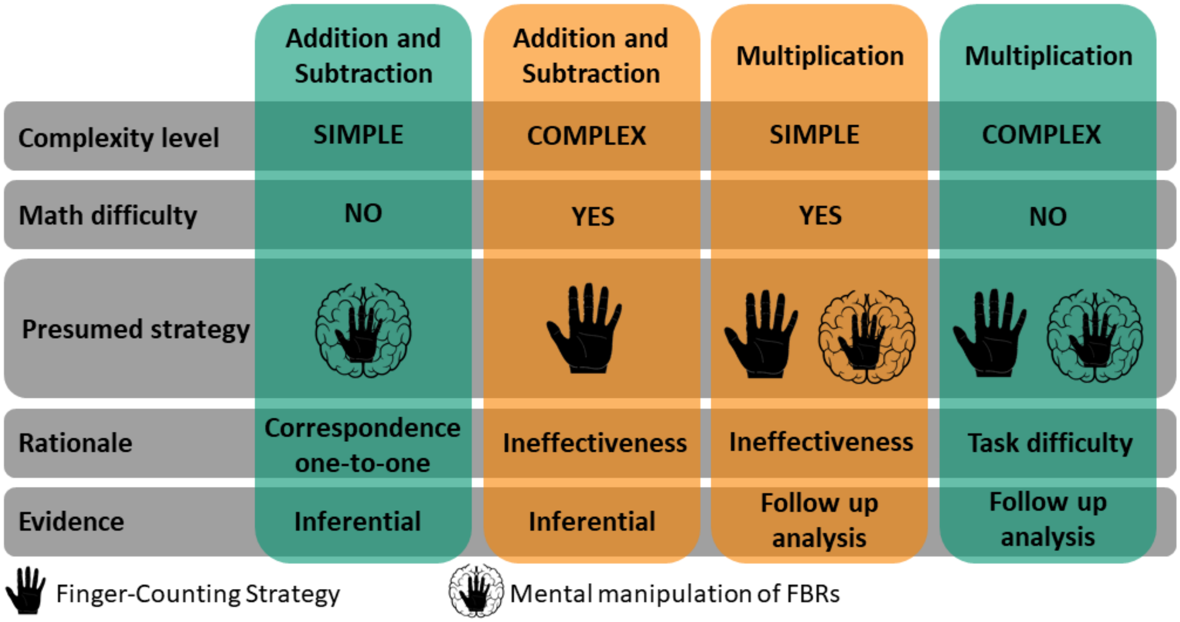

Figure 7. Association between S5 effect and math achievement: summary of results and possible interpretations.

On the other hand, complex addition and subtraction problems have results that cannot be manipulated externally by fingers using one-to-one correspondence. Children with difficulties learning math may persist in using external FCS even for complex addition and subtraction calculations, wherein this strategy is ineffective, and are more likely to commit S5 errors.

In multiplication blocks, few children committed S5 errors in simple problems. In this case, children are likely to rely on FCS, which is ineffective in solving multiplication and can be an indicator of difficulties in learning math. A follow-up analysis demonstrated that the mean rate of unresponded items in GS5 was 51.85% (sd=26.17) and the rate of incorrect answers in the responded items was 80.57% (sd=29.46) demonstrating that most items remained unresponded and incorrect answers were given for the greater part of the responded items.

Complex multiplication problems are difficult for most children, and those with difficulties learning math are likely not even to be able to solve these calculations yet. A follow-up analysis selecting children with the performance below percentile 25th in TDE arithmetic, indicated that

24.10% of these children did not solve any item. From those children who solved at least one item, the mean rate of unresponded items was 70.86% (sd=21.97).

vii) Weaknesses and Strengths of the study

Two main weaknesses should be recognized. As the strategies used by children were not recorded, interpretations remain largely speculative. As the BAOT task used for evaluating the basic arithmetic operations is a fluency task, with time limit, children that are not proficient and were using more procedural strategies (the ones we were most interested in), left some items unresponded. Consequently, there was a gap, in which it was not possible to assess whether the children would make or not S5 errors in these unresponded items. For example, the two last items of the Simple Addition block were those in which the correct answer for multiplication deviates exactly by five from the correct response of addition, and some children left these items unresponded.

The main strengths of the present study are the relatively large sample and range of investigated grades. The S5 effect for multiplication and the Split effect were also investigated. In addition, an association between the S5 effect and math achievement was investigated.

Conclusion

The results of the present study corroborate the existence of the S5 effect, as a signature of the FBRs influence on the basic arithmetic performance. In addition to the previous findings of Domahs et al. (2008) study, this master thesis expanded the evidence of S5 effect demonstrating its occurrence for the multiplication problems. Also, results are in line with Domahs and colleagues' study, in the sense that we reported the developmental influence in S5 errors, as it occurs mostly in younger children. However, it is important to emphasize that the developmental influence in S5 errors may be mediated by the use of strategies relying on FBRs, considering that these strategies are related to S5 errors, and these strategies are more frequently used by young children.

Furthermore, the first evidence of an association between S5 error and math achievement were presented, indicating that S5 errors may be suggestive of the persistent use of strategies relying on FBRs for complex problems and multiplication, even when this strategy is ineffective for the problem's resolution. However, the interpretations about which type of strategy were used when children commit S5 errors is merely speculative, since we did not measure the strategies used by children during BAOT's execution. Further studies should include a direct measurement of strategies, to deeply investigate in which degree FBRs influences calculation leading children to commit S5 errors. Previous findings indicate that S5 error may occur with the implementation of different types of strategies that rely in FBRs.

Further studies should also investigate the cognitive mechanisms related to the S5 error. Considering the present results, S5 error may be associated with an impairment in working memory or attention, occurring as a specific children's difficulty in keeping track of chunks of five, that is how many hands they reused during the resolution of the arithmetical problem.

Finally, the present study has also educational implications. Qualitative analyses of the errors can provide additional information about children's basic arithmetic performance. Particularly the occurrence of S5 errors can be an indicative of persistent and ineffectiveness use of strategies relying on FBRs, and as children with math learning difficulty start using their fingers later and persistently rely on this strategy (Butterworth, 2019), it is important to be aware of type of error. The early identification of children with mathematical learning difficulties are an important investment, preventing children of having increasingly learning gaps (Balt et al., 2019). Additionally, response to intervention's programs can benefit these children, as it follows the individual progress on math learning trajectories, assessing the preexisting numerical knowledge for providing tailored interventions that help them progress on their learning, for example teaching how to progress from finger counting-based strategies to more mature strategies (Balt et al., 2019; Freitas et al., in press).

References

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. Cognitive development, 14(1), 37-56. https://doi.org/10.1016/S0885-2014(99)80017-3

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4*(5), 527–538. https://doi.org/10.1037/0278-7393.4.5.527

Ashcraft, M. H., & Fierman, B. A. (1982). Mental addition in third, fourth, and sixth graders. Journal of Experimental Child Psychology, 33(2), 216–234. https://doi.org/10.1016/0022-0965(82)90017-0

Balt, M., Ehlert, A., & Fritz, A. (2019). Assessment in inclusive mathematics education. Approaches to designing progress assessments for numeracy learning. In D. Kollosche, R. M. J. de Souza, M. Knigge, M. G. Penteado, & O. Skovsmose (Eds.), Inclusive mathematics education (pp. 197–216). Springer. https://doi.org/10.1007/978-3-030-11518-0_14

Balt, M., Fritz, A., & Ehlert, A. (2020). Insights into First Grade Students' Development of Conceptual Numerical Understanding as Drawn From Progression-Based Assessments. Frontiers in Education. 5: 80. https://doi.org/10.3389/feduc.2020.00080

Beller, S., & Bender, A. (2011). Explicating numerical information: when and how fingers support (or hinder) number comprehension and handling. Frontiers in psychology, 2.,214. https://doi.org/10.3389/fpsyg.2011.00214

Butterworth, B. (1999). *What counts: How every brain is hardwired for math*. The Free Press.

Butterworth, B. (2019). Dyscalculia. From science to education. Routledge.

Butterworth, B., Marchesini, N., & Girelli, L. (2003). Multiplication facts: Passive storage or dynamic reorganization? In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetical concepts and skills*. (pp. 189-202). Erlbaum.

Costa, A. J., Silva, J. B., Chagas, P. P., Krinzinger, H., Lonneman, J., Willmes, K., Wood, G., & Haase, V. G. (2011). A handful of numbers: a role for offloading in arithmetics learning?. *Frontiers in psychology*, *2*, 368. https://doi.org/10.3389/fpsyg.2011.00368

Crollen, V., & Noël, M. P. (2015). The role of fingers in the development of counting and arithmetic skills. Acta psychologica, 156, 37-44. https://doi.org/10.1016/j.actpsy.2015.01.007

Dehaene, S. (2011). *The number sense. How the mind creates mathematics* (Revised and expanded edition). Oxford University Press.

Dehaene S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1-2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-n

Di Nuovo, A. (2018). Long-short term memory networks for modelling embodied mathematical cognition in robots. In *2018 International Joint Conference on Neural Networks (IJCNN)*, 1-7. IEEE. https://doi.org/10.1109/IJCNN.2018.8489140.

Domahs, F., Krinzinger, H., & Willmes, K. (2008). Mind the gap between both hands: evidence for internal finger-based number representations in children's mental calculation. Cortex, 44(4), 359-367. https://doi.org/10.1016/j.cortex.2007.08.001

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years?. Frontiers in psychology, 7, 508. https://doi.org/10.3389/fpsyg.2016.00508

Dwyer, C. A. (1975). Book Reviews: Maccoby, E. E., and Jacklin, C. N. The Psychology of Sex Differences Stanford, Calif.: Stanford University Press. *American Educational Research Journal*, *12*(4), 513–516. https://doi.org/10.3102/00028312012004513

Ferreira, F. O., Wood, G., Pinheiro-Chagas, P., Lonnemann, J., Krinzinger, H., Willmes, K., & Haase, V. G. (2012). Explaining school mathematics performance from symbolic and nonsymbolic magnitude processing: similarities and differences between typical and low-achieving children. Psychology & Neuroscience, 5(1). https://doi.org/10.3922/j.psns.2012.1.06

Freitas, F. R., Herzog, M., Haase, V. G., Fritz, A. (submetido). Compreensão conceitual do número no diagnóstico e intervenção para as dificuldades de aprendizagem na aritmética.

Fritz, A., Ehlert, A., & Balzer, L. (2013). Development of mathematical concepts as basis for an elaborated mathematical understanding. South African Journal of Childhood Education, 3(1), 38-67. https://doi.org/10.4102/sajce.v3i1.31

Fuson, K. C., & Kwon, Y. (1992). Learning addition and subtraction: Effects of number words and other cultural tools. Lawrence Erlbaum Associates, Inc.

Geary, D. C., & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. In J. I. D. Campbell (Ed.), Handbook of mathematical cognition (pp. 253-267). New York: Psychology Press.

Geary, D. C. (2004). Mathematics and learning disabilities. Journal of Learning Disabilities, 37, 4-15. https://doi.org/10.1177/00222194040370010201

Geary, D. C., Hoard, M. K., & Bailey, D. H. (2012). Fact retrieval deficits in low achieving children and children with mathematical learning disability. *Journal of learning disabilities*, *45*(4), 291–307. https://doi.org/10.1177/0022219410392046

Gomides, M. R. D. A., Starling-Alves, I., Paiva, G. M., Caldeira, L. D. S., Aichinger, A. L. P. N., Carvalho, M. R. S.,Bahnmueller, J., Moeller, K., Silva, J., & Haase, V. G. (2021). The quandary of diagnosing mathematical difficulties in a generally low performing population. *Dementia & Neuropsychologia*, *15*, 267-274. https://doi.org/10.1590/1980-57642021dn15-020015

Gomides, M. R. A., Lopes-Silva, J. B., Moura, R., de Salles, J. S. & Haase, V. G. (in press). Bateria de avaliação do processamento numérico e cálculo - PRONUMERO. Vetor.

Hyde, J. S. (2016). Sex and cognition: gender and cognitive functions. *Current opinion in neurobiology*, *38*, 53-56. https://doi.org/10.1016/j.conb.2016.02.007

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2008). Development of number combination skill in the early school years: when do fingers help? Developmental Science, 11(5), 662-668. https://doi.org/10.1111/j.1467-7687.2008.00715.x

Klein, E., Moeller, K., Willmes, K., Nuerk, H. C., & Domahs, F. (2011). The influence of implicit hand-based representations on mental arithmetic. Frontiers in psychology, 2, 197. https://doi.org/10.3389/fpsyg.2011.00197

Kuder, G.F., & Richardson, M.W. (1937). The theory of the estimation of test reliability. Psychometrika, 2, 151–160. https://doi.org/10.1007/BF02288391

Lafay, A., Thevenot, C., Castel, C., & Fayol, M. (2013). The role of fingers in number processing in young children. *Frontiers in psychology*, *4*, 488. https://doi.org/10.3389/fpsyg.2013.00488

Langhorst, P., Ehlert, A. & Fritz, A. (2012). Non-numerical and Numerical Understanding of the Part-Whole Concept of Children Aged 4 to 8 in Word Problems. Journal für Mathematik-Didaktik, 33(2), 233-262. https://doi.org/10.1007/s13138-012-0039-5

Lehtinen, E., Hannula-Sormunen, M., McMullen, J., & Gruber, H. (2017). Cultivating mathematical skills: From drill-and-practice to deliberate practice. *ZDM*, *49*(4). https://doi.org/10.1007/s11858-017-0856-6

Ministério da Educação. (2018). Base Nacional Comum Curricular. https://basenacionalcomum.mec.gov.br/

Moeller, K., Martignon, L., Wessolowski, S., Engel, J., & Nuerk, H. C. (2011). Effects of finger counting on numerical development–the opposing views of neurocognition and mathematics education. *Frontiers in psychology*, *2*, 328. https://doi.org/10.3389/fpsyg.2011.00328

Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Math= male, me= female, therefore math≠ me. *Journal of personality and social psychology*, *83*(1), 44. https://doi.org/10.1037/0022-3514.83.1.44

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... & Greenwald, A. G. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593-10597. https://doi.org/10.1073/pnas.0809921106

Orbach, L., Herzog, M., & Fritz, A. (2019). Relation of state-and trait-math anxiety to intelligence, math achievement and learning motivation. Journal of Numerical Cognition, 5(3), 371-399. https://doi.org/10.5964/jnc.v5i3.204

Park, J., & Brannon, E. M. (2013). Training the approximate number system improves math proficiency. *Psychological science*, *24*(10), 2013-2019.https://doi.org/10.1177/0956797613482944

Park, J., & Brannon, E. M. (2014). Improving arithmetic performance with number sense training: an investigation of underlying mechanism. *Cognition*, *133*(1), 188–200. https://doi.org/10.1016/j.cognition.2014.06.011

Parsons, S., & Bynner, J. (2005). Does numeracy matter more?. London: University of London, Institute of Education National Research and Development Centre for Adult Literacy and Numeracy. Accessed: http://discovery.ucl.ac.uk/1566245/.

Raven, J., Raven, J. C., Court, J. H., Paula, J. J., Alves, G. A. S., Malloy-Diniz, L. F. & Schlottfeldt C. G. M. F. (2018). CPM RAVEN - Matrizes Progressivas Coloridas de Raven: validação e normatização brasileira (1ª ed). São Paulo: Pearson.

Reeve, R., & Humberstone, J. (2011). Five-to 7-year-olds' finger gnosia and calculation abilities. Frontiers in psychology, 2, 359. https://doi.org/10.3389/fpsyg.2011.00359

Shusterman, A., Slusser, E., Halberda, J., & Odic, D. (2016). Acquisition of the Cardinal Principle Coincides with Improvement in Approximate Number System Acuity in Preschoolers. *PloS one*, *11*(4), e0153072. https://doi.org/10.1371/journal.pone.0153072

Siegler, R. S. (1999). Strategic development. Trends in Cognitive Science, 11(3), 430–435. https://doi.org/10.1016/S1364-6613(99)01372-8

Soylu, F., Lester Jr., F. K., & Newman, S. D. (2018). You Can Count on Your Fingers: The Role of Fingers in Early Mathematical Development. *Journal of Numerical Cognition*, *4*(1), 107-135. https://doi.org/10.5964/jnc.v4i1.85

Stein, L. M. (1994). TDE - Teste de desempenho escolar: Manual para aplicação e interpretação. Stein, L. M. (Eds). TDE - Teste de desempenho escolar: Manual para aplicação e interpretação. São Paulo: Casa do Psicólogo.

Szkudlarek, E., & Brannon, E. M. (2017). Does the approximate number system serve as a foundation for symbolic mathematics? *Language Learning and Development, 13*(2), 171–190. https://doi.org/10.1080/15475441.2016.1263573

Thompson, I. (1999). Mental calculation strategies for addition and subtraction. Part 1. *Mathematics in school*, *28*(5), 2-4. Accessed in: https://www.researchgate.net/publication/345870473_MENTAL_CALCULATION_STRATEGIES_FOR_ADDITION_AND_SUBTRACTION_WITH_INTEGERS

Wasner, M., Moeller, K., Fischer, M.H., & Nuerk, H. (2015). Related but not the same: Ordinality, cardinality and 1-to-1 correspondence in finger-based numerical representations. *Journal of Cognitive Psychology, 27*, 426 - 441. https://doi.org/10.1080/20445911.2014.964719

Supplementary material A – Descriptive analyses of participants' characteristics

Table 1. Descriptive analyses of participants' characteristics who responded at least one item in the specific BAOT block.

	N	Mean (sd) age	% female	% BH	% public school
Simple Addition					
General sample	1093	8.28 (1.49)	53.52	88.20	87.75
1st graders	178	5.90 (0.39)	47.19	100	43.82
2nd graders	128	7.00 (0.51)	53.91	100	100
3rd graders	277	8.10 (0.57)	50.18	78.70	98.92
4th graders	356	9.22 (0.49)	57.02	80.34	97.19
5th graders	154	10.25 (0.59)	58.44	100	93.51
Complex Addition					
General sample	911	8.74 (1.15)	54.77	85.95	97.48
2nd graders	127	7.00 (0.50)	53.54	100	100
3rd graders	274	8.08 (0.57)	50.36	78.83	98.91
4th graders	356	9.22 (0.49)	57.02	80.34	97.19
5th graders	154	10.25 (0.59)	58.44	100	93.51
Simple Subtraction					
General sample	958	8.60 (1.28)	55.01	86.53	96.14
1st graders	52	5.92 (0.33)	55.77	100	73.08
2nd graders	121	7.01 (0.51)	55.37	100	100
3rd graders	276	8.09 (0.57)	50.00	78.62	98.91
4th graders	355	9.22 (0.49)	57.18	80.28	97.18
5th graders	154	10.25 (0.59)	58.44	100	93.51
Complex Subtraction					
General sample	876	8.78 (1.14)	55.02	97.49	86.96
2nd graders	114	7.00 (0.50)	53.51	100	100
3rd graders	261	8.09 (0.56)	50.19	78.93	98.85
4th graders	349	9.23 (0.50)	57.59	80.52	97.13
5th graders	152	10.26 (0.58)	58.55	100	94.08
Simple Multiplication					
General sample	727	9.10 (0.93)	55.43	85.01	97.39
3rd graders	233	8.12 (0.56)	50.22	81.61	100
4th graders	351	9.22 (0.48)	57.26	80.63	97.15
5th graders	153	10.26 (0.58)	58.82	100	94.12
Complex Multiplication					
General sample	636	9.19 (0.90)	57.39	87.42	97.01
3rd graders	159	8.14 (0.55)	54.09	86.79	100
4th graders	328	9.22 (0.48)	58.54	82.01	96.95
5th graders	149	10.25 (0.58)	58.39	100	93.96

Table 2. Descriptive analyses of participants' characteristics for Split Error Analyses (Results sections ii to v).

	N	Mean (sd) age	% female	% BH	% public school
Simple Addition					
General sample	405	7.53 (1.55)	52.84	91.60	80.00
1st graders	142	5.92 (0.38)	47.89	100	45.07
2nd graders	61	6.90 (0.47)	52.46	100	100
3rd graders	82	7.98 (0.57)	46.34	75.61	97.56
4th graders	86	9.16 (0.48)	60.47	83.72	100
5th graders	34	10.15 (0.70)	70.59	100	97.06
Complex Addition					
General sample	393	8.54 (1.25)	52.42	89.57	98.73
2nd graders	91	6.98 (0.49)	56.04	100	100
3rd graders	107	8.05 (0.59)	43.93	72.90	99.07
4th graders	131	9.15 (0.44)	52.67	90.84	99.24
5th graders	64	10.34 (0.60)	60.94	100	95.31
Simple Subtraction					
General sample	420	8.34 (1.42)	53.81	89.29	95.48
1st graders	40	5.90 (0.38)	55.00	100	70.00
2nd graders	75	6.99 (0.51)	50.67	100	100
3rd graders	116	8.07 (0.62)	47.41	76.72	99.14
4th graders	123	9.16 (0.49)	56.10	85.37	98.37
5th graders	66	10.29 (0.65)	63.64	93.94	100
Complex Subtraction					
General sample	527	8.65 (1.19)	53.32	88.99	97.53
2nd graders	91	6.99 (0.51)	48.35	100	100
3rd graders	161	8.07 (0.56)	47.20	80.75	99.38
4th graders	185	9.20 (0.49)	57.30	85.41	96.22
5th graders	90	10.26 (0.59)	38.89	100	94.44
Simple Multiplication					
General sample	248	8.96 (0.97)	57.66	89.11	97.18
3rd graders	102	8.14 (0.51)	55.88	89.22	100
4th graders	100	9.14 (0.45)	60.00	84.00	97.00
5th graders	46	10.39 (0.54)	56.52	100	91.30
Complex Multiplication					
General sample	416	9.14 (0.96)	55.77	90.38	97.36
3rd graders	119	8.07 (0.52)	50.42	89.92	100
4th graders	198	9.21 (0.48)	58.08	85.86	96.97
5th graders	99	10.29 (0.59)	57.58	100	94.95

Table 3. Descriptive analyses of participants' characteristics and percentage of errors in responded items for the group of children who did not commit any S5 error (GNoS5)

	N	%	Mean (sd) age	% female	% BH	% public school	% error in responded items
Simple Addition							
General sample	352	86.91	7.61 (1.52)	52.27	91.19	80.40	22.68
1st graders	113	79.58	5.96 (0.34)	46.02	100	41.59	30.18
2nd graders	55	90.16	6.91 (0.48)	54.55	100	100	19.06
3rd graders	76	92.68	7.96 (0.58)	46.05	76.32	97.37	22.88
4th graders	76	88.37	9.16 (0.49)	59.21	82.89	100	18.65
5th graders	32	94.12	10.13 (0.71)	68.75	100	96.88	11.53
Complex Addition							
General sample	328	83.46	8.69 (1.22)	52.13	89.94	98.48	25.30
2nd graders	62	68.13	7.02 (0.50)	59.68	100	100	29.39
3rd graders	89	83.18	8.08 (0.55)	43.82	74.16	98.88	31.56
4th graders	116	88.55	9.16 (0.44)	50.86	91.38	99.14	22.37
5th graders	61	95.31	10.36 (0.61)	59.02	100	95.08	17.56
Simple Subtraction							
General sample	359	85.48	8.40 (1.41)	54.04	88.02	95.26	28.68
1st graders	31	77.5	5.87 (0.43)	51.61	100	67.74	44.69
2nd graders	58	77.33	7.02 (0.55)	53.45	100	100	31.48
3rd graders	101	87.07	8.03 (0.62)	45.54	75.25	99.01	36.13
4th graders	109	88.62	9.18 (0.51)	57.80	83.49	98.17	20.60
5th graders	60	90.91	10.27 (0.66)	63.33	100	93.33	19.86
Complex Subtraction							
General sample	312	83.87	8.67 (1.22)	53.85	88.14	98.08	36.71
2nd graders	51	73.63	6.98 (0.51)	52.94	100	100	45.69
3rd graders	98	86.34	8.05 (0.60)	44.90	78.57	98.98	46.26
4th graders	106	83.78	9.15 (0.45)	57.55	84.91	98.11	29.47
5th graders	57	90.00	10.33 (0.66)	63.16	100	94.74	25.72
Simple Multiplication							
General sample	221	89.11	9.00 (0.51)	56.56	89.14	96.83	13.26
3rd graders	87	85.29	8.17 (0.51)	52.87	87.36	100	15.70
4th graders	91	91.00	9.11 (0.43)	60.44	85.71	96.70	12.40
5th graders	43	93.48	10.42 (0.55)	55.81	100	90.70	10.16
Complex Multiplication							
General sample	393	94.47	9.12(0.96)	56.49	90.59	97.20	63.53
3rd graders	115	96.64	8.06(0.52)	50.43	89.57	100	85.22
4th graders	185	93.43	9.20(0.46)	58.38	86.49	96.76	58.73
5th graders	93	93.94	10.28(0.46)	60.22	100	94.62	46.26

Table 4. Descriptive analyses of participants' characteristics and percentage of errors in responded items for the group of children who committed at least one S5 error (GS5)

	N	%	Mean (sd) age	% female	% BH	% public school	% error in responded items
Simple Addition							
General sample	53	13.09	7.00 (1.61)	56.6	94.34	77.36	49.18
1st graders	29	20.42	5.79 (0.49)	55.17	100	59.62	55.78
2nd graders	6	9.84	6.83 (0.41)	33.33	100	100	62.50
3rd graders	6	7.32	8.17 (0.41)	50.00	66.67	100	21.74
4th graders	10	11.63	9.20 (0.42)	70.00	90.00	100	45.83
5th graders	2	5.88	10.50 (0.71)	100	100	100	12.50
Complex Addition							
General sample	65	16.54	7.82 (1.13)	53.85	87.69	100	49.75
2nd graders	29	31.87	6.90 (0.49)	48.28	100	100	62.63
3rd graders	18	16.82	7.89 (0.76)	44.44	66.67	100	46.26
4th graders	15	11.45	9.07 (0.46)	66.67	86.67	100	31.71
5th graders	3	4.69	10.00 (0.00)	100	100	100	36.36
Simple Subtraction							
General sample	61	14.52	7.95 (1.40)	52.46	96.72	96.72	47.97
1st graders	9	22.50	6.00 (0.00)	66.67	100	77.78	73.15
2nd graders	17	22.67	6.88 (0.33)	41.18	100	100	68.30
3rd graders	15	12.93	8.33 (0.49)	60.00	86.70	100	36.10
4th graders	14	11.38	9.00 (0.00)	42.86	100	100	27.72
5th graders	6	9.09	10.50 (0.55)	66.67	100	100	21.84
Complex Subtraction							
General sample	85	16.13	8.40 (1,18)	60.00	90.60	97.65	64.88
2nd graders	24	26.37	6.96 (0.36)	45.83	100	100	83.75
3rd graders	22	13.66	8.14 (0.47)	72.73	77.27	100	72.69
4th graders	30	16.22	9.23 (0.57)	63.33	90.00	96.67	54.45
5th graders	9	10.00	10.11 (0.33)	55.56	100	88.89	30.19
Simple Multiplication							
General sample	27	10.89	8.67 (0.96)	66.67	88.89	100	31.48
3rd graders	15	14.71	7.93 (0.46)	73.33	100	100	29.28
4th graders	9	9.00	9.44 (0.53)	55.56	66.67	100	34.84
5th graders	3	6.52	10.00 (0.00)	66.67	100	100	32.38
Complex Multiplication							
General sample	23	5.53	9.48 (0.95)	43.48	86.96	100	78.71
3rd graders	4	3.36	8.25 (0.50)	50.00	100	100	87.82
4th graders	13	6.57	9.38 (0.65)	53.85	76.92	100	76.82
5th graders	6	6.06	10.50 (0.55)	16.67	100	100	76.95

Supplementary material B – Performance analyses (extras)

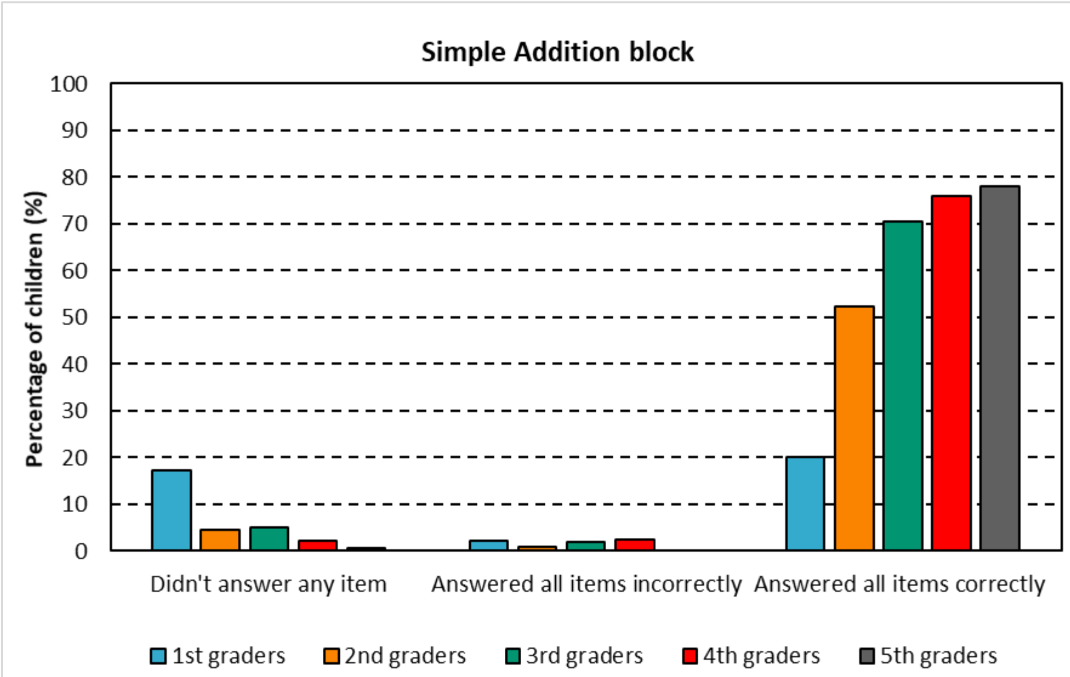

Figure 1. Percentage of children that did not respond to any item, responded to all items correctly and responded to all items incorrectly in the Simple Addition block, stratified by grade.

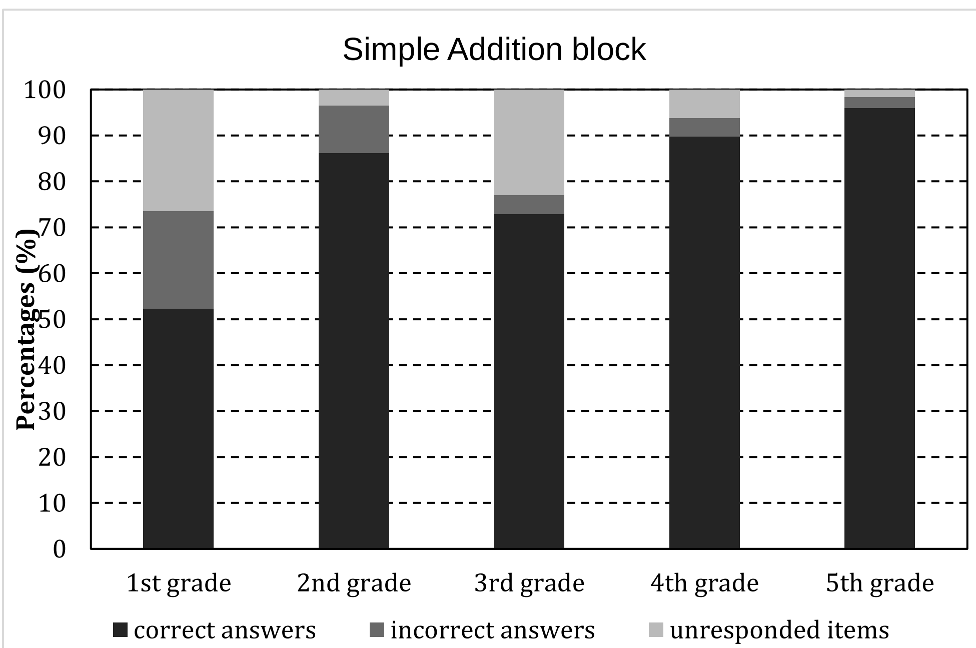

Figure 2. Percentage of unresponded, correct and incorrect items in the Simple Addition block, stratified by grade.

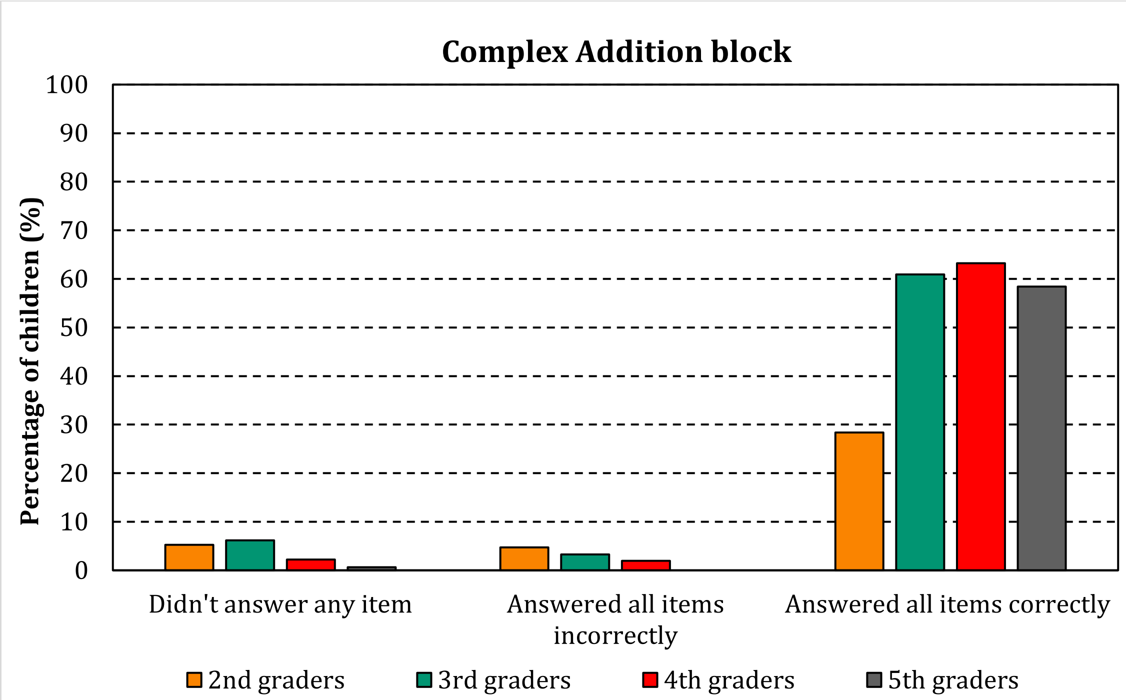

Figure 3. Percentage of children that did not respond to any item, responded to all items correctly and responded to all items incorrectly in the Complex Addition block, stratified by grade.

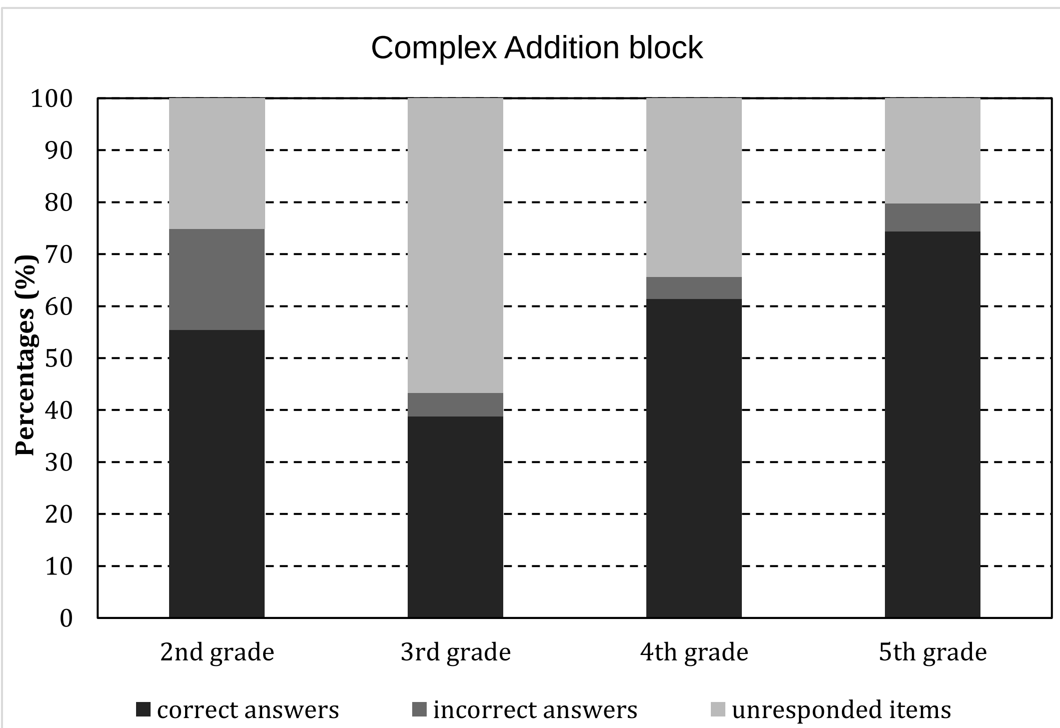

Figure 4. Percentage of unresponded, correct and incorrect items in the Complex Addition block, stratified by grade.

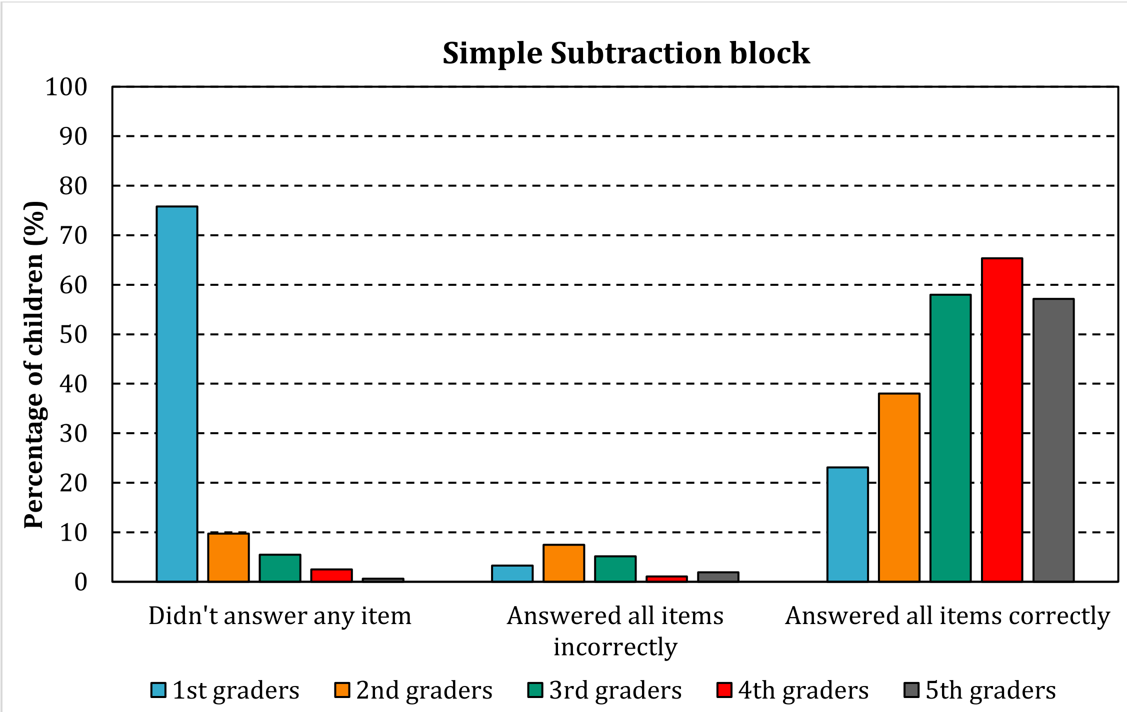

Figure 5. Percentage of children that did not respond to any item, responded to all items correctly and responded to all items incorrectly in the Simple Subtraction block, stratified by grade.

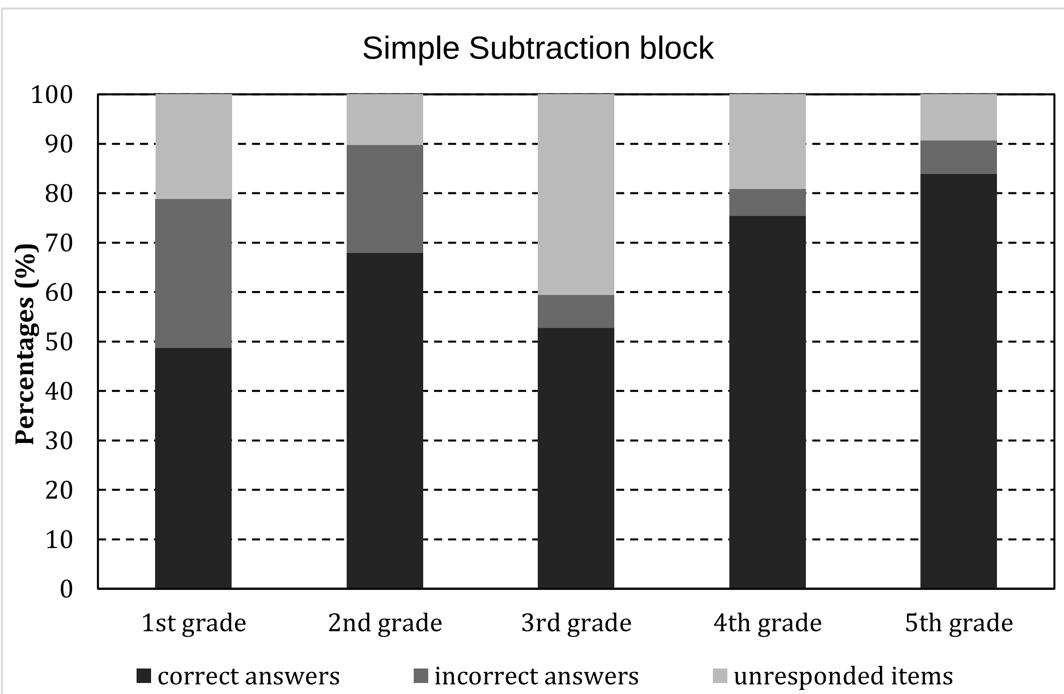

Figure 6. Percentage of unresponded, correct and incorrect items in the Simple Subtraction block, stratified by grade.

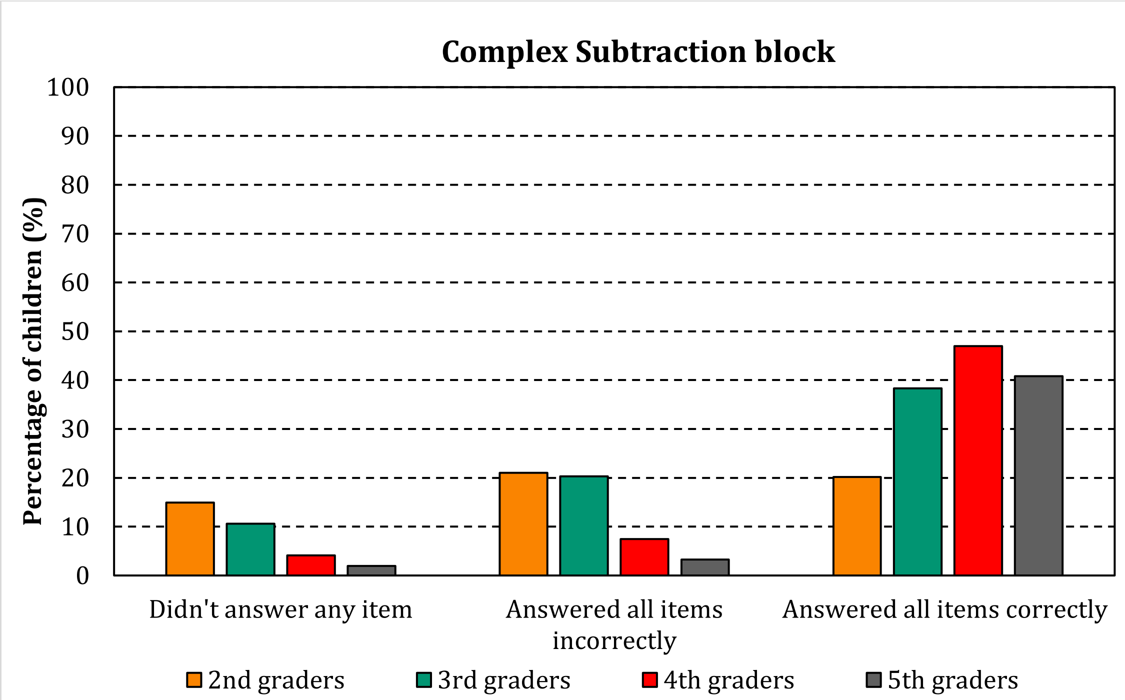

Figure 7. Percentage of children that did not respond to any item, responded to all items correctly and responded to all items incorrectly in the Complex Subtraction block, stratified by grade.

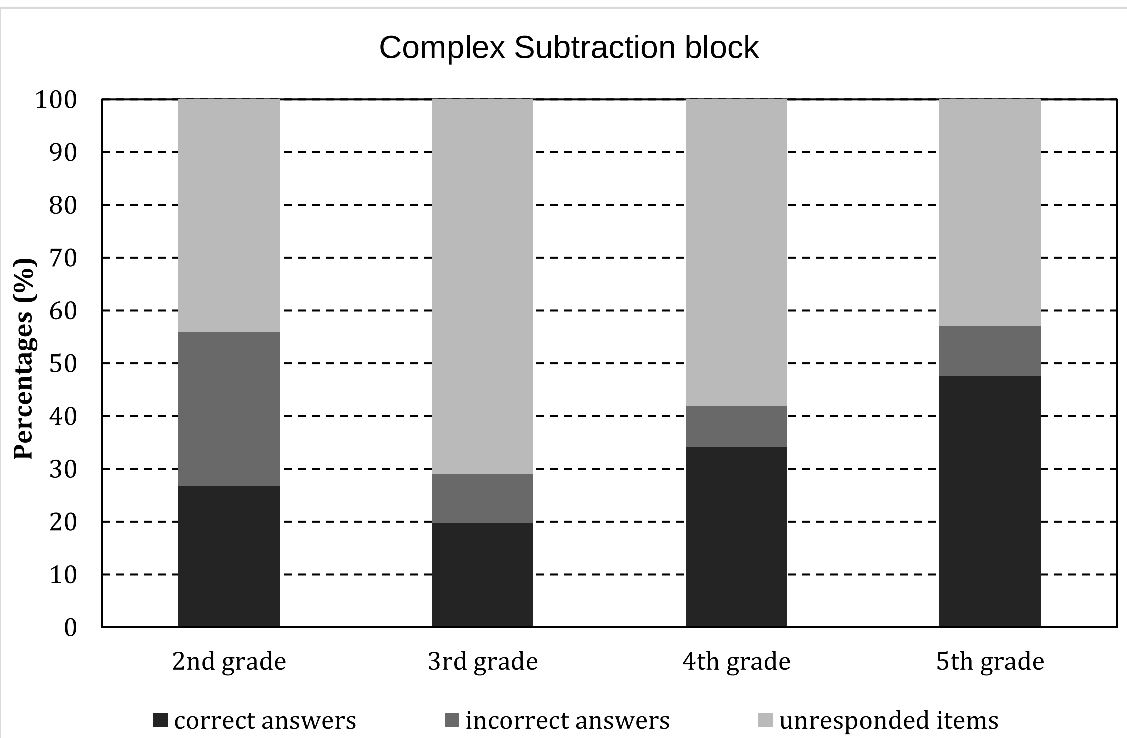

Figure 8. Percentage of unresponded, correct and incorrect items in the Complex Subtraction block, stratified by grade.

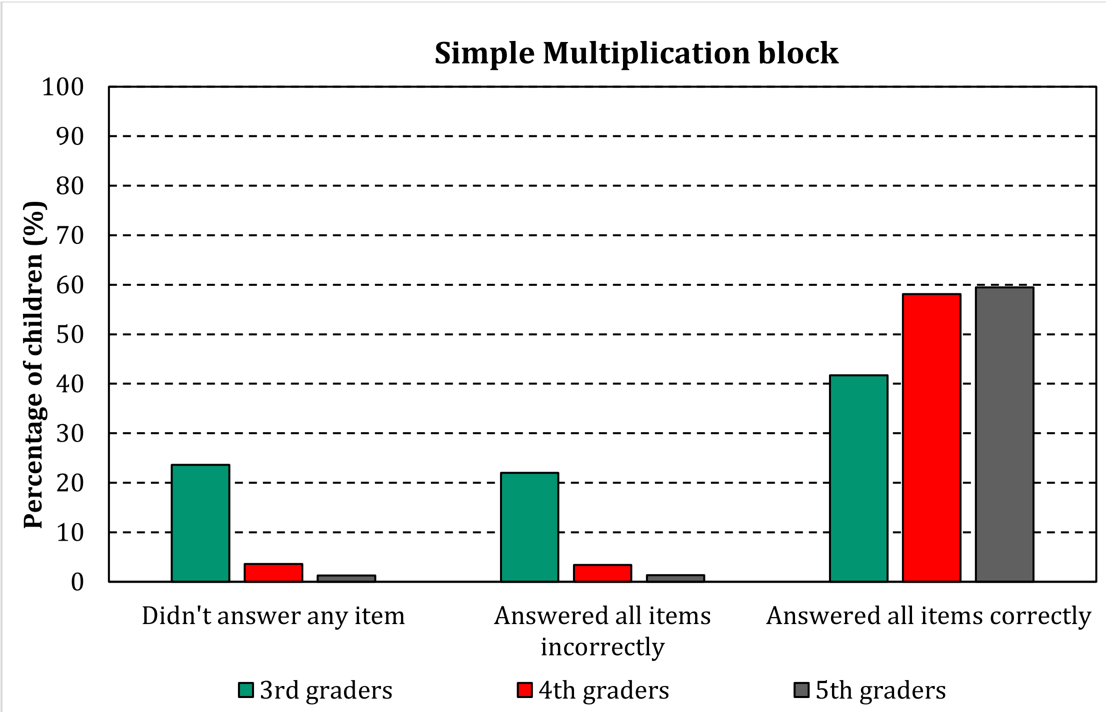

Figure 9. Percentage of children that did not respond to any item, responded to all items correctly and responded to all items incorrectly in the Simple Multiplication block, stratified by grade.

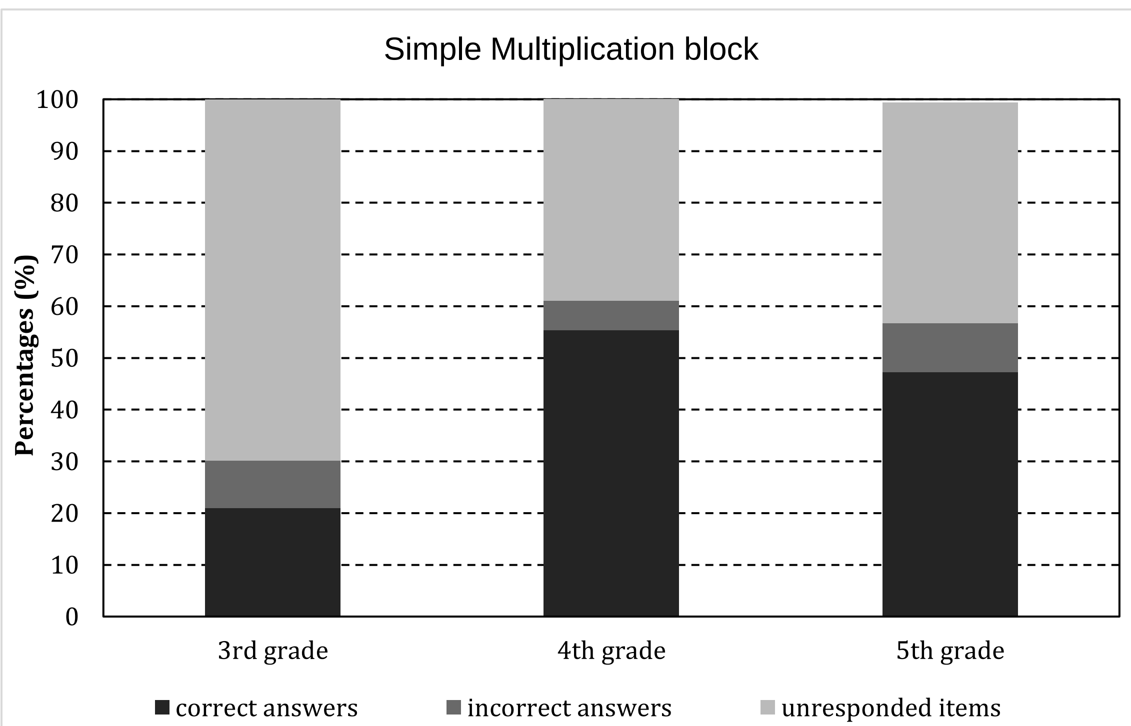

Figure 10. Percentage of unresponded, correct and incorrect items in the Simple Multiplication block, stratified by grade.

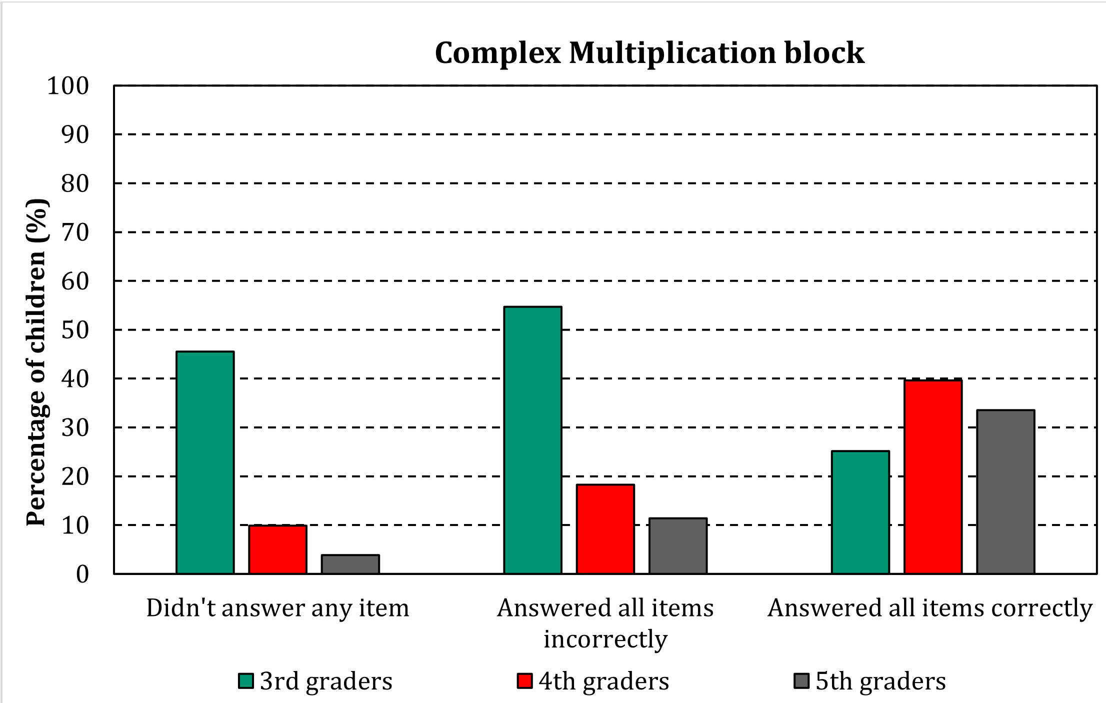

Figure 11. Percentage of children that did not respond to any item, responded to all items correctly and responded to all items incorrectly in the Complex Multiplication block, stratified by grade.

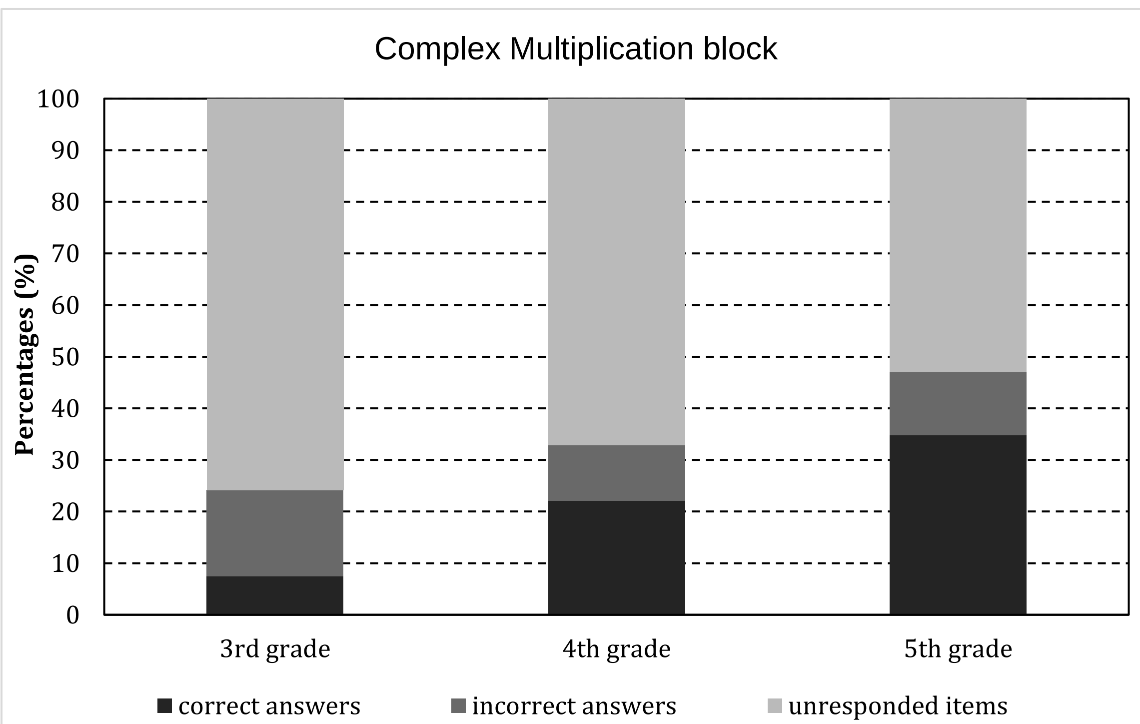

Figure 12. Percentage of unresponded, correct and incorrect items in the Complex Multiplication block, stratified by grade.

Supplementary material C – Mean percentage and stardard desviation of Splits errors for GNoS5 and GS5

Table 1. Mean and standard desviation of the percentage of Split errors (ranging from 1 to >11) for GNoS5 and GS5 in the Simple Addition block.

Children who did not commit S5 errors (GNoS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	352	Mean	65.80	13.94	10.08	3.50	0.00	1.22	1.17	0.77	1.86	0.21	0.62	0.85
		sd	41.42	29.41	25.55	16.12	0.00	8.16	7.09	6.63	11.77	2.98	5.93	6.92
1st grade	113	Mean	59.11	19.66	12.26	3.76	0.00	0.30	0.57	0.89	2.80	0.44	0.00	0.22
		sd	38.49	32.21	22.73	15.31	0.00	2.47	3.56	9.41	12.83	4.70	0.00	2.35
2nd grade	55	Mean	61.67	16.12	10.27	5.27	0.00	1.21	0.45	0.45	1.82	0.00	0.00	2.73
		sd	43.63	31.75	27.32	19.93	0.00	8.99	3.37	3.37	13.48	0.00	0.00	14.96
3rd grade	76	Mean	63.57	11.92	11.59	4.17	0.00	1.07	3.26	1.32	1.50	0.33	0.41	0.88
		sd	43.49	29.07	30.18	17.66	0.00	6.21	13.52	6.95	11.57	2.87	2.50	6.02
4th grade	76	Mean	69.84	9.75	9.43	2.63	0.00	3.25	0.97	0.60	1.64	0.00	1.13	0.76
		sd	44.13	26.95	27.96	16.11	0.00	14.12	4.41	4.07	11.61	0.00	5.03	3.89
5th grade	32	Mean	92.19	4.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.13	0.00
		sd	22.39	14.81	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	17.68	0.00
Children who commited S5 errors (GS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	53	Mean	11.27	10.09	6.45	5.12	49.15	6.82	2.19	2.87	1.16	0.00	1.93	2.96
		sd	16.04	12.61	10.69	8.52	35.15	12.97	5.83	7.16	3.43	0.00	6.09	8.74
1st grade	29	Mean	12.75	11.15	7.03	7.32	45.23	4.77	2.73	3.65	0.66	0.00	0.92	3.79
		sd	14.90	11.94	10.32	9.80	34.63	8.37	7.22	8.16	2.46	0.00	3.55	11.29
2nd grade	6	Mean	15.08	16.55	10.34	1.39	37.10	7.41	0.00	7.69	2.78	0.00	0.00	1.67
		sd	22.17	13.80	8.52	3.40	34.36	8.36	0.00	9.94	6.80	0.00	0.00	4.08
3rd grade	6	Mean	8.33	12.50	0.00	4.17	59.72	15.28	0.00	0.00	0.00	0.00	0.00	0.00
		sd	20.41	20.92	0.00	20.21	32.67	27.09	0.00	0.00	0.00	0.00	0.00	0.00
4th grade	10	Mean	3.69	3.69	7.58	2.58	56.26	8.69	3.69	0.00	2.58	0.00	7.58	3.69
		sd	4.82	4.82	15.45	4.15	39.94	15.24	4.82	0.00	4.15	0.00	11.43	4.82
5th grade	2	Mean	25.00	0.00	0.00	0.00	75.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		sd	35.36	0.00	0.00	0.00	35.36	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 2. Mean and standard desviation of the percentage of Split errors (ranging from 1 to >11) for GNoS5 and GS5 in the Complex Addition block.

Sample	N		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S>11
Children who did not commit S5 errors (GNoS5)														
General Sample	328	Mean	67.97	13.02	6.87	3.64	0.00	2.42	1.27	0.49	0.23	0.86	0.05	3.18
		sd	41.06	27.73	21.72	15.72	0.00	13.54	10.06	4.23	2.94	8.25	0.92	16.44
2nd grade	62	Mean	65.33	12.91	8.39	5.81	0.00	1.34	2.02	0.97	0.94	0.27	0.27	1.75
		sd	37.00	22.61	21.79	16.68	0.00	7.57	12.86	6.46	6.42	2.12	2.12	12.73
3rd grade	89	Mean	61.32	15.73	7.58	5.06	0.00	3.18	1.12	0.75	0.00	1.12	0.00	4.12
		sd	45.82	33.08	23.36	20.01	0.00	16.06	10.60	4.97	0.00	10.60	0.00	19.35
4th grade	116	Mean	69.66	13.46	5.68	2.87	0.00	1.32	1.03	0.29	0.14	1.44	0.00	4.11
		sd	41.22	28.45	20.19	14.79	0.00	9.92	9.45	3.09	1.55	10.21	0.00	18.91
5th grade	61	Mean	77.16	8.33	6.56	0.82	0.00	4.51	1.48	0.00	0.00	0.00	0.00	1.48
		sd	36.00	21.94	22.37	6.40	0.00	19.10	6.86	0.00	0.00	0.00	0.00	8.13
Children who commited S5 errors (GS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	65	Mean	15.41	5.44	5.41	4.40	55.90	3.34	1.89	0.99	1.52	0.64	0.85	4.20
		sd	21.48	11.49	9.94	11.12	35.20	8.86	7.77	3.03	4.77	2.33	3.30	12.83
2nd grade	29	Mean	20.46	6.71	8.43	4.52	38.71	4.85	2.52	1.79	3.40	1.43	1.91	5.27
		sd	21.32	10.42	11.22	11.32	30.71	8.88	7.10	3.79	6.74	3.36	4.77	11.13
3rd grade	18	Mean	6.90	6.90	2.28	6.81	68.25	1.49	0.00	0.69	0.00	0.00	0.00	6.67
		sd	17.15	16.17	7.19	14.56	34.38	2.34	0.00	2.95	0.00	0.00	0.00	19.70
4th grade	15	Mean	16.94	1.67	3.06	0.83	70.83	3.33	3.33	0.00	0.00	0.00	0.00	0.00
		sd	25.42	6.45	8.97	3.23	30.42	12.91	12.91	0.00	0.00	0.00	0.00	0.00
5th grade	3	Mean	10.00	3.33	6.67	6.67	73.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		sd	17.32	5.77	11.55	11.54	46.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 3. Mean and standard desviation of the percentage of Splits errors (ranging from 1 to >11) for GNoS5 and GS5 in the Simple Subtraction block.

Children who did not commit S5 errors (GNoS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	359	Mean	66.55	11.27	5.88	5.57	0.00	3.12	1.03	1.25	0.19	2.45	0.10	2.60
		sd	42.87	27.37	20.56	18.40	0.00	13.18	8.04	7.05	1.57	14.22	1.30	11.62
1st grade	31	Mean	53.35	17.35	15.51	5.45	0.00	3.14	1.10	1.08	0.67	0.00	0.00	2.36
		sd	40.15	27.68	24.82	13.79	0.00	7.79	3.62	4.69	2.66	0.00	0.00	7.83
2nd grade	58	Mean	60.50	13.64	4.97	2.60	0.00	4.55	2.49	1.42	0.49	7.81	0.00	1.54
		sd	42.83	29.74	16.66	7.17	0.00	18.69	13.59	4.84	2.14	23.88	0.00	6.23
3rd grade	101	Mean	62.96	10.56	4.10	10.77	0.00	4.44	0.36	0.88	0.20	0.00	0.14	4.70
		sd	46.12	28.20	19.62	26.38	0.00	13.32	2.58	3.98	1.99	7.00	1.42	16.39
4th grade	109	Mean	74.82	10.05	5.13	2.31	0.00	1.78	0.49	1.31	0.00	1.77	0.18	2.18
		sd	40.34	26.89	20.47	11.63	0.00	10.57	3.71	9.84	0.00	10.89	1.92	11.82
5th grade	60	Mean	70.28	9.44	6.11	5.69	0.00	1.94	1.67	1.67	0.00	2.22	0.00	0.97
		sd	41.07	24.43	22.54	20.15	0.00	13.05	12.91	7.94	0.00	13.19	0.00	5.34
Children who commited S5 errors (GS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	61	Mean	15.48	7.10	9.67	8.60	49.47	5.55	1.09	1.01	0.00	0.00	0.00	2.02
		sd	20.96	11.57	16.03	13.94	34.17	12.59	3.54	3.87	0.00	0.00	0.00	7.57
1st grade	9	Mean	8.50	10.23	13.51	4.88	40.07	15.15	1.94	2.02	0.00	0.00	0.00	3.70
		sd	13.30	12.32	18.35	7.72	36.31	12.42	3.85	4.01	0.00	0.00	0.00	11.11
2nd grade	17	Mean	17.33	10.37	18.72	10.62	30.13	6.03	2.89	1.08	0.00	0.00	0.00	2.84
		sd	20.40	11.52	16.65	11.28	27.95	12.46	5.76	3.06	0.00	0.00	0.00	9.79
3rd grade	15	Mean	14.44	3.89	4.44	6.67	69.44	0.00	0.00	0.00	0.00	0.00	0.00	1.11
		sd	23.46	8.25	13.31	15.17	35.31	0.00	0.00	0.00	0.00	0.00	0.00	4.30
4th grade	14	Mean	11.91	4.76	3.57	10.12	58.93	7.14	0.00	1.79	0.00	0.00	0.00	1.79
		sd	20.07	12.10	13.36	18.54	28.95	18.16	0.00	6.68	0.00	0.00	0.00	6.68
5th grade	6	Mean	31.67	6.67	5.56	9.72	46.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		sd	25.83	16.33	13.61	15.29	29.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 4. Mean and standard desviation of the percentage of Splits errors (ranging from 1 to >11) for GNoS5 and GS5 in the Complex Subtraction block.

Children who did not commit S5 errors (GNoS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	442	Mean	53.89	21.26	7.37	6.27	0.00	3.84	0.72	1.82	0.30	1.41	0.33	2.79
		sd	42.57	34.38	21.14	19.33	0.00	15.45	6.16	9.57	2.63	7.96	3.60	13.84
2nd grade	67	Mean	53.57	15.71	7.39	8.43	0.00	4.00	1.44	3.51	0.37	1.92	0.42	3.22
		sd	40.31	27.25	17.26	20.04	0.00	14.80	6.03	14.39	3.05	7.70	2.63	13.68
3rd grade	139	Mean	53.06	20.40	6.97	4.96	0.00	4.11	0.29	2.88	0.65	1.04	0.72	4.93
		sd	43.84	34.50	21.58	18.34	0.00	16.56	2.39	12.48	3.93	5.87	5.98	19.04
4th grade	155	Mean	57.40	21.26	8.21	4.78	0.00	3.25	0.86	1.03	0.00	1.80	0.11	1.31
		sd	43.40	35.86	24.00	16.54	0.00	14.09	8.45	5.38	0.00	10.31	1.34	9.05
5th grade	81	Mean	48.86	27.35	6.46	9.57	0.00	4.370	0.62	0.14	0.21	0.86	0.00	1.58
		sd	40.68	36.20	17.39	24.49	0.00	16.70	5.56	1.23	1.85	5.96	0.00	10.22
Children who commited S5 errors (GS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	85	Mean	15.02	11.16	6.32	5.34	47.39	3.49	2.53	2.45	2.34	1.15	0.54	2.27
		sd	20.06	17.56	12.30	9.63	35.80	8.02	5.99	6.52	7.42	7.56	2.61	8.60
2nd grade	24	Mean	15.57	13.39	10.21	9.87	19.20	7.01	6.49	6.72	4.19	1.28	1.11	4.84
		sd	15.00	11.96	9.98	10.40	20.83	7.25	8.07	9.67	7.51	4.34	3.50	9.81
3rd grade	22	Mean	19.02	6.52	7.88	3.94	49.70	1.29	2.05	1.82	3.56	0.00	0.76	3.48
		sd	24.07	15.10	14.74	8.96	31.20	3.49	6.00	5.88	11.32	0.00	3.55	13.11
4th grade	30	Mean	15.27	11.74	3.30	3.68	61.19	1.44	0.48	0.00	0.67	2.22	0.00	0.00
		sd	22.10	19.36	12.66	35.63	35.63	6.29	2.61	0.00	3.65	12.17	0.00	0.00
5th grade	9	Mean	2.96	14.63	2.22	2.22	70.93	6.30	0.00	0.74	0.00	0.00	0.00	0.00
		sd	8.89	27.94	6.67	6.67	36.28	16.54	0.00	2.22	0.00	0.00	0.00	0.00

Table 5. Mean and standard desviation of the percentage of Split errors (ranging from 1 to >11) for GNoS5 and GS5 in the Simple Multiplication block.

Sample	N		S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S>11
Children who did not commit S5 errors (GNoS5)														
General Sample	221	Mean	16.29	14.74	12.57	13.75	0.00	9.99	2.08	1.51	6.17	1.66	0.83	7.97
		sd	35.35	33.29	30.38	31.83	0.00	27.95	13.57	10.63	22.30	10.90	5.92	24.63
3rd grade	87	Mean	17.00	19.83	12.61	14.43	0.00	5.59	4	1.98	7.63	2.81	1.85	5.31
		sd	36.28	36.88	31.35	32.08	0.00	19.00	18.65	12.01	24.77	13.45	9.23	18.69
4th grade	91	Mean	17.16	6.72	10.60	14.66	0.00	12.21	1.22	0.55	5.49	1.34	0.12	13.50
		sd	35.88	23.81	26.35	32.67	0.00	31.01	10.53	5.24	20.35	10.71	1.16	32.56
5th grade	43	Mean	13.05	21.45	16.67	10.47	0.00	14.21	0.00	2.58	4.65	0.00	0.26	1.68
		sd	32.84	39.56	36.19	30.00	0.00	35.00	0.00	15.30	21.31	0.00	1.69	8.27
Children who commited S5 errors (GS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	27	Mean	3.24	4.82	3.09	2.26	50.81	4.57	3.28	2.01	7.69	0.74	1.23	7.95
		sd	8.52	11.07	11.36	6.89	30.52	11.66	7.27	5.87	12.87	3.85	6.42	16.52
3rd grade	15	Mean	0.00	5.78	3.33	0.00	56.67	8.22	1.11	1.33	10.22	1.33	2.22	4.67
		sd	0.00	13.83	12.91	0.00	30.39	14.85	4.30	5.16	15.30	5.16	8.61	10.22
4th grade	9	Mean	5.56	3.44	3.70	5.40	46.88	0.00	3.81	3.81	1.85	0.00	0.00	16.08
		sd	11.79	6.85	11.11	10.92	34.09	0.00	7.69	7.69	5.56	0.00	0.00	24.20
5th grade	3	Mean	12.50	4.17	0.00	4.17	33.33	0.00	12.50	0.00	12.50	0.00	0.00	0.00
		sd	12.50	7.22	0.00	7.22	14.43	0.00	12.50	0.00	12.50	0.00	0.00	0.00

Table 6. Mean and standard desviation of the percentage of Split errors (ranging from 1 to >11) for GNoS5 and GS5 in the Complex Multiplication block.

Children who did not commit S5 errors (GNoS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	393	Mean	13.28	8.97	5.87	5.01	0.00	8.57	2.75	9.33	4.57	2.95	1.01	37.70
		sd	30.47	24.90	20.20	18.60	0.00	24.44	13.47	26.08	17.91	14.35	7.03	42.30
3rd grade	115	Mean	6.41	5.90	2.39	2.83	0.00	5.25	3.07	9.59	2.54	2.27	2.18	57.58
		sd	22.24	20.31	12.13	12.87	0.00	20.05	14.63	27.48	12.25	12.22	11.64	44.57
4th grade	185	Mean	16.99	9.01	8.77	5.66	0.00	8.34	1.85	10.82	5.05	3.94	0.56	29.02
		sd	33.79	25.32	24.93	10.79	0.00	23.85	10.35	27.77	19.53	17.60	3.83	38.86
5th grade	93	Mean	14.39	12.69	4.39	6.43	0.00	13.11	4.15	6.04	6.13	1.83	0.44	30.40
		sd	31.14	28.68	16.74	21.84	0.00	29.58	17.03	20.13	20.25	8.30	3.31	38.00
Children who commited S5 errors (GS5)														
Sample	**N**		**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**	**S8**	**S9**	**S10**	**S11**	**S>11**
General Sample	23	Mean	3.91	3.64	3.26	0.00	43.18	2.94	2.78	4.91	2.98	3.04	0.00	29.35
		sd	11.58	7.69	11.44	0.00	36.54	12.54	6.20	10.58	6.10	8.08	0.00	32.15
3rd grade	4	Mean	0.00	0.00	0.00	0.00	59.38	0.00	9.38	3.13	0.00	6.25	0.00	21.88
		sd	0.00	0.00	0.00	0.00	47.19	0.00	11.97	6.25	0.00	12.50	0.00	29.54
4th grade	13	Mean	3.85	0.85	5.77	0.00	50.40	0.00	1.45	1.54	2.55	1.92	0.00	31.67
		sd	13.87	3.08	14.98	0.00	37.09	0.00	3.60	5.55	5.02	6.93	0.00	34.10
5th grade	6	Mean	6.67	12.12	0.00	0.00	16.73	11.28	1.28	13.40	5.90	3.33	0.00	29.29
		sd	10.33	10.99	0.00	0.00	7.26	24.06	3.14	16.64	9.25	8.16	0.00	34.28