Elementary students' conditional reasoning skills: The case of Mathematics

Dissertation zur Erwerb des Doctor of Philosophy (Ph.D)

am Munich Center of the Learning Sciences

der Ludwig−Maximilians−Universität München

Vorgelegt von

Anastasia Datsogianni

15.06.2020

1st Supervisor / Erstgutachter: Prof. Dr. Stefan Ufer

2nd Supervisor / Zweitgutachter: Prof. Dr. Beate Sodian

Tag der mündlichen Prüfung: 23.07.2020

Acknowledgements

This journey has come to the end as some people stood by me during the whole duration of it and now it is the time to 'officially' express my gratitude to them.

First of all, I cannot really express by words how grateful I am to my first supervisor, Prof. Dr. Stefan Ufer for his unlimited support, assistance and encouragement. He was always responsive, and willing to share his admirable knowledge. He is undoubtedly for me a role model of researcher and instructor in the field of Mathematics Education. Thank you from the bottom of my heart, Prof. Ufer, for making this 4-year PhD experience at LMU so unique.

I would also like to thank my second supervisor, Prof. Dr. Beate Sodian, for her contribution to my research and her constructive input from the field of Developmental and Educational Psychology.

Let me truly express my gratitude to my international supervisor Prof. Dr. Henry Markovits. Without having his extensive previous work as a foundation of my thesis, I would probably not decide to explore this incredible research field of conditional reasoning. He was always supportive and boosted my research with his ideas and advice, and I am so honoured that I had the opportunity to work in his lab during my research stay in Canada. He is definitely for me the most inspirational researcher in the area of logical reasoning. I would like also to thank all the school principals, teachers and students' parents who endorsed my research and all the students in Cyprus who participated voluntarily in my studies there.

I could not close this text without mentioning my family and friends. I would like to say a huge thank you to my loved grandparents here, Dimitrios and Anastasia, who supported me every second of my life in Munich. They are among the most powerful people I have ever met and dedicated their life to supporting me and my brother in our studies and in general; words cannot express my gratitude to them. I would like to thank also my brother for always providing me with his valuable advice, and my lovely aunts and cousins here in Munich for their love and support. My deepest thanks to my parents, Ntina and Kostas who stood by me and gave me courage in every difficulty I might faced, being always so optimistic; blessed to be their daughter.

This academic journey would not be so exceptional without my REASON friends. In particular, I would like to thank my friend Olga, my lovely witty office-mate, who was always so positive with this great sense of humour; I will never forget the moments we were laughing so hard. My friend Despoina, the honest and trustworthy friend, with whom I spent so much beautiful time in and out of office; she was always the voice of logic for me. I thank both of them for their -incredible- support during my life in Munich. Many thanks to Diana, Mary, Gina and Arianne for spending together memorable lunch breaks, unofficial meetings and always having such interesting discussions.

Last but not least, is my partner, Charis, who always encourages me to follow my dreams; with whom I can smile in every difficult moment and experience the real meaning of life; thank you so much for everything.

Table of Contents

ABSTRACT

Reasoning about conditional “if..then” statements is a central component of logical reasoning. However, a research link between conditional reasoning and mathematics has been reported only for late adolescence and adults (Attridge & Inglis, 2013; Stylianides, Stylianides, & Philippou, 2004; Durand-Guerrier, 2003), despite claims about the pivotal importance of conditional reasoning, i.e. reasoning with if-then statements, in mathematics. To address this issue and shed some light on the the area of conditional reasoning within mathematics in elementary school, three studies were conducted to measure students conditional reasoning and alternative generation skills in two contexts (everyday and mathematical) and investigate various factors (i.e, age, logical form, working memory capacity, alternative generation skills) that might affect conditional reasoning skills at these ages, as well as the potential scaffolding function of different trainings on these skills. Firstly, after having approached the background that frames conditional reasoning in mathematics and everyday context, we reported on a study that explored if it is feasible to survey conditional reasoning skills in everyday contexts and mathematics with primary school students. The findings shown that the applied instrument was accessible to students, and reflected central predictions of Mental Model Theories of conditional reasoning for differences between the two contexts. Moreover, the question, if the ability to generate examples of mathematical concepts, and to generate multiple alternative models for a given premise, has an influence on students’ conditional reasoning with these concepts, was raised at that point. In this direction this pilot study also aimed at investigating students’ alternative generation outcomes in both contexts. Based on the aforementioned pilot study, the first study addressed the open question, to which extent conditional reasoning with mathematical concepts differs from conditional reasoning in familiar everyday contexts. This study also examined the role of alternatives generation skills on conditional reasoning within an everyday and a mathematical context. The results of study 1 suggest that,

consistently with previous findings, even 2nd graders were able to make correct inferences on some logical forms. Controlling for WM, there were significant effects of grade and logical form, with stronger growth on MP and AC than on MT and DA. The main effect of context was not significant, but context interacted significantly with logical form and grade level. The pattern of results was not consistent with the predictions of MMT. The study also indicates that deductive reasoning skills arise from a combination of knowledge of domain-general principles and domain-specific knowledge. In addition, it extends results concerning the gradual development of primary students' conditional reasoning with everyday concepts (Markovits & Barrouillet, 2002) to reasoning with mathematical concepts adding to our understanding about the link between mathematics and conditional reasoning in primary school. Moreover, alternatives generation skills predict correct conditional reasoning in both contexts, but interesting differences occurred. The findings from the everyday context mirror previous results, predicting correct AC and DA reasoning and inhibiting correct MT reasoning. In the mathematical context, alternatives generation predicted correct reasoning in all forms. The main contribution of study 1 is the emphasis on the specific role of mathematical knowledge in conditional reasoning with mathematical concepts. The results of the latter studies inspired the development of a short-term educational intervention. This goal was addressed in study 2 by investigating the effects of two short-term trainings based on alternative generation priming within two contexts (contrary-to-fact and mathematical contexts). The results of this study were mixed, revealing a decrease of definite reasoning scores after the short interventions and an increase in DA reasoning; however, further analysis is required. Ultimately, the studies in this dissertation aimed to gain some evidence in the area of conditional reasoning within mathematics in primary school and contribute to future research on this research field.

1. GENERAL INTRODUCTION

Logical reasoning skills even from very early ages are considered an essential tool not only for encountering daily phenomena and situations logically based on appropriate decision making processes but also for mathematics learning and success. In particular, logical reasoning is considered to be at the heart of critical thinking, analytical thinking and problem solving skills (Liu, Ludu & Holton, 2015) and a pivotal part of advanced thinking amidst human species (Markovits & Barrouillet, 2002). One of the main aspects of logical reasoning is conditional reasoning, i.e. reasoning with if-then statements. Research in the field of developmental psychology has outlined relevant schemes in young students and some clear developmental trajectories, but also a great deal of variation in performance due to a wide range of external parameters.

Moreover, a link between conditional reasoning and mathematics has been established only in the case of late adolescence and adults (Attridge & Inglis, 2013; Inglis & Simpson, 2009). Our knowledge about elementary students' conditional reasoning within mathematics is still vague, despite claims in the literature about the paramount importance of logical reasoning for mathematics learning and success (e.g., Morsanyi & Szücs, 2014; Nunes et al., 2007). Given that current theories describe conditional reasoning as a process that involves domain-specific knowledge, it is an open question to which extent students can apply their existing logical reasoning knowledge to learning environments that involve mathematical concepts.

Therefore, this thesis aims to shed some light on the topic of conditional reasoning within mathematical concepts in primary school. To this end, pilot study tested the feasibility of the designed mathematical conditional reasoning tasks for the age group between grade 2 and grade 6. Followingly, study 1 systematically investigated elementary students' conditional reasoning skills within mathematics and everyday context as well as students' alternative generation skills in these two contexts. Factors that affect students' performance were investigated thoroughly. After study 1, a short intervention study followed in order to explore the instructional methods that might help students to develop their conditional reasoning skills.

Before the description of the two studies, a General Introduction precedes presenting the background information that theoretically frames the two studies. The concepts and processes presented in this section will enhance the interpretation and discussion of the results in pilot study, study 1 and study 2. The chapter begins with section 1.1, which describes the concepts of deductive reasoning and conditional reasoning. In section 1.2 the theoretical accounts of conditional reasoning are discussed, whereas section 1.3 approaches conditional reasoning particularly in primary school from a developmental perspective and different contexts. The latter section also discusses the main process that underlies conditional reasoning which is alternative generation process separately as well as in connection to creativity. Subsequently, section 1.4 includes a description of trainings on students' conditional reasoning based on prior research and this section introduces not only the background but also the rationale behind study 2.

1.1.DEDUCTIVE REASONING

It is generally accepted that conditional reasoning as a deductive reasoning type is driven by a mix of psychological and social factors all of which should be taken into account when investigating this type of reasoning. In order to capture most of those factors, we need firstly to define the terms of deductive reasoning and conditional reasoning with regard to primary school students within the everyday context but also the mathematical one which is the core -under investigation- domain of this thesis. First of all, according to Johnson Laird (1999) reasoning is a systematic process of thought that starting from percepts, thoughts, or assertions derives a conclusion. Deductive reasoning requires the use of intent to generate valid conclusions from the given true premises (Wharton & Grafman, 1998) and based on research from psychology, students' deductive reasoning skills develops with age (e.g., Byrne & Overton, 1988).

1.1.1. Deductive reasoning and Mathematics

The ability to make valid deductions is considered to be a core part also of mathematical thinking (Moshman, 1990; Markovits & Lortie-Forgues, 2011), problem-solving skills (English & Halford, 1995), learning and success (Morsanyi & Szücs, 2014; Nunes et al., 2007). In fact, in mathematics, deduction is strongly associated with the development of students' ability to successfully process formal proofs (Foltz, Overton, & Ricco, 1995), and it is perceived as the core aspect of reasoning in mathematical theories. According to NCTM (1991) fostering the development of students' deductive reasoning skills is crucial in supporting them to derive conclusions and justify statements on their own (NCTM, 1991). A detailed example of deductive reasoning is conditional reasoning, which occurs not only in scientific and mathematical context, but also in everyday language and communication discourse.

From the perspective of mathematics, conditional reasoning has been thoroughly discussed with adults (Attridge & Inglis, 2013; Inglis & Simpson, 2009) and secondary school students, with results stressing difficulties in reasoning with conditionals in mathematics (e.g., Küchemann, & Hoyles, 2002). In spite of the potential importance of this form of reasoning, there is very little evidence about elementary students' abilities to reason with conditionals within mathematical instruction. However, even basic mathematical concepts are characterised by a strict logical structure, which makes deduction a principal component of reasoning in mathematics at every educational instruction. When it comes to conditional reasoning with mathematical concepts, it is suggested that knowledge of these concepts is mainly realised during primary school. Thus, at this age level, we see a parallel development of knowledge about mathematical concepts, as well as a general improvement in conditional reasoning skills. Learning to deal with mathematical concepts involves understanding and applying their properties correctly, which often have the form of conditional statements (e.g., '*Assume I have added two numbers. If I increase one of the two numbers by one and I decrease the other by one, then the two new numbers will have the same sum.*'). It is still to be answered to which extent making inferences with such statements does not only require conditional reasoning skills, but also mathematical knowledge about the data obtained. Although previous results have shown that secondary students face difficulties when dealing with mathematical conditional reasoning (e.g, Küchemann & Hoyles, 2002), psychological studies show that even very young children can resort to reasonable assumptions with conditionals with familiar everyday contents (e.g., Markovits & Thompson, 2008). This notion raises the question if and how conditional reasoning with mathematical concepts differs from conditional reasoning with everyday contents.

1.1.2. Conditional reasoning

Conditional reasoning tasks usually consist of a conditional rule of the form "if p then q" and a minor premise. Four different minor premises differentiate four possible logical forms of inference: p is true (Modus Ponens, MP), p is false (Denial of the Antecedent, DA), q is true (Affirmation of the Consequent, AC), and q is false (Modus Tollens, MT). Our work is based on the traditional interpretation of conditionals: p is sufficient, but not necessary for q (Evans & Over, 2004). Depending on the logical form (minor premise), different inferences can be made: Definite conclusions can be drawn for MP ("q is true") and MT ("p is false"), while AC resp. DA do not allow for definite conclusions about p resp. q.

Example (see also table 1): *If Mary jumps into the swimming pool, she will be wet* (Major premise/ major rule)

John met Mary at the swimming pool and she was wet. (Minor premise; Affirmation of the consequence)

Based on what he knows, what can John say for sure?

-*Mary jumped into the swimming pool.*

-*Mary did not jump into the swimming pool*

-*He cannot be sure whether Mary jumped into the swimming pool or not.*

John just met Mary and she was not wet. (Minor premise; Modus Tollens)

John was told that Mary did not jump into the swimming pool. (Minor premise; Denial of the antecedent)

John was told that Mary just jumped into the swimming pool. (Minor premise; Modus Ponens)

1.1.2.1.Logical forms (conditional inferences)

The current studies of this dissertation will be focused on four types of conditional - deductive syllogisms. These syllogisms have the following structure Conditional inferences are based on a conditional rule of the form 'if p, then q' that were described in the previous section

If a glass is dropped on the ground, then there is a sound. [If p, the q]

There is a sound. [q]

Conclusion: *We cannot be sure whether a glass was dropped on the ground or not.* [p or not p]

The first two sentences are the premises whereas the third one is the conclusion. Between the two premises the first one is the major and the second is the minor premise. The major premise consists of the first part, the antecedent (e.g. If a glass is dropped on the ground) and the consequent (e.g. there is a sound). Depending of the form of the minor premise, four different types of syllogisms arise. In other words, different minor premises differentiate four possible logical forms. These four forms are presented also in table 1.

Table 1

The logical forms (minor premises) under the major rule 'if p, then q'

p	Q	Modus Ponens (definite conclusion)
q	p or not p	Acceptance of the consequence (indefinite conclusion)
not p	q or not p	Denial of the antecedent (indefinite conclusion)
not q	not p	Modus Tollens (definite conclusion)

Based on the traditional interpretation of conditionals that p is sufficient, but not necessary for q (Evans & Over, 2004), and depending on the logical form (minor premise), different conclusions can be drawn. Definite conclusions are possible for Modus Ponens ('p is true') and Modus Tollens ('q is false'), while Acceptance of Consequent ('q is true') and Denial of Antecedent ('p is false) do not allow definite conclusions about p and q, respectively, and the correct conclusion in both cases is indefinite.

Across studies on reasoning in everyday contexts, the expected correct answer is provided more frequently for MP than for MT and AC, and more frequently for MT and AC than for DA (Schroyens, Schaeken, & d'Ydewalle, 2001). Correct AC reasoning remains hard even for educated adults (Cummins et al., 1991; Markovits and Doyon, 2004), and this reasoning outcomes depend on the difficulty of generating alternative answers to the given premises (Cummins, Lubart, Alksnis, & Rist, 1991; Markovits & Vachon, 1990). However, when increased AC reasoning scores are observed, these are often considered as the emergence of deductive reasoning in children.

In mathematics, proofs are based mainly on MP and MT forms (Stylianides, & Stylianides, 2008). According to the latter researchers, direct proof (using picture, algebra, or everyday language) and proof by mathematical induction is based on correct MP form whereas indirect proof (proofs by contraposition and by contradiction) is based on correct MT form.

1.2.CONDITIONAL REASONING: THEORY ACCOUNTS

Conditional reasoning in psychology has been approached by a plethora of theories. Thus, it becomes apparent that to address the complexity of this reasoning type, the study should be founded on the theory that describes most appropriately conditional reasoning at the specific range of ages. In this section the main theories are briefly and separately described. While doing so, the most appropriate theory for the current studies of this thesis will be discussed and justified based on our goals and our sample.

1.2.1. Mental logic theory

One of the theories that was discussed to describe conditional reasoning is Mental Logic Theory according to which making conclusions is based on formal mental rules (Schroyens, Schaeken, & d'Ydewalle, 2001). This theory primarily focus on the logical structure of the task hence the reasoner extracts the logical forms of the premise and generate conclusions by using mental rules (e.g. Braine & O'Brien 1991). This theory puts emphasis on the wording of the conditionals and words such as 'if, or, all, only, some etc.' Since this theory is founded on the importance of the logical structure of the conditionals, one could claim that this theory would not predict specific effects of knowledge about the context for different logical forms.

1.2.2. Probabilistic theories

According to this account of theories, reasoners cope with conditionals by reflecting the conditional probability of the consequent based on their subjective approach (e.g., Evans & Over, 2004; Oaksford, Chater, & Larkin, 2000). Probabilistic theories assume that deriving a conclusion is mainly a statistical process that is based on the use of an individual's knowledge and refers to an estimation of the probability that the given conclusion is true. For instance, correct AC reasoning according to this theory, depends on the number of 'p' and 'not p' cases (alternative antecedents) under the given premise that q is true.

1.2.3. Mental model theory

However, there is a theory that assumes that conditional reasoning performance is founded on the logical structure of the conditional, while accounting for content-based variation, this theory is the Mental Model Theory (MMT) (Johnson-Laird & Byrne, 1991, 2002). MMT is one of the most influential theories that have been used to describe conditional reasoning in young children. This theory suggests that conclusions are drawn by constructing mental models that encode information about the meaning of the conditional (Johnson-Laird & Byrne, 2002) and represent possibilities. Such models are generated from a semantic analysis of rules and represent possible states of affairs under these rules. If a given model represents a potential counterexample to a putative conclusion, this conclusion will be denied, otherwise it will be accepted. In the view of that, conditional reasoning depends on knowledge about the specific content of a conditional. In other words, individuals use the meaning of premises and their knowledge about the content to think about what is possible given the premises (Nickerson, 2015). In other words, the conditional reasoning process is semantic because the generation of the mental models depends on meaning of the conditional and the knowledge of the reasoner (Garnham & Oakhill, 1996).

The development of MMT theory in general provided many theoretical tools regarding the interpretation of different phenomena in the area of conditional reasoning (Markovits, 2000). Prior studies have shown that MMT accurately describes conditional reasoning processes among elementary students using everyday contents (e.g., Markovits, 2000; Markovits & Thompson, 2008). As Evans & Over (1996) mention, this theory is the most complete and thorough referring to the interpretation of deductive process, the interpretation of systematic errors, and the context effect on conditional reasoning scores. Thus, we will use MMT in the following as the basis for our analyses of conditional reasoning in the primary school age.

1.3 CONDITIONAL REASONING IN PRIMARY SCHOOL

1.3.1 Development of conditional reasoning

According the research in developmental psychology, there is generalized increase in children's skills to give the logically correct answers to various conditional reasoning tasks (Daniel & Klaczynski, 2006; Janveau-Brennan & Markovits, 1999). Based on MMT, we can expect the development of conditional reasoning to depend on at least two mechanisms: (1) the acquisition of general schemata, which allow adequate interpretation of the different logical forms and guide strategies of model generation, and (2) an increase in knowledge about different contents, which is necessary to build up mental models in general, and more specifically disablers and alternatives. While the first mechanism can be assumed to work successfully independent of the content of the conditionals, the second parameter allows for the construction or retrieval of mental models for specific conditionals, which is content dependent.

Based on working memory considerations, as Barrouillet and Lecas (1999) effectively support, the development of such general schemata of conditionals starts from the conjunctive-like interpretation (Model: ‘p and q’), developing to the biconditional interpretation (Models: ‘p and q;’ ‘not-p and not-q’) and then to the full conditional interpretation (Complete three-model representation: ‘p and q‘ ;’not-p and not-q‘ ;’not-p and q’). This developmental pattern occurs due to the increase of the reasoner’s working memory capacity.

According to Markovits (2014) the conditional reasoning performance is also affected by the semantic properties of the given premises. In detail, the latter researcher (Markovits, 2000) suggests that students even at the age of 7 or 8 show correct MP and AC reasoning when dealing with category-based conditionals (e.g., ‘If an animal is a dog, then it has four legs’). Subsequently, students at the age of 10-12 develop conditional reasoning skills with familiar causal conditionals (e.g., ‘‘If a rock is thrown at a window, then the window will break’’) (Janveau-Brennan & Markovits, 1999). Conditional reasoning skills in an imaginary context (contrary- to-fact) follow (e.g., ‘‘If a feather is thrown at a window, then the window will break’’) and are not observed until the age of 14–15 (Markovits & Vachon, 1989). Lastly, conditional reasoning skills based on abstract premises (‘‘If P then Q’’) are revealed only during early adulthood (Markovits & Lortie-Forgues, 2011; Markovits & Vachon, 1990; Venet & Markovits, 2001). This developmental pattern is presented in the table 2 which is proposed by Markovits (2014).

Table 2

Developmental patterns of conditional reasoning by premise category and alternatives generation (Markovits, 2014).

Age level	Premise category	Alternative	Example
7–8 years	If category P then property Q (categorical)	Category A that also has property Q	If an animal is a dog, then it has four legs. Cats also have four legs.
10–12 years	If cause P then effect Q (familiar causal)	Cause A can also lead to Q	If a rock is thrown at a window, then the window will break. Chairs can also break windows.
14–16 years	If cause P then effect Q (CF causal)	Cause A can also lead to Q (where A is not necessarily a cause of Q)	If a feather is thrown at a window, then the window will break. Dust can also break windows (if feathers can).
20+ years	If X then Y (abstract)	A could lead to Y (A could be anything)	If Dolx, then Blax. Something else could also lead to Blax.

1.3.2 Differences by logical form

According to MMT, to make a valid deduction on MP forms, an instance of the base model of the condition, ‘p and q’, is sufficient. To derive valid conclusion for MT tasks, an additional model ‘not-p and not-q’ is necessary. To conclude that AC does not allow a valid conclusion, minimally models ‘p and q’ and ‘not-p and q’ are necessary, while for DA models ‘not-p and not-q’ and ‘not-p and q’ are required. Critical to the ability to deny the AC and DA inferences is the ability to generate the model ‘not-p and q’. These refer to alternative antecedents, which are counterexamples for the typical errors in AC and DA inferences. Another class of alternative models are instances of ‘p and not-q’, which are counterexamples for the correct MP and MT inferences, called disablers (Cummins et al., 1991). Higher availability of disablers is related to lower rate of MP and MT acceptances (Cummins, 1995; Janveau-Brennan & Markovits, 1999), while higher availability of alternative antecedents is related to higher rates of correct reasoning in AC and DA forms (Cummins et al., 1991; Markovits & Vachon, 1990). Janveau-Brennan and Markovits (1999) found that conditional reasoning in young children (ages 7 to 12) is affected by rates of both alternative and disabler generation in a practice that follows their effect in adults. In similar lines, studies indicate that correct DA and AC reasoning correlates negatively with correct MP and MT reasoning (Newstead et.al., 2004; Morsanyi, McCormack & O' Mahony, 2018). This could mean an association between alternative generation (supporting correct DA and AC reasoning) and disabler generation (leading to incorrect MP and MT reasoning).

1.3.3 The role of context

First of all, according to Greatorex (2014) the problem context is difficult to be defined. In the mathematics education field, the terms 'cover history, thematic content, situation and setting' are used alternatively to the term problem context (Salgado, 2017). For this reason, the terms context and content in this thesis are used interchangeably.

Delving into the literature, it is apparent that there some context-independent mechanisms as well as some context-specific mechanisms. Starting from the first type of mechanisms, within MMT, the general ability to construct interpretations of conditionals that are more complex has been hypothesized to underlie development. Specifically, more complex interpretations require maintaining additional models in memory, which requires increased working memory capacity. As Barrouillet and Lecas (1999) suggested (see section 1.3.1) the development of conditional reasoning abilities is determined by a developmental increase in working memory capacity. While this model suggests relatively sharp developmental differences, other results show a more gradual change. This finding is in accordance with Janveau-Brennan and Markovits'(1999) study, who found a steady age-related development in the ability to show correct AC and DA reasoning between grades 1 and 6, as well as a gradual increase in retrieval of disabling conditions leading to less correct MP reasoning. Many studies have confirmed that the AC and DA forms are usually not mastered before the age of 11–12 years, while even only about one third of adults have been found to systematically make these inferences normatively (Gauffroy & Barrouillet, 2009; Moshman, 2011; Markovits, 2014; Ricco, 2010; Christophorides, Spanoudis, & Demetriou, 2016). Summarising, the current evidence indicates that, possibly connected with working memory capacity, children employ mechanisms of conditional reasoning, which allow correct MP reasoning first, then MT, and later AC and DA reasoning. However, as we shall see, there are clear indications that, due to the necessity to retrieve or generate alternatives and disablers, these reasoning skills are subject to important content effects. Regarding the development of conditional reasoning with positively vs. negatively worded minor premises (i.e. MP and AC vs. MT and DA), the literature provides less information. The negations involved in negatively worded minor premises have been hypothesized to pose specific difficulties (Schroyens, Schaeken, & d'Ydewalle, 2001). However, the MMT account for the effects of positive vs. negative wording in the literature is not as explicit as for definite vs. indefinite forms and is yet to be further explored.

Apart from general individual development, the specific content of the conditional has been found to influence the ability to make conditional inferences (e.g. Janveau-Brennan & Markovits, 1999). Previous studies (Markovits, 2000; Markovits & Thompson, 2008) have shown that even 6 or 7 year old children can reason logically even with the AC and DA inferences, when the content refers to simple categorical premises (e.g. 'If an animal is a cat, then it has legs'). Likewise, Chao and Cheng (2000) found evidence that conditional reasoning in preschool children is domain-specific, since content-rich rules emerge earlier than more formal ones. These kinds of effects have led to the conclusion that conditional reasoning might be domain-specific (e.g., Chao & Cheng, 2000; Cummins, 1996b), especially in early ages. In fact, Markovits (2014) outlined the developmental pattern described in section 1.3.1 which is founded on the content of conditionals (categorical premises, familiar causal premises, counterfactual premises, abstract conditionals).

The fact that conditional reasoning performance depends on the content of the conditionals is in line with critical assumptions of MMT. According to this, mental models need to be retrieved or constructed based on knowledge about the situation contained in the conditional statement. As previously mentioned, studies on the effects of content on conditional reasoning have concentrated on broad categories, which have been shown to affect in particular retrieval of alternative antecedents (e.g. Markovits & Lortie-Forgues, 2011). However, the effects of more specific forms of content variation, such as that involved in reasoning with mathematical concepts, need to be further highlighted.

1.3.3.1 Everyday context

As presented in table 2 proposed by Markovits (2014), regarding the everyday context there is a clear developmental trajectory starting from category-based conditionals, and subsequently transmitting to causal conditionals. In detail students even at the age of 6 can show correct AC and DA reasoning while approaching familiar category-based conditionals for which the generation of alternative models is accessible (e.g, Markovits & &Thompson, 2008). 10-12 years old student cope successfully with causal conditionals in the everyday context (Janveau-Brennan & Markovits, 1999). In general, in the everyday context according to De Neys and Everaerts (2008) and Janveau-Brennan and Markovits (1999) the changes is conditional reasoning performance with concrete premises up to middle adolescence can be explained by age-related improvements in working memory, retrieval efficiency, and inhibitory capacity.

Conditional reasoning based on metacognitive understanding (Daniel & Klaczynski, 2006), or conditional reasoning with abstract premises (Markovits & Vachon, 1990), are observed to be developed later.

1.3.3.2 Contrary-to-fact context

In section 1.3.1 was mentioned among other types of conditionals, the contrary-to-fact context (imaginary context). This category of conditionals refers to conditionals that are based on false (contrary to reality) premises. According to Markovits and Lortie-Forgues (2011), this type of conditional's context facilitates the reasoning development from the familiar context to the abstract one (e.g. *"If a shirt is rubbed with ink, then the shirt will be clean.")* This kind of conditionals require the generation of alternatives in the the contrary-to-fact context used (*i.e., "If a shirt is rubbed with oil, the shirt will be clean"*). Such contrary-to-fact alternatives refer to an intermediate form of alternatives construction (Markovits and Lortie-Forgues, 2011) since both young students and adolescents find it more difficult to perform correct AC and DA reasoning with contrary-to-fact premises than with true premises (Markovits & Vachon, 1989).

1.3.3.3 Mathematical context

One possible type of content, for which knowledge is acquired during primary school age, comprises mathematical concepts. Consider, for example, the conditional "*If a house has three floors with four windows, each, then it has twelve windows"*. Representing this situation per se does not require substantial mathematical knowledge beyond representing cardinal numbers, which is usually acquired by early primary school age (e.g., Litkowski et al. 2020). However, generating an explicit alternative would involve imagining a configuration of floors and windows per floor, which does not consist of twelve windows (for example, two floors with six windows each). Finding pairs of factors that have a given product requires substantial mathematical knowledge about multiplication, which has been found to develop slowly even until the end of primary school (e.g., Robinson, Price, & Demyen, 2018).

Only sparse evidence about conditional reasoning with mathematical concepts is available for primary school. The mathematics education literature has focused mostly on older learners, and has shown, for example, that dealing with mathematics and participating in mathematics instruction can lead to improved conditional reasoning skills by following the 'theory of formal discipline' (Attridge & Inglis, 2013; Handley et al., 2004; Inglis & Simpson, 2008, 2009; Morsanyi, Kahl, & Rooney, 2017; Morsanyi, McCormack, & O'Mahony, 2018). For secondary school students, research has focused mainly on conceptual issues such as the differentiation between a statement and its converse, and less on drawing conditional inferences (Küchemann & Hoyles, 2002).

Given that both general schemata for conditional reasoning and mathematical knowledge develop during primary school age, it is of substantial interest to understand how these two developments interact. It can be assumed that knowledge about familiar conditionals is widely available at this age, while knowledge about mathematical concepts varies substantially (Robinson, Price, & Demyen, 2018). This would suggest that reasoning with mathematical concepts would be more difficult than reasoning with everyday statements. This would be consistent with prior results showing early reasoning skills with everyday contents (e.g. Markovits & Thompson, 2008) and reports on secondary school students' problems dealing with mathematical conditionals (Küchemann & Hoyles, 2002). Different mechanisms could be hypothesized to explain such differences. First, the mathematical content of the conditional might affect decoding and representation of the conditional, which would be reflected in a relatively coherent performance difference between everyday reasoning and reasoning with mathematical content over all logical forms. This effect could also be moderated by the development of mathematical knowledge during primary school age, and thus be greater for younger students. Beyond the initial problem representation, the retrieval or construction of alternative mental models is a second point in the MMT account of conditional reasoning, that is particularly dependent on content-related knowledge. MMT would predict stronger content-related differences here for the two indefinite forms AC and DA, since more models are required to make valid inferences on these forms than for the definite forms MP and MT. Indeed, prior research has found an influence of alternative generation skills on AC reasoning rather than on MP reasoning (Janveau-Brennan & Markovits, 1999; Markovits & Quinn, 2002). Finally, since this effect depends on available knowledge about the conditional's content, it should be more pronounced for students in lower grades, leading to an interaction of content, logical form and grade level. Given the lack of available evidence on elementary school students' conditional reasoning skills with mathematical concepts, more substantial hypotheses are hard to derive. However, the role of alternative generation found in prior studies with elementary school students (e.g., Markovits, 2017), together with progress in mathematical knowledge during elementary school age (Robinson, Price, & Demyen, 2018) speak to expecting the interaction of content, logical form, and grade level.

1.3.4 Alternatives generation process

According to Mental Model Theories, inferences are drawn by constructing mental models that encode information about specific situations in which the conditional is valid (Johnson- Laird & Byrne, 2002). The reasoners' ability to generate alternative models (beyond a model representing "p and q") for the given conditional is considered a crucial prerequisite to draw valid inferences. Studies on conditional reasoning in the everyday context (De Chantal & Markovits, 2017) have shown that the alternative generation skills do predict early development of conditional reasoning. Hence the role of alternatives for conditional reasoning is well-studied in the everyday context, however this role has not been studied for conditionals that involve mathematical concepts (mathematical context; e.g., "*If I arrange three rows of four squares each, then I need 12 squares*."). In this case, alternatives generation concerns the mental construction of mathematical objects that fulfil 'not-p and q' (e. g., "*12 squares could be constructed by six rows of two squares each"*), beyond those that represent 'p and q' (or 'not-p and not-q'). Generating such alternative perspectives to mathematical situations is often discussed in research on multiple solutions (Leikin & Lev, 2007). Based on this perspective, alternatives generation can be assumed to require mathematical knowledge of the conditional content (Leikin & Lev, 2007). Beyond a general link between mathematics skills and conditional reasoning skills (Attridge & Inglis, 2013), this could lead to a specific influence of mathematical knowledge on AC and DA conditional reasoning in the mathematical context. Studies in the field of mathematics with university students show negative correlation between MT form and DA as well as AC form (Attridge & Inglis, 2013; Morsanyi, McCormack & O' Mahony, 2017). This backs up the assumption that conditional reasoning in this context is based on mental model construction, similar to the everyday context.

Overall, the existing literature in primary school pupils investigates the relation between alternatives generation and conditional reasoning only in the everyday context (with different levels of abstraction; e.g. Markovits & Lortie-Forgues, 2011). In the mathematical context, research with primary school students either investigate conditional reasoning (Christoforides, Spanoudis & Demetriou, 2016) or multiple solution tasks (e.g. Sullivan, Bourke & Scott, 1997). Yet, to date research not has addressed alternatives generation in relation to conditional reasoning in two different contexts and this study aims to fill this research gap.

According to MMT, generation of alternatives is based on knowledge about the content of the conditionals. Alternatives generation for a given conditional is considered as a crucial prerequisite to draw valid inferences (Johnson-Laird & Byrne, 2002; Markovits & Barrouillet, 2002). Studies on reasoning with conditionals from an everyday context (e.g. De Chantal & Markovits, 2017) have shown that alternatives generation skills predict conditional reasoning even from pre-school age on. In these studies, alternatives generation skills are associated with correct AC and DA reasoning, in particular (Cummins et.al, 1991; Markovits & Vachon, 1990). In addition to alternatives generation, individuals might also generate disablers (mental models of the form '*p and not-q*', contradicting the major rule) describing inhibitory factors which might prevent q from occurring, even in the presence of p (e.g., 'Anna took a painkiller, so her arm does not hurt, even though it is broken'; Cummins et al., 1991). Disablers might lead to the rejection of valid conclusions for MP and MT inferences (Janveau-Brennan & Markovits, 1999). Many studies report a positive correlation between the numbers of generated alternatives and disablers (Thompson, 2000; De Neys, Shaeken, & D'Ydewalle, 2002). In line with this, studies with university students revealed that correct DA and AC reasoning correlates negatively with correct MT reasoning (Newstead et.al, 2004). Hence, it is likely that alternatives generation is positively linked with AC and DA reasoning, being not or negatively related with MT reasoning. In prior research with young students, disabler generation is considered less relevant for logical reasoning than alternatives generation (Janveau-Brennan & Markovits, 1999; De Chantal & Markovits, 2017).

1.3.5 The link between conditional reasoning and creativity

Alternatives generation procedures, do not only underlie conditional reasoning but also many types of reasoning (Piaget, 1987; Gauffroy & Barrouillet, 2011) such as divergent thinking. In particular, according to Markovits and Brunet (2012) divergent thinking and conditional reasoning share some commonalities which are based on the necessity of alternatives generation under the given content and instructions. Divergent thinking which is a fundamental component of creativity requires this process of alternatives generation which is similar to the one needed in conditional reasoning. In other words, the key process in both thinking forms is the understanding of uncertainty. It is accepted that divergent thinking can be taught, even in young students (Cliatt, Shaw, & Sherwood, 1980), at least some components of it (i.e fluency and flexibility; Leikin, 2013), however the link between divergent thinking and conditional reasoning at the altar of investigating potential scaffold functions of divergent thinking priming has been only lately studied.

In detail, this link between the two concepts has been investigated by Markovits and De'Chantal (2019), and the results they obtained from preschool children (4 -5 years old) show that alternative generation priming affects positively AC reasoning. It is worth noting that AC reasoning as mentioned in the previous studies of this dissertation refers to the individual's ability to go beyond the given representation of the conditional and generate alternative conclusions that are not explicitly described in the relevant task. Another study with elementary students by Markovits and Brunet (2012) revealed that after a short training on divergent thinking, students' solution rates in logical reasoning tasks were improved. However, these results apply only to students of high socio-financial status in line with the authors' assumptions. The results of another study (Daniel & Klaczynski, 2006), show that providing students with an explicit alternative antecedent when presenting premises improves AC reasoning and the understanding of uncertainty.

Based on this evidence that divergent thinking and conditional reasoning are based on remarkably similar reasoning processes, and that a priming of divergent thinking improves conditional reasoning, we wanted to extend these results to the mathematical context. In the mathematics context multiple solution tasks are those that allow some space for alternative solutions to students, and this type of tasks is used in this study, by encouraging of which, we will test conditional reasoning scores in the post-testing.

1.4 TRAININGS ON ELEMENTARY STUDENTS' CONDITIONAL REASONING SKILLS

Based on different theories, prior research has shaped its trainings accordingly. For example, studies which were based on Mental Logic theory, designed their training on truth tables, emphasising the logical structures of conditional reasoning (e.g., Muller et al., 2001). Yet, the results were not that promising. Hence, when it comes to the understanding of logical relations as the main way to master conditional reasoning, it does not seem sufficient.

1.4.1 Alternatives generation priming

Based on MMT (Johnson-Laird's, 2006), the trainings are focused on the mental models that students have to think and handle appropriately. A study by DeChantal & Markovits (2017), showed that alternatives generation priming significantly improved conditional reasoning skill among preschool students, providing evidence for the importance of alternative generation skills even in the early development of logical reasoning

1.4.2 Training of meta logical aspects

Another experimental study aiming at the improvement of elementary students' conditional reasoning skills in AC and DA reasoning, was based on the training of meta logical principles (e.g. logical consistency, inferential identity of each scheme, etc) and the theory of Demetriou* (Christophorides, Spanoudis, & Demetriou, 2016). The results shown that through this type of training, MP and MT reasoning were fully developed by 3rd graders, however, the two invalid forms, AC and DA, were not mastered even by 6th graders.

1.4.3 Rationale of a training through mathematical multiple solution tasks

Looking at the alternative generation process of divergent thinking from the mathematics education view, this process can be evaluated and promoted through multiple solution mathematical tasks. The latter according to Silver (1997) refer to tasks that can be solved by many solution methods and potentially might encourage the development of students' fluency and flexibility. Another definition by Leikin and colleagues (Leikin & Levav-Waynberg, 2008) characterise as multiple solution tasks those assignments that require explicitly from students to solve them in many different ways. According also to Polya (1973) and Schoenfeld (1983) solving multiple solution tasks can facilitate students' knowledge by translating different representations, comparing different strategies, and make possible connections between different mathematical concepts and ideas.

Taken together findings from the psychology field concerning trainings on young students' conditional reasoning the following arise. First of all, results from psychology suggest that instruction can support young students develop some logical forms, as MP, MT (e.g. Christophorides, Spanoudis, & Demetriou, 2016) and AC forms (Markovits and De'Chantal, 2019; Daniel & Klaczynski, 2006) and students can master these forms even earlier than expected according to the developmental trajectory (e.g. Barroulliet & Lecas, 1999).

From the mathematics education domain under the lens of arguments and proofs construction, it seems that elementary students show deductive reasoning skills through the appropriate instructional context (e.g. Reid, 2002; Stylianides, 2007b; Zack, 1997). However, no evidence is found in the research field of mathematics education regarding any trainings on students' mathematical conditional reasoning skills in primary school. Consequently, the main question aim of this study is the investigation of the role of mathematical instruction through multiple solution tasks in nurturing the development of students' conditional reasoning skills.

2. OPEN QUESTIONS AND GOALS OF THE PhD PROJECT

Despite the increase in publications in psychology on the topic of conditional reasoning in primary school, there is almost no evidence regarding conditional reasoning and alternative generation skills within mathematical concepts by students at this group age. Therefore, this thesis aspires to open a discussion about this topic. This aim is addressed through the pilot study and are in line with the following research questions:

(1) How do students' conditional reasoning skills in mathematics and everyday contexts develop during elementary school age?

(2) In which way are students' conditional reasoning skills dependent on the required logical form of inference in each context?

(3) Do students' conditional skills develop similarly for the four logical forms of inference similarly in both contexts, or are there specific differences?

(4) Does the ability to generate examples of mathematical concepts, and to generate multiple alternative models for a given premise, have an influence on students' conditional reasoning with these concepts?

By answering the above questions, the study aims at the design of some mathematical conditional reasoning tasks parallel to those in the literature about everyday context and the test of their feasibility. Subsequently, this pilot study provides also some first evidence concerning elementary students' conditional reasoning skill in everyday and mathematical context based on differences in mathematical knowledge.

The main goal of study 1 is to investigate the extent to which reasoning about mathematical concepts specifically affects primary school students' conditional reasoning in the four different logical forms, and its development. Beyond replicating findings on the development in logical reasoning with familiar everyday statements, the following questions are addressed:

(5a) Is there a general disadvantage of reasoning with mathematical content, compared to everyday content, for primary age students?

(5b) Is such a general disadvantage greater for children in lower grades, as compared to upper primary school grades?

(6a) Is there a specific disadvantage of reasoning with mathematical concepts, as compared to everyday conditionals, for the indefinite logical forms AC and DA?

(6b) Is the effect under (6a) dependent on students' grade level?

Study 1 also aimed to transfer results on the role of alternatives generation in primary students' conditional reasoning from everyday conditionals to conditionals from a mathematical context. Hence, the following research questions are addressed:

(7) Is the influence of alternatives generation skills on conditional reasoning specific to the respective context (everyday vs. mathematical)?

(8) Do alternatives generation skills predict correct reasoning differently across the four logical forms?

By answering these questions, this study provides some evidence in the field of mathematics education concerning the effect of context, grade level, logical form, alternatives generation skills and working memory capacity on elementary students' conditional reasoning, since this area was not documented so far.

As mentioned, there is limited evidence concerning elementary students' conditional reasoning skills in mathematics (Datsogianni, Ufer, & Sodian, 2018; Datsogianni, Sodian, Markovits, & Ufer, in press), so our knowledge of how we could support students to develop these skills is based on even less data. Therefore, the aim of study 2 is to investigate the type of instruction that might lead to improvement of elementary students' conditional reasoning in mathematics and beyond. Hence the research questions of this study are the following:

(9) Do students' scores in conditional reasoning tasks improve after a short-term training with alternatives generation tasks which refer to the contrary-to-fact context (Experimental Condition A)?

(10) Do students' scores in conditional reasoning tasks improve after a short-term training with mathematical multiple solution tasks (Experimental Condition B)?

To answer these questions we conducted a short-term interventional study which was based on the alternatives generation priming within the contrary-to-fact (experimental condition A) and mathematical context (experimental condition B). Control conditions refer to solving of single solution mathematical tasks (condition C) and text reading (condition D). At the same time the results of Study 2 will contribute to future discussions regarding scaffolding elementary students' conditional reasoning, not only at a theoretical, but also at a practical level.

3. PILOT STUDY

3.1 PILOT STUDY: PART I- ELEMENTARY STUDENTS' CONDITIONAL REASONING SKILLS IN MATHEMATICAL AND EVERYDAY CONTEXTS

3.1.1. Brief Introduction

One of the main aspects of logical reasoning is conditional reasoning, i.e. reasoning with if-then statements. Research from developmental psychology has identified corresponding skills in young students and some clear developmental patterns, but also a great deal of variation in performance due to various external factors. Moreover, a link between conditional reasoning and mathematics has been found only in the case of late adolescence and adults (Attridge & Inglis, 2013; Inglis & Simpson, 2009). Our knowledge about elementary students' conditional reasoning within mathematics is still weak, despite claims in the literature about the importance of logical reasoning for mathematics learning and success (e.g., Morsanyi & Szücs, 2014; Nunes et al., 2007). Given that current theories describe conditional reasoning as a process that involves domain-specific knowledge, it is an open question to which extent students can transfer their existing logical reasoning skills to contexts that involve mathematical concepts. The main goal of this exploratory study is to investigate if it is feasible to survey elementary school students' skills to reason with conditionals that involve mathematical concepts.

In the sequel, we will focus on the MMT account of conditional reasoning. According to this, individuals initially represent conditionals by the conjunctive model that de- scribes a case in which p and q are valid, and check if this model is consistent with the minor premise. For MP, this is the case, and the conclusion q is true can be drawn. If the model is not consistent with the minor premise, further alternative models have to be constructed. Based on working-memory considerations, Barrouillet & Lecas (1999) propose that individuals' treatment of conditionals evolves from a conjunctive-like interpretation (only p and q), to a biconditional (p and q; not-p and not-q), and then a conditional interpretation (p and q; not-p and not-q; not-p and q). Changes from a conjunctive interpretation to a biconditional interpretation of implication statements would thus be reflected in increased solution rates for MT, further changes towards a conditional interpretation would be reflected in increased solution rates for DA and AC.

Research in developmental psychology has shown that even very young children possess basic abilities in at least some forms of conditional reasoning when tasks are presented in a familiar everyday context (e.g., Markovits & Thomson, 2008). However, there is a great deal of variation in performance due to external factors such as the type of instructions (Saelen, Markovits & Klein, 2009), the context in which the inferences are drawn and evaluated (familiar premises, false premises, or abstract conditional premises; Markovits & Lortie-Forgues, 2011), or how strongly the antecedent and the consequent are associated (Markovits et al., 1998).

The reasoners' ability to generate alternative models (beyond p and q) for the given conditional is considered a crucial prerequisite to draw valid inferences. This is particularly important to infer the uncertainty of AC and DA inferences, where two models are consistent with the minor premise that lead to different conclusions (Markovits & Lortie Forgues, 2011), but also for MT where the initial model is not sufficient to draw inferences. A range of studies illustrate that this generation of alternatives is not a generic process of recalling what might be possible given the premises, but involves recalling concrete instances from semantic memory of what is possible in the given context (Markovits, 2014). For example, a study of Janveau- Brennan and Markovits (1999) shows that performance in conditional reasoning tasks was related to performance on a task that required students to generate alternative causes for the conclusion of the conditional, besides the given prerequisite. Thus, reasoning with conditionals, that involve mathematical concepts, will plausibly vary with available knowledge about these concepts.

Problems about students' deductive reasoning in mathematics are often reported in the context of proof (e.g., Stylianlides & Stylianides, 2007), often focusing on the difference between an implication and its converse (Küchemann & Hoyles, 2002). How- ever, conditional reasoning is not restricted to proofs. For example, the application of simple mathematical conditionals during problem solving (e.g., “If I arrange three rows of four squares each, I need 12 squares.”) can involve conditional reasoning processes in all logical forms. We use the term conditional reasoning in mathematical contexts to describe reasoning about such conditionals that describe mathematical structures. MMTs of conditional reasoning propose that, even though similar processes might be involved and similar strategies (cf. Barrouillet & Lecas, 1999) may be applied for conditional reasoning in mathematical and everyday contexts, the availability of knowledge about the underlying concepts will modulate reasoning in mathematical contexts. In particular, it can be expected that increasing concept knowledge will support conditional reasoning skills when alternative models are necessary, i.e. for AC, DA, and MT tasks.

Logical and conditional reasoning are considered to be closely linked with mathematics learning (e.g., Morsanyi & Szücs, 2014), and it is plausible that conditional reasoning is involved when students discuss about mathematical concepts, their properties and relations in the classroom. However, in light of the tension between early conditional reasoning skills in everyday contexts (Markovits & Thomson, 2008) and frequent claims about students’ fallacies when dealing with conditionals in mathematics (e.g., Küchemann & Hoyles, 2002), it is an open question how conditional reasoning in the two contexts is related.

3.1.2 Goals and Questions

The goal of the current study was primarily to study if it is feasible to survey elementary students’ conditional reasoning skills in everyday and mathematical contexts. Moreover, we aimed to compare grade 2 and grade 4 students’ conditional reasoning skills in everyday and mathematical contexts because of differences in mathematical knowledge. The study serves as a feasibility study for more detailed investigations in the future and, in particular, focused the following questions:

(1) How do students' conditional reasoning skills in mathematics and everyday contexts develop during elementary school age?

(2) In which way are students' conditional reasoning skills dependent on the required logical form of inference in each context?

(3) Do students' conditional skills develop similarly for the four logical forms of inference similarly in both contexts, or are there specific differences?

3.1.3 Methods of the study

A total of 55 elementary students (4th graders n=13: M=9.5 years, 6th graders n=42: M=11.5 years) from a public school of Cyprus participated in a cross-sectional survey study. A questionnaire was constructed with four conditional reasoning tasks within one common cover story. Two of the tasks focused on plausible conditionals in everyday situations, such as having fever, and two of the tasks focused on conditionals, which described multiplicative or additive structures in the context of the cover story. For example, one task in the mathematical context concerned dwarf houses. It was explained that these always consist of several rows, all with the same number of quadratic rooms of the same size. The major rule in this task was "If a dwarf's house has exactly 2 rows of 4 rooms each, then it has 12 windows". All four tasks contained one item for each logical form (MP, MT, DA, AC). The verbal structure of the tasks, items, and answer alternatives closely followed existing studies on conditional reasoning (e.g., Markovits & Lortie-Forgues, 2011). Tasks and items were arranged in a fixed, random order. The first author administered questionnaire in 40 minute class- room sessions. While the administrator presented the items to the class using a computer presentation, the students had booklets to provide their answers.

For each item, students were presented with the major and the minor premise for the respective logical form as well as one possible conclusion. They were asked to indicate if it can be inferred from the provided premises that the conclusion is true for sure, if it can be inferred that the conclusion is false for sure, or if it cannot be inferred for sure, if the conclusion is true or false from the available information. No reasons for the answers were requested to gather students' intuitive responses.

Students' answers for each item were coded as correct or false according to the interpretation of conditionals described above. The data was analysed using Generalized Linear Mixed Models (GLMM), a generalization of logistic regression that allows analyzing the data on item level but still taking into account dependencies between answers provided by each student. Apart from the questionnaire study, the items were also used in one-to-one-interviews with a 16 second and 8 fourth graders.

3.1.4 Results

Regarding the feasibility of the measurement, the group test sessions as well as the interviews showed that also children in grade 2 understood the tasks from both contexts. Out of 880 possible answers in the questionnaire survey (55 participants x 4 tasks x 4 items), participants provided 870 answers. Overall, 69.9% of the answers were correct, illustrating early conditional reasoning skills under specific conditions. The most frequent wrong answers were that no conclusion can be taken for MP (14.7% of all MP items) and MT (20.3%), that the proposed conclusion is correct for AC (33.3%), or that it is incorrect for DA (26.4%).

Regarding question (1), we expected higher solution rates for the everyday context than in the mathematical context, because it would be harder to generate mental models for the mathematical situations. Moreover, we expected this difference to be more pronounced in grade 4 than in grade 6 because we that students in grade 6 would have acquired more connected knowledge of addition and multiplication, allowing them to generate alternative mental models more easily. A GLMM analysis with in- dependent variables grade level and context showed that students provided more cor- rect answers in the everyday context than in the mathematical context in grade 4, while performance was similar in grade 6 (fig. 1a).

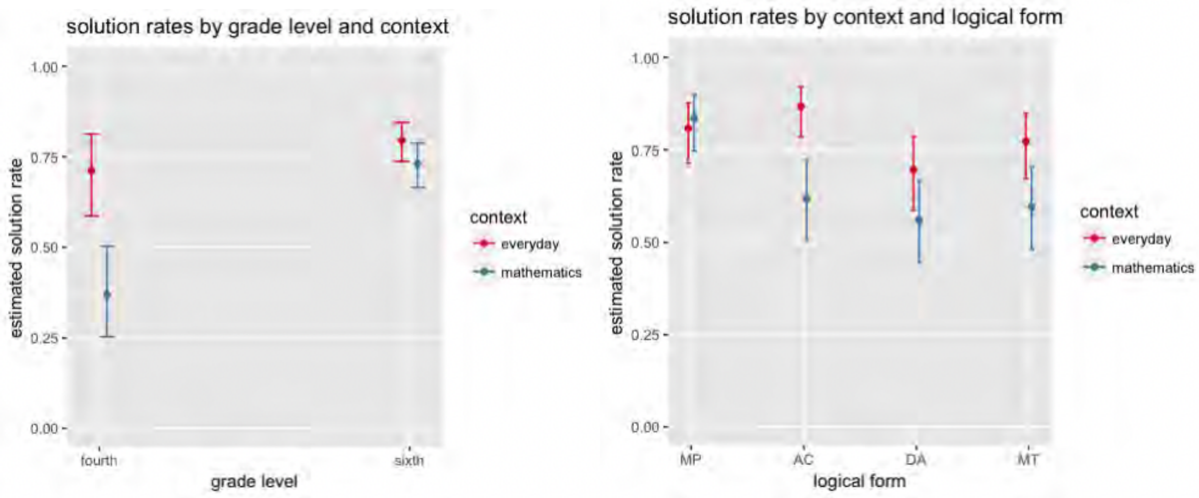


Fig. 1a and b: estimated solution rates and 95% confidence intervals by grade level and context (1a, left) resp. by context and logical form (1b, right)

Regarding question (2), we expected only small differences between the four logical forms in the everyday context, because early reasoning skills have been observed in the past for plausible everyday conditionals. However, we expected that in particular those forms of reasoning would be harder in the mathematical context, which involve construction of alternative models (MT, AC, DA). A GLMM analysis with independent variables inference type and context showed no significant differences between the solution rates for the four logical forms in the everyday context (fig. 1b). In the mathematical context, however, AC, DA and MT each turned out to be harder than MP (fig. 1b). The largest difference between the two contexts was found for AC, which requires generating a mental model that satisfies the consequent and does not satisfy the antecedent, the last stage of the Barrouillet & Lecas (1999) model.

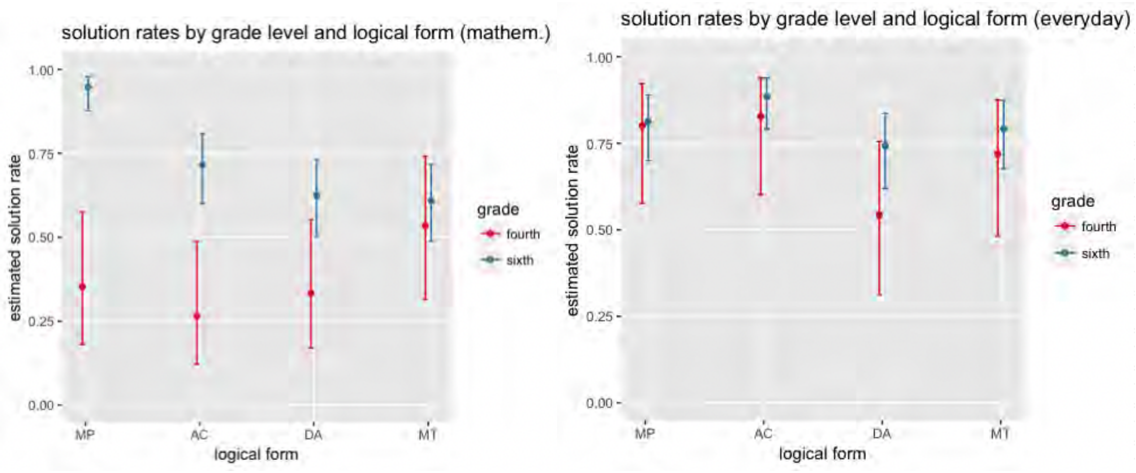


Fig. 2a and b: estimated solution rates and 95% confidence intervals by grade level and logical form (2a, left: everyday context; 2b, right: mathematical context)

We calculated GLMMs for each context separately, with independent variables grade level and logical form. For the everyday context, no significant differences between grade levels or between logical forms occurred (fig. 2a). For the mathematical context, MP and AC items were solved significantly better by sixth graders than by fourth graders (fig. 2b). This indicates that, although a certain amount of students seems to be able to interpret implications as strictly conditional statements, a transfer to mathematical contexts did not seem to occur necessarily, unless sufficient knowledge of the involved concepts is attained. However, MT tasks, which require to build up only one additional model (not-p and not-q) seemed to be unaffected by these differences.

3.1.5 Discussion

The first goal of this study was to investigate if it is feasible to survey elementary school students' conditional reasoning skills in both, everyday and mathematical contexts. Results from

interviews and one-to-one interviews indicate that the developed instrument is accessible to elementary school students. Moreover, the questionnaire survey results replicate early conditional reasoning skills reported in studies from developmental psychology (Markovits & Thompson, 2008). The second motivation for our study was to explore to what extend students could apply these skills when dealing with mathematical concepts in the classroom and beyond.
We assumed that grade 6 students' knowledge of addition and multiplication – the concepts underlying the tasks in the mathematical context – would be substantially higher than that of grade 4 students. Even though we could not report data on students' mathematics skills in this short report, the results are consistent with this assumption and MMTs: Only for the mathematical context did grade 6 students outperform their peers from grade 4, but results were comparable for the everyday context (1). This underpins the role of specific knowledge for conditional reasoning in mathematical contexts.

Moreover, our exploratory results indicate that the differences between the two contexts are not only due to task understanding: While the four logical forms were of similar difficulty for the everyday context, the logical forms requiring to analyse alternative models (AC, AD, MT) were significantly harder than MP in the mathematical context (2).

Finally, results indicate that the 6th graders deeper knowledge of mathematical concepts does not show equally for all logical forms, but primarily for MP and AC items (3). This is rather unexpected, since – following the model of Barrouillet & Lecas (1999) – an increase in AC and DA should occur only after a conditional interpretation of the conditional (and thus also MT) has been mastered. At the current state of re- search, reasons can only be hypothesized. For example, DA and MT involve negations in the problem formulation, but MP and AC not. However, if negations themselves were the problem, similar effects would be observable also in the everyday context.

Further research will be necessary to describe these effects in more detail, in particular focusing on the role of mathematical knowledge.

Even though our study shows that it is possible to measure elementary students' conditional reasoning skills in both contexts, and obtain plausible results in many aspects, the further results of our study have to be considered carefully. Most importantly, the small sample size – even though tackled statistically by analyzing on item level – forbids taking far-reaching conclusions apart from backing up the feasibility of the measurement and rising hypotheses for further research. Moreover, future research will have to collect really longitudinal data and include explicit measures of mathematical knowledge to support the interpretations.

However, the results obtained with the new instrument are consistent with theoretical expectations in some aspects that indicate educational implications – even though further research must back up these implications with further empirical data. Most centrally, the results illustrate that students are not necessarily able to transfer conditional reasoning skills from everyday contexts (Markovits & Thomson, 2008) to contexts involving mathematical concepts. Most likely the main reason for this is that knowledge of these concepts is necessary to analyze the necessary mental models when drawing inferences. If we agree that valid reasoning with mathematical concepts is a goal of classroom instruction, this implies the necessity to practice simple deductions on all logical forms when dealing with mathematical concepts. Moreover, it will be necessary to discuss why certain conclusions can or cannot be drawn – particularly for the more complex logical forms such as AC, DA, and MT (cf. Schroyens, Schaeken, & d'Ydewalle, 2001) – with a focus on the models underlying these inferences.

3.2 PILOT STUDY: PART II- A PILOT STUDY ON ELEMENTARY PUPILS' CONSITIONAL REASONING SKILLS AND ALTERNATIVE GENERATION SKILLS IN MATHEMATICS

3.2.1 Brief Introduction

One of the main aspects of logical reasoning is conditional reasoning, i.e. reasoning with if-then statements. Conditional reasoning tasks usually present a rule of the form "if p then q" as a major premise, and a minor premise. Four minor premises differentiate four possible logical forms of inference: p is true (Modus Ponens, MP), p is false (Denial of the Antecedent, DA), q is true (Affirmation of the Consequent, AC), and q is false (Modus Tollens, MT). Definite conclusions can be drawn for MP ("q is true") and MT ("p is false"), while AC resp. DA do not allow definite conclusions. Even very young children can show conditional reasoning skills in familiar everyday contexts (Markovits & Thomson, 2008). However, a link between conditional reasoning and mathematics skills has been found only in adolescents and adults in the context of proof (Stylianides & Stylianides, 2007).

According to Mental Model Theories, inferences are drawn by constructing mental models that encode information about specific situations in which the conditional is valid (Johnson- Laird & Byrne, 2002). The reasoners' ability to generate alternative models (beyond a model representing "p and q") for the given conditional is considered a crucial prerequisite to draw valid inferences. Studies on conditional reasoning in the everyday context (De Chantal & Markovits, 2017) have shown that the alternative generation skills do predict early development of conditional reasoning. This raises the question (4), if the ability to generate examples of mathematical concepts, and to generate multiple alternative models for a given premise, has an influence on students' conditional reasoning with these concepts.

3.2.2 Methods

Our study aimed to examine the relation between primary students' skills in conditional reasoning with mathematical concepts, and their performance on the corresponding, alternative generation tasks. Participants were 55 elementary students (4th graders n=13: M=9.5 years, 6th graders n=42: M=11.5 years) from a public school in Cyprus. Their conditional reasoning skills were assessed in two conditional reasoning tasks dealing with mathematical concepts (Datsogianni, Ufer, & Sodian, 2018). Each task contained one item for each logical form (MP, MT, DA, AC). For example, one conditional reasoning task focused on circumference of rectangles and was formulated in a context concerning dwarf houses. It was explained that dwarf houses always consist of several aligned rows, all with the same number of same-sized quadratic rooms. The major rule in this task was "If a dwarfs house has exactly 2 rows of 4 rooms each, then it has 12 windows". One alternative generation task for each of the mathematical concept asked students to generate as many examples satisfying the conclusion of the major premise (i.e. "Draw as many dwarf houses as possible, that have 20 windows!").

3.2.3 Results and discussion

Students' solved 62.7% of the mathematics conditional reasoning items correctly, illustrating early conditional reasoning skills under specific conditions. Regarding the first alternative generation task 62% of students gave 4 to 6 correct alternative solutions (M= 3.42, Mdn= 4.00). The second alternative generation task seemed to be more difficult since only 20% of students gave 4 to 5 correct solutions (M= 1.81, Mdn= 1.00). This indicates that the alternative generation tasks are feasible, but not too easy for our target sample. Supporting our assumption, the number of generated alternatives correlated with students reasoning performance for each of the tasks focusing on mathematical concepts (task 1: ρ=0.358,Sig.=0.001, task 2: ρ=0.343, Sig.=0.003), supporting the idea that mathematical reasoning is based on the generation of alternatives.

The results showed firstly that the applied instrument is accessible to students while they replicate early conditional reasoning skills reported in studies from developmental psychology (e.g. Markovits & Thompson, 2008). Secondly, this study replicates previous results (De Chantal & Markovits, 2017) regarding the significant relation between alternative generation skills and students' conditional reasoning skills, extending these results to tasks that involve mathematical concepts. The small sample size forbids taking far-reaching conclusions. However, future research may benefit from the developed tasks to describe the role of mathematical knowledge in conditional reasoning with mathematical concepts in more detail. For example, alternative generation training might be one approach to support students' conditional reasoning in mathematics.

4. STUDY 1

4.1 STUDY 1: PART 1- REASONING WITH CONDITIONALS ABOUT EVERYDAY AND MATHEMATICAL CONCEPTS IN PRIMARY SCHOOL

4.1.1 Brief Introduction

The ability to make valid deductions is considered of central importance for scientific reasoning, hypotheses generation and evaluation (Kuhn et al., 1988) as well as for mathematical thinking (Moshman, 1990; Markovits & Lortie-Forgues, 2011). In particular in mathematics, deduction is strongly associated with the development of students' ability to understand formal proofs (Foltz, Overton, & Ricco, 1995), and it is considered as the central mode of reasoning in mathematical theories. Even basic mathematical concepts are characterized by a strict logical structure, which makes deduction a central mode of reasoning in mathematics at every educational level. One of the key components of deductive reasoning is conditional reasoning, which occurs not only in scientific and mathematical discourse, but also in everyday language and communication. Deductive reasoning, and more specifically conditional reasoning, is assumed to lie at the heart of mathematical thinking. When it comes to conditional reasoning with mathematical concepts, we can assume that knowledge of these concepts develops substantially during primary school. Thus, at this age level, we see a parallel development of knowledge about mathematical concepts, as well as a general improvement in conditional reasoning skills. Learning to deal with mathematical concepts involves understanding and applying their properties correctly, which often have the form of conditional statements (e.g., 'Assume I have added two numbers. If I increase one of the two numbers by one and I decrease the other by one, then the two new numbers will have the same sum.'). It is an open question, to which extent making inferences with such statements does not only require conditional reasoning skills, but also mathematical knowledge about the concepts involved. Conditional reasoning in mathematics has been investigated primarily with adults (Attridge & Inglis, 2013; Inglis & Simpson, 2009) and secondary school students, with results suggesting difficulties in reasoning with conditionals in mathematics (e.g., Küchemann, & Hoyles, 2002). In spite of the potential importance of this form of reasoning, there is very little evidence about elementary students' abilities to reason with conditionals with a mathematical content. Although previous results have shown that secondary students' have real difficulties with mathematical conditional reasoning (e.g, Küchemann & Hoyles, 2002), psychological studies show that even very young children can reason correctly with conditionals with familiar everyday contents (e.g., Markovits & Thompson, 2008). This leads to the question if and how conditional reasoning with mathematical concepts is different from conditional reasoning with everyday contents.

The present study focuses on primary school students from grades 2, 4, and 6, an age in which sufficient mathematical concepts are learned, which makes it possible to study conditional inferences about these concepts. We will specifically compare elementary students' conditional reasoning skills about everyday situations and about mathematical concepts. At this age conditional reasoning skills with everyday contexts have been found in the past (e.g., Markovits & Thompson, 2008; Markovits, 2017), but it is an open question, to which extend such skills can be transferred to reasoning about mathematical concepts, which are acquired in this age.

Conditional inferences: Conditional inferences are based on a conditional rule of the form *'if p, then q'* as a major premise, and a minor premise. Four different minor premises differentiate four possible logical forms. These four forms can be described systematically by the wording of the minor premise (positive vs. negative) and the type of normative correct conclusion (table 3): Based on the traditional interpretation of conditionals that p is sufficient, but not necessary for q (Evans & Over, 2004), and depending on the logical form (minor premise), different conclusions can be drawn. Definite conclusions are possible for Modus Ponens ('p is true') and Modus Tollens ('q is false'), while Acceptance of Consequent ('q is true') and Denial of Antecedent ('p is false) do not allow definite conclusions about p and q, respectively, and the correct conclusion in both cases is indefinite.

Table 3

Logical forms in conditional reasoning for the major premise 'if p, then q'.

Name of form (abbreviation)	*Minor premise*	*Normatively correct conclusion*	*Minor premise wording*	*Conclusion type*
Modus Ponens(MP)	*'p is true'*	*'so q is true'*	positive	Definite
Modus Tollens (MT)	*'q is false'*	*'so p is false'*	negative	Definite
Denial of the Antecedent (DA)	*'p is false'*	*'so q or not q'*	positive	Indefinite
Affirmation of the Consequent (AC)	*'q is true'*	*'so p or not p'*	negative	Indefinite

Mental model theory of conditional inference: This theory suggests that conclusions are drawn by constructing mental models that encode information about the meaning of the conditional (Johnson-Laird & Byrne, 2002). Such models are generated from a semantic analysis of rules and represent possible states of affairs under these rules. If a given model represents a potential counterexample to a putative conclusion, this conclusion will be denied, otherwise it will be accepted. Within this perspective, conditional reasoning depends on knowledge about the specific content of a conditional. In other words, individuals use the meaning of premises and their knowledge about the content to think about what is possible given the premises (Nickerson, 2015). Prior studies have shown that MMT accurately describes conditional reasoning processes among elementary students using everyday contents (e.g., Markovits, 2000; Markovits & Thompson, 2008). Thus, we will use MMT in the following as the basis for our analyses of conditional reasoning in the primary school age.

According to MMT, to make a valid deduction on MP tasks, an instance of the base model of the condition, ‘p and q’, is sufficient. To arrive at a valid conclusion for MT tasks, an additional model ‘not-p and not-q’ is necessary. To derive that AC does not allow a valid conclusion, minimally models ‘p and q’ and ‘not-p and q’ are necessary, while for DA models ‘not-p and not-q’ and ‘not-p and q’ are required. Critical to the ability to deny the AC and DA inferences is the ability to generate the model ‘not-p and q’. These refer to alternative antecedents, which are counterexamples for the typical errors in AC and DA inferences. Another class of alternative models are instances of ‘p and not-q’, which are counterexamples for the correct MP and MT inferences, called disablers (Cummins et al., 1991). Higher availability of disablers is related to lower rate of MP and MT acceptances (Cummins, 1995; Janveau-Brennan & Markovits, 1999), while higher availability of alternative antecedents is related to higher rates of correct reasoning in AC and DA forms (Cummins et al., 1991; Markovits & Vachon, 1990). Janveau-Brennan and Markovits (1999) found that conditional reasoning in young children (ages 7 to 12) is affected by rates of both alternative and disabler generation in a way that is similar to their effect in adults. In addition, studies indicate that correct DA and AC reasoning correlates negatively with correct MP and MT reasoning (Newstead et.al., 2004; Morsanyi, McCormack & O' Mahony, 2018). This could reflect a connection between alternative generation (supporting correct DA and AC reasoning) and disabler generation (leading to incorrect MP and MT reasoning).

Development of conditional reasoning: Based on MMT, we can expect the development of conditional reasoning to depend on at least two mechanisms: (1) the acquisition of general schemata, which allow adequate interpretation of the different logical forms and guide strategies of model generation, and (2) an increase in knowledge about different contents, which is necessary to build up mental models in general, and more specifically disablers and alternatives. While the first mechanism can be assumed to have effects independent of the content of the conditionals, the second mechanism allows for the construction or retrieval of mental models for specific conditionals, which is content dependent.

Content-independent mechanisms: Within MMT, the general ability to construct interpretations of conditionals that are more complex has been hypothesized to underlie development. Specifically, more complex interpretations require maintaining additional models in memory, which requires increased working memory capacity. According to Barrouillet and Lecas (1999) model (see section 1.3.1) the development of conditional reasoning abilities is determined by a developmental increase in working memory capacity. While this model suggests relatively sharp developmental differences, other results show a more gradual change. For example, Janveau-Brennan and Markovits (1999) found a steady age-related development in the ability to make correct AC and DA inferences between grades 1 and 6, as well as a gradual increase in retrieval of disabling conditions leading to less correct MP reasoning. Many studies have found that the AC and DA forms are usually not mastered before the age of 11–12 years, while even only about one third of adults have been found to systematically make these inferences normatively (Gauffroy & Barrouillet, 2009; Moshman, 2011; Markovits, 2014; Ricco, 2010; Christophorides, Spanoudis, & Demetriou, 2016). In addition, many studies (e.g. Janveau-Brennan & Markovits, 1999) have found that specific content strongly affects conditional reasoning. Summarizing, the current evidence indicates that, possibly connected with working memory capacity, children acquire schemata of conditional reasoning, which allow correct MP reasoning first, then MT, and later AC and DA reasoning. However, as we shall see, there are clear indications that, due to the necessity to retrieve or generate alternatives and disablers, these reasoning skills are subject to important content effects. Regarding the development of conditional reasoning with positively vs. negatively worded minor premises (i.e. MP and AC vs. MT and DA), the literature provides less information. The negations involved in negatively worded minor premises have been hypothesized to pose specific difficulties (Schroyens, Schaeken, & d'Ydewalle, 2001). Mental Model Theories usually assume, that mental models only represent *possibilities* which can occur, given the premises – not what is impossible given the premises (principle of truth, Johnson-Laird, 2001), which could lead to problems if a negation leads to an unspecified situation (e.g., while 'not wet' has a similar meaning as 'dry', 'blue' has not clear opposite not 'not blue'). However, the account for the effects of positive vs. negative wording in the literature is not as explicit as for definite vs. indefinite forms.

Content-specific mechanisms: As for (2), apart from general individual development, the specific content of the conditional has been found to influence the ability to make conditional inferences. Previous studies (Markovits, 2000; Markovits & Thompson, 2008) have shown that even 6 or 7 year old children can reason logically even with the AC and DA inferences, when the content refers to simple categorical premises (e.g. 'If an animal is a cat, then it has legs'). Chao and Cheng (2000) also found evidence that conditional reasoning in preschool children is domain-specific, since content-rich rules emerge earlier than more formal rules. These kinds of effects have led to the conclusion that conditional reasoning might be domain-specific (e.g., Chao & Cheng, 2000; Cummins, 1996b), especially in early ages. In fact, Markovits (2014) reported a developmental pattern of conditional reasoning that suggests that 7 and 8 year old students possess conditional reasoning skills with categorical premises, 10 to 12 year old children can make logical conditional inferences with familiar causal premises, 14 to 16 year olds can do so with causal and counterfactual premises, while adults older than 20 also perform well with abstract conditionals. That conditional reasoning performance depends on the content of the conditionals is in line with core assumptions of MMT. According to this, mental models need to be retrieved or constructed based on knowledge about the situation contained in the conditional statement. As previously mentioned, studies on the effects of content on conditional reasoning have concentrated on broad categories, which have been shown to affect in particular retrieval of alternative antecedents (e.g. Markovits & Lortie-Forgues, 2011). However, the effects of more specific forms of content variation, such as that involved in reasoning with mathematical concepts, have not often been studied.

Conditional reasoning about mathematical concepts: One possible type of content, for which knowledge is acquired during primary school age, comprises mathematical concepts. Consider, for example, the conditional ‘If a house has three floors with four windows, each, then it has twelve windows’. Representing this situation per se does not require substantial mathematical knowledge beyond representing cardinal numbers, which is usually acquired by early primary school age (e.g., Litkowski et al. 2020). However, generating an explicit alternative would involve imagining a configuration of floors and windows per floor, which does not consist of twelve windows (for example, two floors with six windows each). Finding pairs of factors that have a given product requires substantial mathematical knowledge about multiplication, which has been found to develop slowly even until the end of primary school (e.g., Robinson, Price, & Demyen, 2018).

Only sparse evidence about conditional reasoning with mathematical concepts is available for primary school. The mathematics education literature has focused mostly on older learners, and has shown, for example, that dealing with mathematics and participating in mathematics instruction can lead to improved conditional reasoning skills by following the 'theory of formal discipline' (Attridge & Inglis, 2013; Handley et al., 2004; Inglis & Simpson, 2008, 2009; Morsanyi, Kahl, & Rooney, 2017; Morsanyi, McCormack, & O'Mahony, 2018). For secondary school students, research has focused mainly on conceptual issues such as the differentiation between a statement and its converse, and less on drawing conditional inferences (Küchemann & Hoyles, 2002).

Given that both general schemata for conditional reasoning and mathematical knowledge develop during primary school age, it is of substantial interest to understand how these two developments interact. It can be assumed that knowledge about familiar conditionals is widely available at this age, while knowledge about mathematical concepts varies substantially (Robinson, Price, & Demyen, 2018). This would suggest that reasoning with mathematical concepts would be more difficult than reasoning with everyday statements. This would be consistent with prior results showing early reasoning skills with everyday contents (e.g. Markovits & Thompson, 2008) and reports on secondary school students' problems dealing with mathematical conditionals (Küchemann & Hoyles, 2002). Different mechanisms could be hypothesized to explain such differences. First, the mathematical content of the conditional might affect decoding and representation of the conditional, which would be reflected in a relatively coherent performance difference between everyday reasoning and reasoning with mathematical content over all logical forms. This effect could also be moderated by the development of mathematical knowledge during primary school age, and thus be greater for younger students. Beyond the initial problem representation, the retrieval or construction of alternative mental models is a second point in the MMT account of conditional reasoning that is particularly dependent on content-related knowledge.

MMT would predict stronger content-related differences here for the two indefinite forms AC and DA, since more models are required to make valid inferences on these forms than for the definite forms MP and MT. Indeed, prior research has found an influence of alternative generation skills on AC reasoning rather than on MP reasoning (Janveau-Brennan & Markovits, 1999; Markovits & Quinn, 2002). Finally, since this effect depends on available knowledge about the conditional's content, it should be more pronounced for students in lower grades, leading to an interaction of content, logical form and grade level. Given the lack of available evidence on elementary school students' conditional reasoning skills with mathematical concepts, more substantial hypotheses are hard to derive. However, the role of alternative generation found in prior studies with elementary school students (e.g., Markovits, 2017), together with progress in mathematical knowledge during elementary school age (Robinson, Price, & Demyen, 2018) speak to expecting the interaction of content, logical form, and grade level.

The current study: In this study, we contrast conditional reasoning about premises involving mathematical concepts with reasoning about familiar causal premises. To this end, we study conditionals about easily accessible situations that contain mathematical structures. We chose structures related to multiplication and addition, concepts which are been introduced in first years of primary school. We assume that increasing knowledge about these concepts will affect conditional reasoning performance on top of the well-described development of conditional reasoning skills with familiar premises.

4.1.2 Goals and questions

The main goal of this study is to investigate to what extent reasoning about mathematical concepts specifically affects primary school students' conditional reasoning in the four different logical forms, and its development. Beyond replicating findings on the development in logical reasoning with familiar everyday statements, the following questions are addressed:

(5a) Is there a general disadvantage of reasoning with mathematical content, compared to everyday content, for primary age students?

Understanding the situations, in which we embedded the mathematical content for the conditionals, did not require substantial mathematical knowledge. Thus, we did not put forward a specific hypothesis for this main effect.

(5b) Is such a general disadvantage greater for children in lower grades, as compared to upper primary school grades?

Again, since we embedded the mathematical concepts and structures in easily accessible situations, we also did not put forward a specific hypothesis for this specific interaction.

(6a) Is there a specific disadvantage of reasoning with mathematical concepts, as compared to everyday conditionals, for the indefinite logical forms AC and DA?

The role of alternative generation for reasoning with indefinite forms has been shown for everyday reasoning. Moreover, generating alternatives for the mathematical structures reflected in our conditionals requires well-connected mathematical knowledge. Thus, we expected a corresponding interaction of content and the type of inference (definite for MP and MT vs. indefinite for AC and DA).

(6b) Is the effect under (6a) dependent on students' grade level?

Since knowledge about additive and multiplicative structures develops over primary school age, we expected that the effect discussed under (2a) would be more pronounced for students in lower grades, as compared upper primary school grades.

4.1.3 Methods

Sample and design: Around 300 students and their parents were approached for participation in in this cross-sectional survey. A total of 102 elementary students (average age 10 years, 1 month) from grade 2, grade 4 and grade 6 living in Cyprus participated. Parents' written consent and children's oral assent were obtained for all participants. Further information about the sample is displayed in table 4. On the 'books at home' question (Paulus, 2009), the median category was 'one complete bookcase (26 to 100 books)' in all grades, and distribution over the five answer alternatives (from 'no or very few books' to 'over 200 books') did not differ significantly over grades ($\chi^2(8) = 13.4$, $p = .10$).

Table 4

Sample size, mean, and standard deviation (SD) of age in months and working memory scores by grade level.

grade	*N*	*age in months M (SD)*	*working memory score M (SD)*
2	31	98.6 (8.6)	2.16 (0.99)
4	33	119.3 (6.4)	3.18 (1.04)
6	37	141.8 (6.4)	4.19 (1.66)

Each participant took part in one individual 45-60 minute face-to-face interview during regular school hours with the first author in a separate room of the school. The factors relating to *content* (everyday vs. mathematical) and logical form (*positive vs. negative* wording of minor premise, *definite* vs. *indefinite* conclusion, and the interaction of the two) were varied within subject, with randomized sequence of two content blocks, randomized sequence of conditionals within each content block, and randomized sequence of four minor premises (each relating to a different logical form) within each conditional.

Procedure: Initially, students were asked for their age, native language, and approximate number of books at their house ('books at home' question; Paulus, 2009). General instructions followed and participants were clearly informed about the anonymity and confidentiality of their replies as well as their voluntary participation, clarifying that they were free to withdraw from the interview process anytime without any negative consequences. Then, participants were familiarized with the three answer alternatives ('yes, this is certainly so', 'no, this is certainly not so', 'you cannot say for sure, whether it is so, or not') and shown how to select their answers on the tablet computer screen. A short game with three questions about hidden marbles was used to check for comprehension of these answer alternatives (see Appendix). An explanation was given in case of wrong answers. Afterwards, two blocks (everyday vs. mathematical) of four conditionals each were presented (10 - 15 min. per block). Each conditional was presented separately, and students were asked to make four conclusions based on four different minor premises, corresponding to the four logical forms. Then, a block of alternative generation tasks followed, which is not examined in this paper. In the end of the interview, students' responses to a short working memory test and an arithmetic calculation test (not examined in this paper) were gathered.

All tasks were displayed using a tablet-based interview system and children were expected to select their preferred answer by touching the screen on the respective part of the visual representation of answer alternatives. The interview system also randomized the sequence of blocks, conditionals and minor premises. By ensuring this full randomization of the questions' order, we systematically controlled for possible order effects. No justifications for the answers were requested as we were interested in students' intuitive responses.

Conditional reasoning tasks: Eight conditionals were used to measure conditional reasoning skills (four conditionals per content condition). The verbal structure of the tasks was parallel in both content blocks. All conditionals were presented verbally and in a written form on a tablet computer. Participants were told for each conditional that they should assume that it was really true. For each of the four conclusions to be made on each conditional, students were presented with the major and the minor premise on the screen. They were asked if they could conclude that a given conclusion was true for sure, if it was not true for sure, or if no definite conclusion was possible. For example, for the conditional 'If someone's finger is cut deeply while cooking, then it bleeds', the minor premise 'George's finger is not bleeding.' would have been presented to test the logical form MT. The students would have been asked 'Based on what he knows, what can Peter [the central character in our cover story] say for sure?', and the answer options 'George's finger has just been cut deeply while cooking' (yes), 'George's finger has not just been cut deeply while cooking' (no) and 'He cannot be sure whether George's finger has just been cut deeply while cooking or not' (uncertain).

The everyday conditional reasoning tasks contained familiar causal conditionals (with the antecedent and the consequent being the cause and effect, respectively). The verbal structure of the instructions, conditionals, and answer alternatives were based on previous studies on conditional reasoning (e.g. Markovits & Lortie-Forgues, 2011).

The conditionals with mathematical content dealt with situations, which contained mathematical structures. The specific structures and related concepts were multiplication and addition, since these concepts are included in the national curriculum up to grade 2. Comprehension questions were included for the conditionals, to control if students understood the instructions and the situation, in which the major rule was embedded. For example, the instructions for one of these situations were the following:

'Peter is walking with the little explorers and they just found some treasure boxes. We know that the boxes contain some blue and red diamonds. Each blue diamond is worth 3 gold coins. Each red diamond is worth 2 gold coins'.

The corresponding comprehension question was: '*In a treasure box there is 1 blue diamond and 2 red diamonds. How many gold coins is this worth?'* In case the child provided the correct answer (7), the reasoning tasks followed. After a wrong answer the researcher repeated the instructions and posed the comprehension question for a second and last time. The answer was recorded, and the researcher continued without providing any feedback or hints to students. For mathematical conditionals, only answers on reasoning tasks were included in the analyses, for which the corresponding comprehension question was answered correctly. In grade 2, answers to 62 out of 128 presented conditionals were excluded, with eight students being excluded on all conditionals. In grade 4, 26 of 132 conditionals were excluded and in grade 6, 15 out of 148 conditionals. In grade 4 and grade 6, at least two conditionals were included for each student.

The conditional (major rule) in the example before was *'If the box contains exactly 2 blue diamonds and 3 red diamonds, then the diamonds in the box are worth 12 gold coins. It is certain that this is really true.'* For example, for the logical form MT the following instructions were presented: *'This is Stelios. The diamonds in his box are not worth 12 gold coins. Based on what he knows, what can Peter say for sure?'* The alternative answers were, parallel to the everyday conditionals: '*The box contains exactly 2 blue diamonds and 3 red diamonds'* (yes), *'The box does not contain exactly 2 blue diamonds and 3 red diamonds'* (no), and '*He cannot be sure whether the box contains exactly 2 blue diamonds and 3 red diamonds or not'* (uncertain). All tasks were tested through a pilot study ensuring their appropriateness for this age range (Datsogianni, Ufer, & Sodian, 2018).

Working memory test: Working memory capacity was chosen as a control variable, because it has been found to predict mathematics skills (Holmes, & Adams, 2006), as well as logical reasoning skills (Nakamichi, 2007; 2011). To measure working memory capacity, a backward digit span test (WISC-IV Digit Span Subtest) was used, as in previous studies on young children's conditional reasoning skills (e.g. Nakamichi, 2007; 2011). Specifically, sequences of digits were read out loud and students were asked to repeat them in the reversed order. In the beginning they were provided with an example of a sequence of 3 digits (e.g., 9-2-7) and were asked to reproduce it backwards; in case of a correct reply the test continued. In case of a wrong reply, the correct response was given (7-2-9), then a second example was presented. Regardless of the reply to the second example, no feedback on this was given. The test consisted of seven pairs of digit sequences, with digit sequences within each pair consisting of the same number of digits (e.g. item 1 consisted of trial 1 (digits: 2-5) and trial 2 (digits: 6-3). The number of digits increased by one after each pair. The test was discontinued after failure on both trials of any pair. No hints were given on any of the items. Each item was scored 2 (if the child passed both trials), 1 (if the child passed only one item) or 0 (if the child failed both trials).

Analyses: The answers to the conditional reasoning tasks were analyzed using Generalized Linear Mixed Models (GLMM), a generalization of logistic regression. It allows analyzing the data on item level, but still takes into account dependencies between answers provided by each student and on each task. The package lme4 in R was used (Bates, Mächler, Bolker, & Walker, 2015). Grade level (grades 2, 4, and 6) was included as a between-subject factor, and content (everyday vs. mathematical), wording of the minor premise (positive vs. negative) and type of conclusion (definite vs. indefinite) as within-subject factors. Wald Chi-Square tests were used to compare models during model selection, and to analyze omnibus effects. Planned contrasts of estimated marginal means were used to compare performance in different cells of the experimental design. Bonferroni correction was applied when analyzing multiple contrasts along the same factors.

Ethics statement: The ethics approval was obtained from the Centre of Educational Research and Evaluation of Cyprus Pedagogical Institute as well as the Cyprus Ministry of Education and Culture. Back-and-forth professional translation, from the original English language of the interview protocol into the Greek language and back, was held. Parents and students were informed that the participation in the study was completely voluntary, that answers would be handled confidentially, and that they could stop their participation at any time without any further consequences.

4.1.4 Results

Descriptive results: Overall, 59.5% of students' answers in the conditional reasoning tasks were correct. Figure 3 shows the frequency of each answer option by conditionals' content and logical form. MP tasks were mostly answered correctly, while uncertain conclusions were more frequent for MT tasks. For AC and DA tasks, the most frequent wrong responses were a positive (AC) resp. a negative (DA) response.

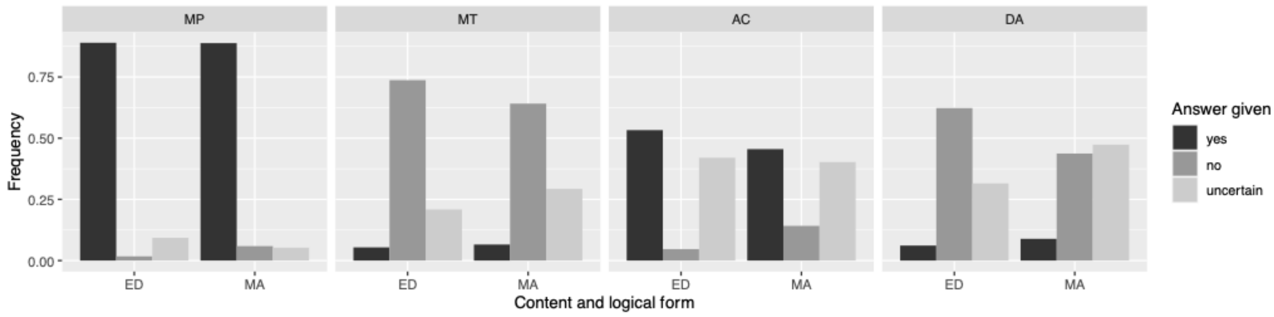


Figure 3

Frequency of answers given by conditional content and logical form. ED: everyday content, MA: mathematical content.

Model selection: In an initial step of model selection, we decided on the random intercepts and slopes to be included. Given the low number of conditionals per person, we analyzed only the random slopes for main effects of grade level and logical form over the different conditionals. Initial analyses indicated that including the random slope of grade level over conditionals lead to a zero variance component for the intercept, so we excluded the random slope for this factor. Initial analyses indicated further, that including random slopes for interaction effects over individuals lead to singular model fit. Since singular models are prone to misinterpretation (Bates, Kliegl, Vasishth, & Baayen, 2018), we decided to analyze only random slopes of main effects of content and logical form over individuals. Chi-square difference tests indicated that leaving out the random slope for grade level over conditionals from a model containing random intercepts and random slopes for all mentioned main effects over conditionals and persons did not affect model fit significantly. However, removing the random slopes for the remaining main effects over persons and removing the random slopes for the remaining main effects over conditionals significantly each affected model fit significantly. Thus, we decided to select the model with random intercepts and random slopes for *wording of the minor premise* and *type of conclusion* over conditionals, as well as random intercepts and random slopes for *wording of minor premise*, *type of conclusion*, and *content* over individuals.

Table 5

Chi-square statistics for the fixed main and interaction effects in the final model, in the order of occurrence in the analysis section.

relates to question	*fixed effect*	*df*	χ^2(df)	*p*	
	working memory	1	2.22	.13	
1a	C: content	1	0.83	.36	
1b	G: grade level	2	21.84	<.001	***
1b	G x C	2	0.62	.73	
2a	W: wording of minor premise	1	5.09	.02	*
2a	T: type of correct conclusion	1	90.81	<.001	***
2a	W x T	1	34.20	<.001	***
2a	C x W	1	1.20	.27	
2a	C x T	1	11.00	<.001	***
2a	C x W x T	1	13.88	<.001	***
2b	G x W	2	9.54	<.01	**
2b	G x T	2	0.02	.99	
2b	G x C x W	2	.02	.99	
2b	G x C x T	2	21.56	<.001	***
2b	G x W x T	2	3.55	.17	
2b	G x C x W x T	2	2.57	.28	

Chi-square statistics for the fixed main and interaction effects in the final model are given in table 5. Working memory as a control variable did not predict conditional reasoning scores significantly (table 5). Estimated marginal means and confidence intervals are available in figure 4.

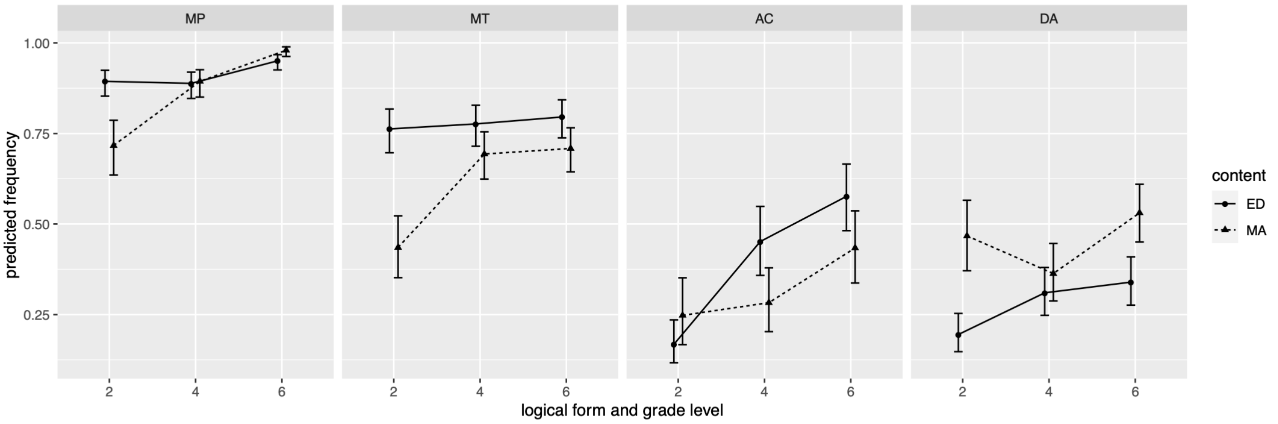


Figure 4

Predicted solution rates and standard error of estimated marginal means (transformed to the 0-to-1 scale for solution rates) of students' conditional reasoning scores by grade level, logical form, and conditionals' content.

General content effect (question 5a): Students answered 59.1% of the questions with everyday content and 60.2% of the questions with mathematical content correctly. The main effect of *content* was not significant (table 5).

Grade-dependent content effect (question 5b): The main effect of grade level was significant (cf. table 5). Over both content conditions, students from grade 6 made significantly more correct inferences than students from grade 4 (67.0% vs. 59.3%, planned contrast $b = .51$, $p < .001$) and students from grade 4 made more correct inferences than students from grade 2 (59.3% vs. 49.2%, planned contrast $b = .60$, $p < .001$). The interaction between *content* and *grade level* was not significant (table 5), providing no evidence of a grade-dependent effect of content on conditional reasoning in general (i.e. over all logical forms).

Grade-independent, form-specific content effects (question 6a): The main and interaction effects referring to logical form (*wording* of minor premise, *type* of conclusion, and their interaction) were significant (table 5). In line with prior results from the everyday contents, a planned contrast analysis showed significantly more correct MP inferences (88.9% of all MP inferences) than MT inferences (69.6%, b = 1.48, p < .001), and significantly more correct MT than AC (41.3%, b = 1.50, p < .001) and DA (48.4%, b = 1.44, p < .001) inferences. Performance of DA and AC inferences did not differ significantly (b = 0.06, p = 1.00). This indicates that the main effect of *conclusion type* (more correct definite than indefinite responses) was modulated by *wording* of the minor premise only for the definite forms, indicated by lower MT compared to MP performance.

Among the interaction effects with the *content* factor, the interactions with *type* of conclusion, as well as the three-way interaction of *content*, *type* of conclusion, and *wording* of the minor premise were significant (table 5). For both contents, fewer indefinite correct conclusions were drawn than definite correct conclusions (ED: 36.8% vs. 81.3%, b = 5.10, p < .001; MA: 43.8% vs. 76.4%, b = 3.76, p < .001). Contrary to our expectations, this effect was significantly more pronounced with everyday than with mathematical content (b = 1.35, p = .02).

A closer analysis indicated that (averaged over all grade levels) there were no significant content-related differences for the two forms with positive wording of the minor premise (MP: 89.0% in ED vs. 88.8% in MA, b = 0.07, p = 1.00; AC: 42.0% in ED vs. 40.3% in MA, b = 0.27, p = 1.00). However, there were marginally more correct MT inferences with everyday content than with mathematical content (73.7% in ED vs. 64.1% in MA, b = 0.78, p = .07), and significantly fewer correct DA inferences with everyday content than with mathematical content (31.5% in ED vs. 47.4% in MA, b = -0.77, p < .05). This indicates that the weaker difference between definite and indefinite forms for mathematical content, compared to everyday content, was mostly due to the two forms with negative wording. It seems to be mainly caused by better DA and (marginally) lower MT reasoning with mathematical content, compared to everyday content.

Grade- and form-dependent content effects (question 6b): Two significant interactions were observed with a connection to this question. Firstly, as a preliminary result, a significant interaction between *grade level* and *wording* of the minor premise occurred (table 5). Such an effect had not been anticipated due to scarce evidence on the effects of positive vs. negative wording of minor premises. A significant performance difference in reasoning with positively worded logical forms could be observed (grade 2: 52.7%, grade 4: 64.6%, grade 6: 74.2%; grade 2 vs. grade 4: $b = 2.73$, $p = .03$, grade 4 vs. grade 6: $b = 3.80$, $p < .01$). For negatively worded logical forms, only a significant difference between grades 2 and 6 could be found (grade 2: 45.8%, grade 4: 54.0%, grade 6: 59.7%, grade 2 vs. grade 4: $b = 1.43$, $p = .46$, grade 4 vs. grade 6: $b = 1.01$, $p = .84$, grade 2 vs. grade 6: $b = 2.35$, $p = .03$). Even in grade 6, performance on negatively worded logical forms was significantly lower than on positively worded forms ($b = 5.18$, $p < .001$) and, indeed differences between grade 2 and grade 6 were significantly stronger for positively than for negatively worded forms ($b = 4.18$, $p < .001$). This speaks for a slower development of reasoning with negatively worded, as compared to positively worded, minor premises.

Moveover, the three-way interaction of *grade level*, *content*, and *conclusion type* was significant (table 5). Averaging over all grades, having to draw an indefinite conclusion had already turned out to have a smaller negative impact on reasoning performance with mathematical content than with everyday content (see question 6b). Contrast analyses revealed that, for all grades and both content types, except for mathematical content in grade 2, questions with indefinite correct conclusions lead to lower performance than questions with definite correct conclusions (table 6). Moreover, these contrasts differed significantly between the two contents only in grade 2 (table 6). This indicates that the content-dependent effect of definite vs. indefinite conclusions was caused mostly by a stronger difference for everyday content, compared to mathematical content, in grade 2.

In particular, this pattern is reflected in better DA reasoning with mathematical content than with everyday content in grade 2 (b = -1.29, p 0 .02). However, this was combined with significantly lower MT reasoning performance with mathematical content as compared to everyday content in grade 2 (b = 1.43, p < .01; for MP: b = 1.20, p = .08; for AC: b = -0.49, p = 1.00).

Table 6

Solution rates and contrasts between definite and indefinite forms by grade level. The content contrast is the difference of the two definite/indefinite contrasts for the respective grade level.

		solution rates				*definite/indefinite contrast*		*content contrast*	
grade	*content*	*MP*	*MT*	*AC*	*DA*	*b*	*p*	*b*	*p*
2	ED	85.9%	70.9%	21.1%	21.9%	6.32	<.001	4.41	<.001
2	MA	71.6%	44.8%	30.3%	44.8%	1.91	.15		
4	ED	86.4%	73.5%	46.2%	34.1%	4.31	<.001	-.13	1.00
4	MA	88.6%	67.9%	36.8%	40.6%	4.44	<.001		
6	ED	93.9%	76.4%	56.5%	37.7%	4.67	<.001	-.24	1.00
6	MA	97.8%	71.0%	48.1%	54.2%	4.91	<.001		

Exploratory analyses of provided answers: To explain the observed pattern of effects, in particular the better reasoning performance with mathematical content compared to the everyday content for indefinite conclusions in grade 2, we analyzed how often students chose the indefinite “uncertain” answer (table 7). We hypothesized, that the specific pattern of higher DA and lower MT reasoning could be due to a stronger tendency to give an uncertain response when reasoning about mathematical content, possibly due to difficulties in retrieving or constructing a representation of the problem situation. The amount of indefinite answers increased descriptively from grade 2 to grades 4 and 6 for everyday contents (table 7). For mathematical contents, it was already rather high in grade 2, and remained on this level in grade 4 and grade 6. In grade 2, significantly more indefinite responses were provided with mathematical content, than with everyday content. In a similar vein, the stronger differences between the two conclusion types by wording of the minor premise (see analyses for question 2a) has to be seen in the context of significantly more indefinite responses on negative worded questions with mathematical content (38.3%) than with everyday content (26.2%, $\chi^2(1) = 23.18$, $p < .001$), while there was no significant difference on positively worded questions (MA: 22.7% vs. ED: 25.6%, $\chi^2(1) = 1.44$, $p = .23$).

Table 7

Frequencies of indefinite ("uncertain") responses by grade level and content, and Chi-square tests for content differences.

grade	*ED*	*MA*	*χ²(1)*	*p*
2	18.6%	31.5%	15.7	<.001
4	28.0%	27.9%	0.0	1.00
6	30.4%	32.2%	0.34	.56

4.1.5 Discussion

The main goal of this study was to investigate students' reasoning with conditionals about mathematical concepts, which are still emerging during primary school. To this end, reasoning with mathematical contents was contrasted against reasoning with familiar causal premises with everyday contents. Based on Mental Model Theory, we started from the assumption that content differences could arise due to general difficulties in representing the situations embedding the conditionals with mathematical content, or from specific problems generating alternative models based on these contents. We primarily assumed the latter, since representing the embedding situations for our conditionals did not require specific mathematical knowledge (and we controlled for comprehension of the situations), while constructing or retrieving alternatives was strongly contingent on such prior knowledge. Finally, we assumed that both effects could be modulated by the increase in mathematical knowledge during primary school age, leading to more pronounced content effects in earlier compared to later primary school grades. Given the sparse evidence, we did not put forward hypotheses regarding positive or negative wording of the minor premise.

Content-overarching findings: Considering average performance over all logical forms and grades, more than 50% percent of students' replies were correct, which is substantially above a guessing probability of 33.3%. This could be taken as evidence in favor of the claim that elementary students do possess early conditional reasoning skills to some extent (e.g., Markovits & Thompson, 2008). In line with previous studies (e.g., Gauffroy & Barrouillet, 2009; Klaczynski, Schuneman, & Daniel, 2004; Markovits & Barrouillet, 2002; Klaczynski & Narasimham, 1998), the results indicate that students' performance increases with grade level. Moreover, in line with prior research (e.g., Barrouillet et al., 1999), we found that MP reasoning was easier for elementary school students than MT reasoning, which was in turn easier than AC and DA reasoning, averaging over both contexts and forms (Markovits et.al, 1996; Markovits, 2000). However, our results did not completely match our predictions made on the basis of MMT.

General content effects: We did not find a general effect of content, averaging over all logical forms and grade levels. This indicates that conditionals with mathematical content did not pose general difficulties for conditional reasoning (beyond comprehension of the framing situations, which was controlled by the comprehension questions). It is important to note, however, that the conditionals with mathematical content were not symbolic mathematical statements, but statements about situations, which included a mathematical structure (e.g., the equivalent of different kinds of collections of red and blue gems in gold coins, or the number of windows in a "dwarf house"). We had assumed that mathematical knowledge was not primarily necessary to represent the conditionals, but to construct alternative models (e.g., different collections of red and blue gems with the same overall value). In this sense, the non-significant main effect of the conditionals' content is in line with the rationale of the conditional reasoning tasks applied in this study. Similar conclusions can be drawn for the interaction of conditionals' content and grade level. Indeed, we did not find evidence of different grade-related differences of overall conditional reasoning between the two contents. This replicates the first results from our own pilot study (Datsogianni, Sodian, & Ufer, 2018). Again, given our assumptions about the necessity of mathematical knowledge in our tasks, this is in line with our expectations. The missing general content difference (everyday vs. mathematical) seems to contradict results from previous studies that the conditionals' content does play a role in conditional reasoning (Markovits & Lortie-Forgues, 2011). However, previous studies (e.g. Markovits & Lortie-Forgues, 2011; Markovits, 2014; Markovits & Vachon, 1990) compared different kinds of relations between antecedent and consequent in everyday contexts (e.g. categorical, causal, or counterfactual conditionals), that are assumed to require different levels of abstraction. Contrary to this, we used causal conditionals with everyday contents, and corresponding conditionals with a mathematical structural mechanism mediating between antecedent and consequent. We assumed that the availability of mathematical knowledge would influence the retrieval of alternative mental models in the mathematical content condition, specifically. That we did not find general content effects seems to indicate, on first sight, that results on early conditional reasoning skills (Markovits et al., 1996) can be transferred from everyday reasoning with familiar causal conditionals to reasoning with mathematical concepts. However, a detailed analysis of logical forms provided a more differentiated picture.

Definite vs. indefinite conclusions: Along our line of reasoning, we had rather expected a pronounced interaction of conditionals' content with the type of conclusion necessary for a given inference: If mathematical knowledge was primarily necessary to construct alternative models for conditionals with mathematical content, a disadvantage should occur for those inferences that require an indefinite conclusion. Thus, we expected that the difference between performances on tasks that require a definite vs. an indefinite conclusion would be more pronounced for mathematical contents. However, what we found was the opposite pattern: Reasoning on indefinite forms actually turned out easier for mathematical contents than for everyday contents, and this effect was particularly pronounced for negatively phrased minor premises (MT and DA reasoning). This finding is not in line with our a-priori predictions based on MMT, and we had no a priori explanation for such a result in the context of MMT. It must, however, be seen in the context of descriptively lower MT performance with mathematical contents, as compared to everyday contents. One reason for this pattern of results might be a stronger tendency for indefinite conclusions ("you cannot say for sure whether, … or not…") when reasoning about mathematical content: Assume students indeed had problems to retrieve alternative models in the mathematical content, in particular for negative premises, which do not provide a specific mathematical situation (e.g., worth 15 gold coins), but only give an indication what is *not* the case (e.g., not worth 15 gold coins). If, moreover, at least some students had already acquired first appropriate reasoning schemata for some logical forms – and solution rates around 70% for everyday contents indicate this at least for MT – it is possible that problems in implementing these schemata could have led to more indefinite responses. One reason for such implementation problems could be the restricted availability of alternative models for mathematical contents. It could be that this goes along with weaker beliefs of students' about the correctness of their given answers, which could be tested in future research.

Similarly, reasoning on logical forms with indefinite correct conclusions did not turn out to be significantly harder with mathematical contents than with everyday content for grade 2 reasoners. Again, this has to be seen in the light of a stronger tendency of second graders to give indefinite responses when reasoning about mathematical contents than for everyday contents. Thus, a similar mechanism as described before might explain this specific effect for grade 2, if we assume that second-graders had most problems constructing alternative models to implement relevant reasoning schemata.

Even though our initially assumed mechanism did not turn out exactly as we hypothesized, the proposed hypothetical explanation would still be in line with the role of mathematical knowledge for reasoning with indefinite logical forms, that modulates the overall positive development of AC and DA reasoning across ages found in prior studies (Janveau-Brennan, & Markovits, 1999; Markovits & Barroulliet, 2002). Consistent with existing studies (Gauffroy & Barrouillet, 2009; Moshman, 2011; Markovits, 2014; Ricco, 2010), logical form turned out as a key factor in describing conditional reasoning performance, in this case regarding the contrast between conditionals with mathematical and everyday contents.

The existence of (form-specific) content effects, in any case, supports the assumption that conditional reasoning is sensitive to domain differences at least in early stages of its development, as it was hypothesized in prior work (Chao & Cheng, 2000). The proposed mechanism however, would indicate that this domain-specificity might originate from the fact that acquired reasoning schemata are generally applicable, but still dependent on domain-specific knowledge, unless they develop into more abstract reasoning schemata that work without recurrence to domain-specific knowledge. However, given the unexpected pattern of results, this proposed mechanism will have to be investigated in further research. In particular, the result does allow not reject the basic assumptions of Mental Model theory, that the conditional inferences are derived not only from a syntactic analysis of the conditionals (based on knowledge stored in long-term memory) but also from a semantic analysis of the conditionals' contents (Markovits, Fleury, Quinn & Venet, 1998). However, also other accounts of conditional reasoning are discussed in the literature, such as the dual-source model of probabilistic conditional reasoning proposed by Klauer, Beller and Hütter (2010)[1], which could provide alternative explanations for the observed results pattern.

Positive vs. negatively worded minor premises: Regarding positive and negative wording of the minor premise, we found lower performance and a slower increase of performance on negatively phrased minor premises. Even though this wording-grade-interaction was not further qualified by an interaction with the conditionals' content in our analyses, we cannot exclude the existence of such moderation due to the surely restricted power of our study. Given that the power would have been sufficient to identify strong effects, the results indicate that the difficulties of negatively phrased minor premises, which have been mentioned in the literature (Schroyens, Schaeken, & d'Ydewalle, 2001) before, might not differ very strongly between everyday reasoning and the kind of reasoning with mathematical concepts we studied. However, given that negatively phrased forms are investigated less frequently, the mechanisms leading to this difference can only be hypothesized, at this point. One reason could, for example, be a difficulty to represent negations in terms of mental models, which are usually assumed to represent what is *possible* under certain assumptions, not what is *impossible* (Johnson-Laird, 2001).

Limitations: Our study has to be considered in light of a set of limitations caused by its specific design. Firstly, we used specific tasks to study conditional reasoning with mathematical concepts, which do not reflect deductive reasoning within a mathematical theory. We considered reasoning about mathematical structures embedded in meaningful situations to be more appropriate to study the role of mathematical concepts in conditional reasoning. Extending the results to deductive reasoning in mathematical theories, as it occurs in later years of education, however, is not straightforward. Secondly, our study has of course limited statistical power to identify small effects. Insignificant findings cannot be taken as evidence for parallel developments or null effects.

---

[1] According to this model, inferences are based on two sources of evidence: logical form (decontextualized source of evidence) and prior knowledge about the reasoning context.

On the other hand, the identified differences do offer support for accounts that argue for a role of knowledge about the conditionals' contents in conditional reasoning. Thirdly, given the cross-sectional design we cannot draw inferences on the individual development of primary school students' development in conditional reasoning. Beyond the grade-level contrasts investigated in our study, future research should also focus on individual developmental trajectories for both reasoning contents, possibly in interaction with the development of mathematical knowledge and skills. Relatedly, our study focused on the primary school age from grade 2 to 6, which is a key phase for the development of everyday conditional reasoning with causal premises. However, content effects might arise at earlier (e.g., for MP), or at later ages, when AC and DA reasoning become more secure. The latter would also correspond to complaints about problems in conditional reasoning in secondary school students (Küchemann, & Hoyles, 2002). Future research could extend the current findings beyond primary school age. As for earlier ages, the availability of the required knowledge about mathematical concepts would have to be taken into account carefully, since these concepts are usually not introduced before grades 1 and 2 of primary school.

Summary: Our results go beyond previous reports on conditional reasoning with everyday concepts and show, that even elementary students are able to make valid deductions for some logical forms when reasoning about mathematical concepts. We acknowledge that there are considerable discussions among researchers about students' ability to make conditional deductions as well as its central importance for scientific reasoning, hypotheses generation and evaluation (Kuhn et al., 1988) and for mathematical thinking (Moshman, 1990; Markovits & Lortie-Forgues, 2011). Our results provide new perspectives on the role of some knowledge about the concepts involved in the statements used for conditional reasoning. This is in line with findings that students do not use general, abstract reasoning rules at this age (Chao & Cheng, 2000). The proposed mechanism describes how knowledge about the conditional contents and more general conditional reasoning skills could interact and develop over elementary school age. In this account, weak mathematical knowledge might inhibit reasoning in forms (e.g., MT), which are at least partially mastered in more familiar contexts according to the literature (e.g., Markovits, 2000; Markovits & Thompson, 2008), but also lead to more correct answers on other logical forms (e.g., DA). Our study does not find indications that with increasing familiarity with the concepts in higher grades, performance in mathematical concepts would exceed performance in everyday contexts. All in all, our results are still in line with a model that puts both mathematical knowledge and conditional reasoning strategies as necessary and mutually non-compensating prerequisites of conditional reasoning with mathematical concepts.

If mathematical knowledge is necessary for conditional reasoning with these concepts, it is an open question if this connection cannot be used in the other direction: Experiencing conditional inferences with mathematical concepts, and discussing alternatives as well as other models for the involved conditionals might not only help to increase conditional reasoning skills, but also add to students' knowledge about these concepts. This is in line with current standards: (CCSS, 2010) argumentation, proof and reasoning should be incorporated regularly into the mathematics classroom from pre-kindergarten through grade twelve.

4.2 STUDY 1: PART II- THE RELATION BETWEEN ELEMENTARY STUDENTS' CONDITIONAL REASONING AND ALTERNATIVES GENERATION: THE CASE OF MATHEMATICS

4.2.1 Brief Introduction

Logical reasoning is considered a key component of advanced thinking amidst human species (Markovits & Barrouillet, 2002) while if-then statements form the basis of scientific mathematical thinking (Markovits & Lortie-Forgues, 2011). Reasoning with if-then statements (e.g. 'If Anna breaks her arm, then it hurts') refers to conditional reasoning. Current theories describe conditional reasoning in younger students as a process that is based on semantic representations of the statements involved. Thus, it is an open question, to which extent domain knowledge influences conditional reasoning skills. In everyday contexts, reasoners' ability to generate multiple alternative models for a given conditional (e.g. 'For which other reasons might Anna's arm hurt?') has been found as a predictive factor to draw valid inferences even from early age (e.g. De Chantal & Markovits, 2017). For conditionals about mathematical concepts, this generation of multiple alternatives is similar to generation of alternative solutions for mathematical problems which according to Leikin and Lev (2007) is considered an indicator of students' creativity and mathematical knowledge.However, our knowledge about the connection between alternatives generation skills and conditional reasoning in the context of elementary school mathematics is still weak.

Conditional reasoning tasks are formed of a conditional rule "if p, then q" as a major premise, and a minor premise (e.g. "q is not true"). The traditional interpretation of conditionals considers p as sufficient, but not necessary for q (Evans & Over, 2004). Four different minor premises lead to four possible logical forms of inference: *Modus Ponens* (*MP*; "p is true, so q is true"), *Modus Tollens* (*MT*; "q is false, so p is false"), *Denial of Antecedent* (*DA*; "p is false, so q or not q") and *Acceptance of the Consequent* (*AC*; "q is true, so p or not p"). Thus, the uncertain logical forms AC and DA do not allow for definite conclusions about p and q respectively. The other two forms (MP and MT) allow valid definite conclusions.

According to Mental Model Theory (MMT) inferences are drawn through the construction of mental models (Johnson-Laird & Byrne, 2002). MMT has been found to describe conditional reasoning accurately not only in adults but also in the age group of primary school children (e.g. Markovits, 2000). Mental models are semantic representations of the possibilities, given the truth of the premises (Johnson-Laird & Byrne, 2002). To derive conclusions, individuals reconstruct the meaning of premises based on their knowledge, to represent what is possible given the premises (Nickerson, 2015). Based on working-memory considerations, Barrouillet & Lecas (1999) proposed an evolvement of individuals' conditional reasoning skills starting from a conjunctive-like interpretation (only one model '*p and q*'*;* correct MP reasoning), to a biconditional ('*p and q*'; '*not-p and not-q*'; correct MP and MT reasoning), and then a conditional interpretation ('*p and q*'; '*not-p and not-q*'; '*not-p and q*'; correct reasoning in each logical form). This evolvement shows up in increasing solution rates for MT, followed by later changes towards a conditional interpretation with increased solution rates for DA and AC. 'Alternatives' are mental models of the type 'not-p and q', which are necessary to arrive at the indefinite conclusions in the AC and DA forms. Beyond model generation, other authors also state that MT and DA are more cognitively demanding compared to MP and AC forms due to the negation statements involved (Johnson Laird & Byrne, 1993).

Conditional Reasoning & Alternatives generation in everyday contexts. According to MMT, generation of mental models, and in particular of alternatives, is based on knowledge about the content of the conditionals. Alternatives generation for a given conditional is considered as a crucial prerequisite to draw valid inferences (Johnson-Laird & Byrne, 2002; Markovits & Barrouillet, 2002).

Studies on reasoning with conditionals from an everyday context (e.g. De Chantal & Markovits, 2017) have shown that alternatives generation skills predict conditional reasoning even from pre-school age on. In many studies, alternatives generation skills are associated with correct AC and DA reasoning, in particular (e.g. Cummins et.al, 1991; Markovits & Vachon, 1990). In addition to alternatives generation, individuals might also generate disablers (mental models of the form 'p and not-q', contradicting the major rule) describing inhibitory factors which might prevent q from occurring, even in the presence of p (e.g., 'Anna took a painkiller, so her arm does not hurt, even though it is broken'; Cummins et al., 1991). Disablers might lead to the rejection of valid conclusions for MP and MT inferences (Janveau-Brennan & Markovits, 1999). Many studies report a positive correlation between the numbers of generated alternatives and disablers (Thompson, 2000; De Neys, Shaeken, & D'Ydewalle, 2002). In line with this, studies with university students revealed that correct DA and AC reasoning correlates negatively with correct MT reasoning (Newstead et.al, 2004). Hence, it is likely that alternatives generation is positively linked with AC and DA reasoning, being not or negatively related with MT reasoning. In prior research with young students, disabler generation is considered less relevant for logical reasoning than alternatives generation (Janveau-Brennan & Markovits, 1999; De Chantal & Markovits, 2017).

Conditional Reasoning & Alternatives generation in mathematical contexts. While the role of alternatives for conditional reasoning is well-studied in the everyday context, it has not been studied for conditionals that involve mathematical concepts (mathematical context; e.g., "If I arrange three rows of four squares each, then I need 12 squares."). In this case, alternatives generation concerns the mental construction of mathematical objects that fulfill 'not-p and q' (e. g., "12 squares could be constructed by six rows of two squares each"), beyond those that represent 'p and q' (or 'not-p and not-q'). Generating such alternative perspectives to mathematical situations is often discussed in research on multiple solutions (Leikin & Lev, 2007). Based on this perspective, alternatives generation can be assumed to require mathematical knowledge of the conditional content (Leikin & Lev, 2007). Beyond a general link between mathematics skills and conditional reasoning skills (Attridge & Inglis, 2013), this could lead to a specific influence of mathematical knowledge on AC and DA conditional reasoning in the mathematical context. Studies in the field of mathematics with university students show negative correlation between MT form and DA as well as AC form (Attridge & Inglis, 2013; Morsanyi, McCormack & O' Mahony, 2017). This backs up the assumption that conditional reasoning in this context is based on mental model construction, similar to the everyday context.

Overall, the existing literature in primary school pupils investigates the relation between alternatives generation and conditional reasoning only in the everyday context (with different levels of abstraction; e.g. Markovits & Lortie-Forgues, 2011). In the mathematical context, research with primary school students either investigate conditional reasoning (Christoforides, Spanoudis & Demetriou, 2016) or multiple solution tasks (e.g. Sullivan, Bourke & Scott, 1997). Yet, to date research not has addressed alternatives generation in relation to conditional reasoning in two different contexts and this study aims to fill this research gap.

4.2.2 Goals and Questions

This study aimed to transfer results on the role of alternatives generation in primary students' conditional reasoning from everyday conditionals to conditionals from a mathematical context. The following research questions are addressed:

(7) Is the influence of alternatives generation skills on conditional reasoning specific to the respective context (everyday vs. mathematical)? Based on the MMT account, we expected alternatives generation in each context to primarily predict conditional reasoning in the corresponding context (e.g. De Chantal & Markovits, 2017).

(8) Do alternatives generation skills predict correct reasoning differently across the four logical forms? Based on prior results and the MMT account, we expected that alternatives generation skills in the everyday context would predict correct AC and DA reasoning (Markovits & Vachon, 1990). However, taking into account that alternatives generation skills are related with the generation of disablers in the everyday context (Thompson, 2000; De Neys, Shaeken & D'Ydewalle, 2002), this might entail a decrease in MP and MT reasoning. For the mathematical context we expected alternatives generation skills to be related to correct AC and DA reasoning, but not to MT reasoning (Attridge and Inglis, 2013; Morsanyi, McCormack & O' Mahony, 2017).

4.2.3 Methods

In this cross-sectional study, *N*=102 students from grades 2, 4, and 6 in Cyprus (average age 12 years, 0 months) were interviewed individually. The feasibility of the instrument was piloted in a previous study, showing that it is accessible to this age group of students (Datsogianni, Ufer & Sodian, 2018). Ethics approval, parental consent signed form, and students' individual oral assents were obtained.

Participants solved four conditional reasoning tasks on each context (Cronbach's α: .62. for everyday and .68 for mathematical context). The everyday conditionals referred to daily life situations. Conditionals in the mathematical context referred to situations that involved mathematical structures which were supposed to be familiar for the participants (e.g., "If a dwarf's house has exactly 3 rows of 4 rooms each, then it has 12 rooms."). All forms (MP, MT, DA, and AC), were included in each task. The order of two contexts, the order of the conditionals in each context, and the order of logical forms for each conditional were randomized across students. Alternatives generation skills were measured afterwards with specific tasks in each context, using the same situations as in the conditionals (4 mathematical and 4 everyday). The reliability scores were good (Cronbach's α=.86 for everyday and .76 for mathematical context). The experimenter (first author), asked students to find as many instances as they could that matched the model 'not-p and q' by drawing their ideas. Participants did not receive positive or negative feedback. The order of two contexts and the order of the situations in each context were randomized across students.

Example of alternatives generation task in the everyday context: "Remember what Peter found out before. If a glass is dropped on the ground in the kitchen, then there is a sound. Peter is at home and hears a sound in his kitchen. Find as many reasons why a sound in the kitchen may occur, as you can."

Example of alternatives generation task in the mathematical context: "Remember that dwarfs build their houses so that there are rooms which all have this form: □ The houses always have one or more rows of rooms which are all equally long. Remember what Peter found out before. If a dwarf's house has exactly 3 rows of 4 rooms each, then it has 12 rooms. How could a dwarf house with 12 rooms look like? Draw as many different houses as you can." In the end of the interview procedure, students solved a working memory test (backward digit span). Separate linear mixed models for each context were used to analyze the data using the package lme4 in R, controlling for working memory skills. The factor logical form and the alternatives generation scores were included in the model. Insignificant interactions between logical form and the alternatives generation scores were removed from the model prior to the final analysis. In the mathematical context, the random factor controlling for individual differences explained no variance.

4.2.4 Results

Overall, students solved 59.1% of the items correctly in the everyday context (ED), and 57.1% in the mathematical context (MA). In both contexts, MT was solved significantly less of than MP (ED: MP 88.9%, MT 73.7%, $p < .001$; MA: MP 84.3%, MT 61.5%, $p < .001$) and AC less often than MT (ED: AC 42.2%, $p < .001$; MA: AC 37.2%, $p < .001$). DA was solved less often than AC in the everyday context (ED: DA 31.5%, $p < .05$), while the difference was not significant in the mathematical context (MA: DA 45.65, $p = .16$).

Students generated more alternatives per task in the everyday context (range 1-11, $M = 4.25$) compared to the mathematical context (range 0-5, $M = 2.08$). It is worth noting that the number of possible alternative solutions was more limited in the mathematical context compared to the everyday context.

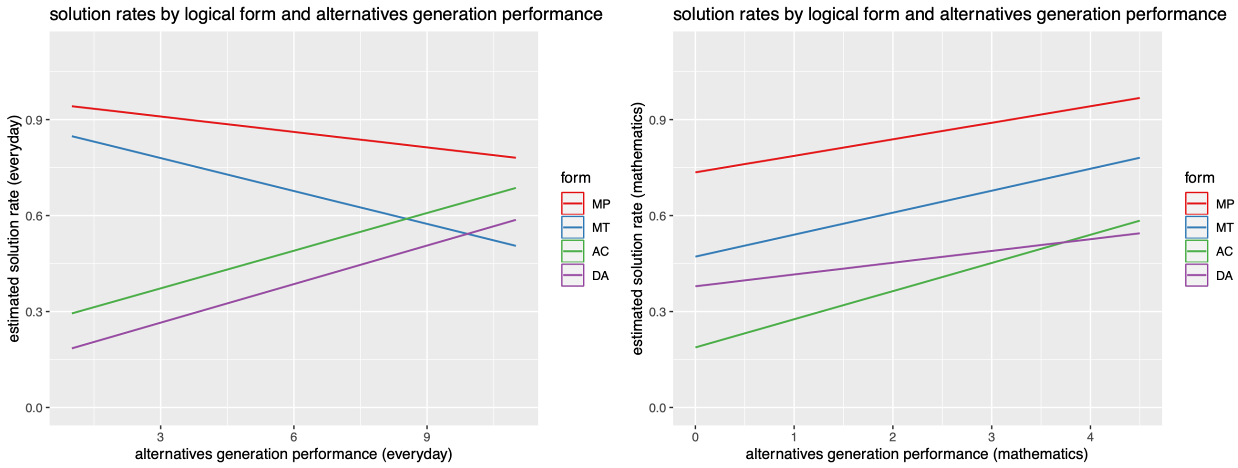


Figure 5a and 5b: Estimated solution rates and 95% confidence intervals by alternatives generation and logical form: (5a; left) everyday, (5b, right) mathematical context.

Regarding research question (7), mathematical alternatives generation ($F(1,398)=13.7$, $p < .001$), but not everyday alternatives generation ($F(1,398)=0.32$, $p = .57$), showed a significant effect on conditional reasoning in the mathematical context. Conditional reasoning in the everyday context was related significantly to mathematical alternatives generation ($F(1,98)=8.35$, $p < .001$) but – over all logical forms – not to everyday alternatives generation ($F(1,98)=0.70$, $p = .41$).

Regarding research question (8), we found a significant interaction between *logical form* and alternatives generation scores only for everyday conditional reasoning ($F(3,300)=6.57$, $p < .001$, fig. 1a). In this context, everyday alternatives generation predicted correct reasoning significantly positively in the AC ($B = 0.039$, $CI_{95\%}[0.008, 0.070]$) and in the DA ($B = 0.040$, $CI_{95\%}[0.010, 0.071]$) form, but negatively for MT ($B = -0.034$, $CI_{95\%}[-0.065, -0.004]$). For example, $B = 0.039$ indicates an estimated increase in the conditional reasoning solution rate of 3.9% per generated everyday alternative. The non-significant interaction ($F(3,398)=0.75$, $p = .52$, fig. 1b) between *logical form* and mathematical alternatives generation for mathematical reasoning indicates, that alternatives generation (positively) predicted conditional reasoning comparably strongly for all logical forms in this context.

4.2.5 Discussion

Regarding research question (7), alternatives generation skills in the everyday context did not have a significant main effect on conditional reasoning in the same context in general (cf. De Chantal & Markovits, 2017). In particular given the significant interaction between everyday alternatives generation scores and logical form, this pattern is in line with prior findings. Mathematical alternatives generation skills predicted logical reasoning in both contexts. Since alternatives generation is mainly based on prior knowledge of the respective content (Leikin & Lev, 2007), this is in line with previous evidence about the relation between logical reasoning and mathematical knowledge that has been found in the literature, before (Attridge & Inglis, 2013). Regarding research question (2) and the results for the everyday context were similar to those found in prior studies. In particular, alternatives generation in the corresponding context was predictive for correct AC and DA reasoning (Cummins et.al, 1991; Markovits & Vachon, 1990). It is also observed that alternatives generation (in this context) inhibits correct MT reasoning; probably students extend the strategy of generating antecedents to generating and (incorrectly) interpreting inhibitors (De Neys, Shaeken, & D'Ydewalle, 2002). However, as for the mathematical context, it seems that alternatives generation is generally predictive of conditional reasoning skills, mostly independent of the logical form. As said above this might reflect a general relation between logical reasoning and mathematical knowledge (Attridge & Inglis, 2013). On the other hand – and we cannot differentiate this explanation in this study – it might also be that knowledge about the mathematical content is necessary to generate a representation of mathematical conditionals and any kind of related mental model (not only of the type 'not-p and q'). If the sole representation of mathematical conditionals is indeed so strongly dependent on corresponding knowledge, this might mask a specific effect of alternatives generation for DA and AC in this context.

Overall, the results of this study indicate that reasoning with mathematical conditionals is, overall, not substantially harder or easier than reasoning with everyday conditionals. However, the analysis of the relation to alternatives generation points to possible differences in the reasoning process. For example, it might be that, in spite of early conditional reasoning skills in the everyday context, students are not able to activate the corresponding strategies in the mathematical context, due to restricted ability to represent the conditionals' meanings in mental models. If indeed problem representation turns out as the primary problem, this implies the necessity not only to practice conditional reasoning, but also to carefully consider students' mathematical knowledge before engaging with basic deductions about mathematical concepts in mathematics instruction. However, reflecting deductions and the meaning of conditionals about mathematical concepts during classroom instruction might also help to build up this prerequisite knowledge.

One possible limitation arising from this study is that alternatives generation tasks addressed only questions with given consequents, for which students had to create as many possible antecedents, as possible. Future studies might separately measure the generation of an initial mental representation of the conditional and alternatives generation. The negative relation between alternatives generation and MT reasoning, moreover, might be explained through investigating the generation of disablers in future research. However, alternative antecedents have been considered more central to conditional reasoning of young students than disablers (Janveau-Brennan & Markovits, 1999). Thus, this study provides first insights for the relation between alternatives generation and conditional reasoning with mathematical concepts, which can be extended in further research.

5. STUDY 2: WHAT KIND OF INSTRUCTION LEADS TO IMPROVEMENT OF ELEMENTARY STUDENTS' CONDITIONAL REASONING?

5.1 BRIEF INTRODUCTION

Conditional reasoning is among the most well-documented reasoning types in the field of developmental psychology. However, as mentioned before, logical reasoning skills of elementary students in the mathematical context have been rarely investigated (Datsogianni, Sodian, Markovits, & Ufer, in press; Datsogianni, Ufer, & Sodian, 2018), while existing evidence shows that students even from grade 2 are capable of approaching conditional reasoning tasks in both everyday and mathematical contexts possessing basic skills at least to some extent. Research in psychology so far is mainly based on elementary students' difficulties in deductive reasoning or the factors that might affect this reasoning performance. Yet, our knowledge about the instructional methods that might help students to make deductions and cope with conditionals more efficiently, is still poor. So, the development of this reasoning type potentially through instructional trainings is of crucial importance. Therefore, the aim of this study is to investigate the effects of three different trainings on students' conditional reasoning skills.

Trainings on conditional reasoning. Based on different theories, prior research has shaped its trainings accordingly. For example, studies which were based on Mental Logic theories (see Chapter 1.2.1), designed their training on truth tables, emphasising the logical structures of conditional reasoning (e.g., Muller et al., 2001). Yet, the results were not that promising. Hence, when it comes to the understanding of logical relations as the main way to master conditional reasoning, it does not seem sufficient.

Another experimental study aiming at the improvement of elementary students' conditional reasoning skills in AC and DA reasoning, was based on the training of meta logical principles (e.g. logical consistency, inferential identity of each scheme, etc) and the theory of Demetriou (Christophorides, Spanoudis, & Demetriou, 2016). The results shown that through this type of training, MP and MT reasoning were fully developed by 3rd graders, however, the two invalid forms, AC and DA, were not mastered even by 6th graders. Based on MMT (see chapter 1.2.3; Johnson-Laird's, 2006), the trainings are focused on the mental models that students have to think and handle appropriately.

Alternatives generation process. According to MMT (see chapter 1.2.3), generation of mental models, and in particular of alternatives, is based on knowledge about the content of the given conditional and is considered as a necessary prerequisite to draw valid inferences (Johnson-Laird & Byrne, 2002; Markovits & Barrouillet, 2002). In detail, the correct AC and DA reasoning is related to the generation of potential alternative antecedents (e.g., Cummins et al., 1991; Markovits & Vachon, 1990). The difficulties in conditional reasoning in primary school are not restricted in the context of mathematics but also in the everyday one (e.g. Datsogianni, Sodian, & Ufer, in press), showing that positive effects of trainings might be appropriate scaffolding tools for students. Students tend to focus on the given premise, and find it difficult to go beyond this information and generate alternative possibilities. These difficulties may arise due to many factors according to the literature but based on the MMT which the whole dissertation is based on, deductive skills mainly depend on the number of generated alternative possibilities (Byrne, 2005; Johnson-Laird, 2001; Markovits & Barrouillet, 2002).

In line with MMT, we consequently assume that these difficulties might arise due to a limited range of alternatives that students have in mind; hence trainings based on alternative generation processes might function as scaffolds for students' conditional reasoning skills. As Klaczynski and Narasimham (1998) claimed MP and MT inferences seem to be developed by early adolescence, which is in line with Baroulliet & Lecas model (1999) and the conductive and biconditional phase, respectively.

The sample of this study refers to 5th and 6th graders who belong to the age range mentioned about the latter model. Hence, the main reason of investigating possible scaffolds is about the development of AC and DA reasoning which require the alternatives generation skills.

Alternatives generation & creativity. Alternatives generation procedures, however, do not only underlie conditional reasoning but many types of reasoning (Piaget, 1987; Gauffroy & Barrouillet, 2011) such as divergent thinking. In particular, according to Markovits and Brunet (2012) divergent thinking and conditional reasoning share some commonalities which are based on the necessity of alternatives generation under the given content and instructions. Divergent thinking which is a fundamental component of creativity, requires this process of alternatives generation which is similar to the one needed in conditional reasoning. In other words, the key process in both thinking forms is the understanding of uncertainty. It is accepted that divergent thinking can be taught, even in young students (Cliatt, Shaw, & Sherwood, 1980), at least some components of it (i.e fluency and flexibility; Leikin, 2013), however the link between divergent thinking and conditional reasoning at the altar of investigating potential scaffold function of divergent thinking priming has been only lately studied.

In detail, this link between the two concepts has been investigated by Markovits and De'Chantal (2019), and the results they obtained from preschool children (4 -5 years old) show that alternative generation priming affects positively AC reasoning. It is worth noting that AC reasoning as mentioned in the previous studies of this dissertation, refers to the individual's ability to go beyond the given representation of the conditional and generate alternative conclusions that are not explicitly described in the relevant task. Another study with elementary students by Markovits and Brunet (2012) revealed that after a short training on divergent thinking, students' solution rates in logical reasoning tasks were improved. However, these results apply only to students of high socio-financial status in line with the authors' assumptions. The results of another study (Daniel &

Klaczynski, 2006), show that providing students with an explicit alternative antecedent when presenting premises improves AC reasoning and the understanding of uncertainty.

Based on this evidence that divergent thinking and conditional reasoning are based on remarkably similar reasoning processes, and that a priming of divergent thinking improves conditional reasoning, we wanted to extend these results to the mathematical context. In the mathematics context multiple solution tasks are those that allow some space for alternative solutions to students, and this type of tasks is used in this study, by encouraging of which, we will test conditional reasoning scores in the post-testing.

Alternatives generation in Mathematical context & creativity. Looking at the alternative generation process of divergent thinking from the mathematics education view, this process can be evaluated and promoted through multiple solution mathematical tasks. The latter according to Silver (1997) refer to tasks that can be solved by many solution methods and potentially might encourage the development of students' fluency and flexibility. Another definition by Leikin and colleagues (Leikin, 2006, 2009; Leikin & Levav-Waynberg, 2008) characterise as multiple solution tasks those assignments that require explicitly from students to solve them in many different ways.

All in all, taken together findings from the psychology field and mathematics education domain concerning trainings on young students' conditional reasoning the following arise. First of all, results from psychology suggest that instruction can support young students develop some logical forms, as MP, MT (e.g. Christophorides, Spanoudis, & Demetriou, 2016) and AC forms (Markovits and De'Chantal, 2019; Daniel & Klaczynski, 2006) and students can master these forms even earlier than expected according to the developmental trajectory (e.g. Barroulliet & Lecas, 2002).

From the mathematics education domain under the lens of arguments and proofs construction, it seems that elementary students show deductive reasoning skills through the appropriate instructional context (e.g., Maher & Martino, 1996; Reid, 2002; Stylianides, 2007b; Zack, 1997). However, no evidence is found in the research field of mathematics education regarding any trainings on students' mathematical conditional reasoning skills in primary school. Consequently, the main question aim of this study is the investigation of the role of mathematical instruction through multiple solution tasks in nurturing the development of students' conditional reasoning skills.

Aim and Research Questions: As mentioned, there is limited evidence concerning elementary students' conditional reasoning skills in mathematics (Datsogianni, Ufer, & Sodian, 2018; Datsogianni, Sodian, Markovits, & Ufer, in press), so our knowledge of how we could support students to develop these skills is based on even less data. Therefore, the aim of this study is to investigate which kind of instruction leads to improvement of elementary students' conditional reasoning in mathematics and beyond. Taking these insights, we might get closer to some potential educational implications in mathematics teaching and learning. Hence the research questions of this study are the following:

(α) Do students' scores in conditional reasoning tasks improve after a short-term training with alternatives generation tasks which refer to the contrary-to-fact context (Experimental Condition A)?

Our hypothesis is that after the training of these tasks there might be an increase in students' search for alternatives in general, meaning significantly higher solution rates in conditional reasoning tasks between the pre and post testing. This increase might appear in both everyday and mathematical context, referring especially to the AC form based on prior research. Regarding the mathematical content there is no evidence about this potential relation, so we do not have a specific hypothesis for this.

(b) Do students' scores in conditional reasoning tasks improve after a short-term training with mathematical multiple solution tasks (Experimental Condition B)?

It is accepted that knowing about the existence of multiple solutions in a mathematics context leads to improved alternative generation in mathematics tasks. Therefore, we expect significantly higher results in the post-test phase compared to the pre-post solution rates, especially in the mathematical context and the two forms AC and DA which are founded on the understanding of uncertainty. Based on the very limited evidence that connects mathematical and everyday context regarding conditional reasoning, we do not have any specific hypothesis about the possible effect of this training on the other two contexts of the post-testing. With regards to the control conditions, our hypothesis is that just dealing with mathematics content, without considering alternatives (condition C), improves conditional reasoning in mathematics whereas just the process of reading a literature text has no observable influence on students' conditional reasoning.

In general, since according to Barroulliet and Lecas (1999), and prior research, students in the ages of our sample are already in the biconditional phase of reasoning, we assume that the alternative generation priming will not affect the MP and MT reasoning, rather than AC and DA reasoning which are mastered only in the phase of conditional reasoning. If students, however, become too suspicious in terms of the uncertainty after the short-term training, and do not accept the necessity of the major rule, we might observe lower MP, and MT reasoning in the post tests of students who participated in these conditions.

5.2 METHODS

Sample and design: In total, 285 elementary students (5th grade *N*= 144, and 6th grade *N*=141) from eleven public schools in Cyprus participated in this study. Parents' written consent and children's oral assent were obtained from all participants.

The study was carried out in a 60' minute classroom activity in small groups (3-12 students) under the instructions of the researcher. The activity consisted of the pre-testing, the experimental and control conditions and the post-testing. The pre-test measured the baseline of students' conditional reasoning skills through one task in the everyday context and one in the mathematical one. These tasks were identical to those used in our previous study (see study 1). The experimental conditions were conditions A and B. In particular, condition A referred to the training with (4) alternative generation tasks in the imaginary (contrary-to-fact) context, whereas condition B was about the training with (4) multiple solution tasks in mathematics. In detail, an example from condition A is the following task adopted from previous studies by Markovits and Lortie-Forgues (2011):

'*Imagine that you have gone to a planet where everything happens very differently to here on Earth. On this planet, if you clean a sweater with ketchup, then the sweater will become clean. Can you imagine other ways of cleaning a sweater on this planet? Give as many responses as you can.'*

At the end of this process the researcher was reading aloud the script regarding the feedback on these 3 tasks of condition A. During the feedback sharing, the researcher provided students with many different solutions. Students were not allowed to change anything on their booklet after getting hints and new possible answers. The structure of Condition B was similar; in this condition the alternative generation was held through the approach of 3 mathematical multiple solutions tasks. An example is the following:

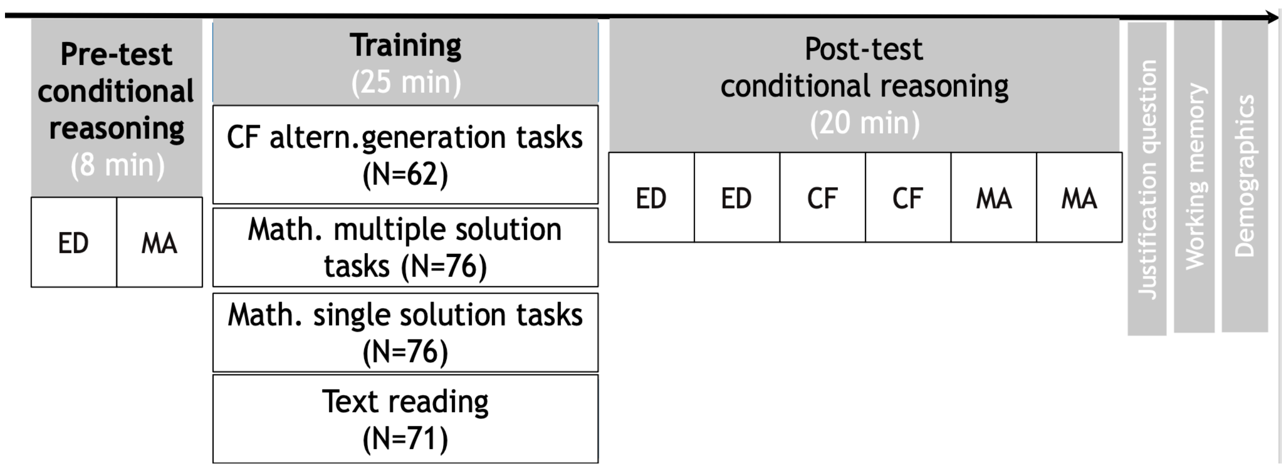
'*A friend gives you a riddle: He says that yesterday in the park, there were some dogs and some children, but no other animals. Altogether, there were 600 legs in the park. How many children and dogs could have been in the park, which have together 600 legs? Find as many and different solutions as you can! Remember that you have to have 600 legs altogether. Each dog has 4 legs and each child has 2 legs*.'

In the end of the feedback section of each experimental condition, the researcher mentioned to all students the importance of alternatives generation in our everyday life and mathematics: '*As we see there are so many and different solutions for every single question or task and all of them can be right. It is really important to think if there any other possibilities for each task because this makes us being creative! Being creative is not only helpful for our everyday life but also for mathematics and other school subjects! In the next tasks try to do the same: Be creative, think about different possibilities that did not come to your mind in the first moment.*

Condition C and D were the control conditions referring to solving of single solution mathematical tasks and text reading, respectively. Immediately afterwards all students were given the post-test. The post-testing consisted of 6 conditional reasoning tasks from the imaginary (2 tasks), everyday (2 tasks) and mathematical context (2 tasks). In order to avoid order effects, between the blocked tasks (everyday and contrary-to-fact context vs. mathematical; block A vs. block B, respectively) we created two versions for each condition, varying the blocks order across the sample. In total eight different versions of booklets were distributed randomly among the groups.

Figure 6

Design of study 2



Procedure: Students from each school were mainly mixed across classes (in agreement with the teachers and school principals) and randomly allocated to the control or experimental conditions.

This classroom activity was carried out in a quiet classroom in each school. In the beginning, after some general information provided, a short pre-test which tested students' conditional reasoning skills through two tasks- one for the everyday context and one for the mathematical context- took place. Researcher read orally the instructions of the each task and students were asked to fill in their answers on their booklets individually. All students had the same time limit to fill in each task and nobody could proceed to the next page earlier (relevant oral and written instructions were emphasized).

In the sequel, students were tested on the post-test conditional reasoning tasks. The post-testing pertained to mathematical, everyday and contrary-to-fact context (2 tasks each). For the last logical form of each booklet referring to the post-testing, students had to justify their answer. After the justification question students had to fill in some working memory tasks. More explicitly, working memory was assessed through the digit-span subtest (Wechsler, 2003), the visual matrix task (Swanson, 1992, 1995), and the mapping and directions tasks (Swanson, 1992), exactly as used by Osterhaus, Koerber and Sodian (2016) for a paper-and-pencil assessment. The structure of the working memory assessment was adopted by the latter study. Regarding the digit-span, students had to recall and write in their booklets two sequences (one by one) of 7 and 10 digits in a forward order, which were read aloud to students by the researcher. In the sequel, students approached for 5 seconds each of the two visual matrix tasks (each 5 x 5 matrix contained 10 dots) that they had to remember and draw on their booklets. Referring to the mapping and directions tasks which followed, participants were provided with two routes of horizontal and four vertical streets, and had to recall and draw them on the relevant part of their booklet. The routes consisted of 9 and 14 route segments respectively. In total, 60 items (segments) were used to assess working memory.

At the end of the classroom activity, students had to provide some demographic information (age, number of books at home and language spoken at home). Working memory scores, qualitative data obtained from the training processes as well as students' justifications mentioned before, will not be presented in this study.

Analyses*:* The answers to the conditional reasoning tasks were analyzed descriptively as well as using Generalized Linear Mixed Models (GLMM), a generalization of logistic regression. It allows analyzing the data on item level, but still takes into account dependencies between answers provided by each student and on each task. The package Lme4 in R was used (Bates, Mächler, Bolker, & Walker, 2015).

Ethics approval**:** Parents and students were informed that the participation in the study was completely voluntary, that answers would be handled confidentially, and that they could stop their participation at any time without any further consequences. Two students wanted to withdraw their participation after having started it; hence their answers are not included in our dataset. The ethics approval was obtained from the Centre of Educational Research and Evaluation of Cyprus Pedagogical Institute as well as the Cyprus Ministry of Education and Culture.

5.3 RESULTS

Firstly, the descriptive analysis revealed a mixed set of results. In detail, table 8 shows that among the experimental groups (condition A & B) there was an increase in students' conditional reasoning scores between the pre and post-test regarding the DA form. On the contrary, the latter groups after the short training provide lower scores in AC reasoning compared to their pre´-test AC scores. The same pattern is observed in MT reasoning while MP scores seem not to vary between the pre and post-test.

Table 8

Means and standard deviations

Condition	Logical form	Pre-test *M(SD)*	Post-test *M(SD)*
A (Contrary-to-fact)	MP	98.38 (12.70)	89.51 (30.75)
	MT	91.93 (27.45)	65.04 (47.87)
	AC	51.61 (50.38)	25.20 (43.59)
	DA	9.67 (29.80)	52.03 (50.16)
B (Mathematical)	MP	94.73 (22.47)	92.76 (25.99)
	MT	81.33 (39.22)	67.76 (46.89)
	AC	68.42 (46.79)	36.84 (48.39)
	DA	14.47 (35.41)	44.73 (49.88)
C (Control-Math)	MP	98.70 (11.39)	88.88 (31.53)
	MT	89.61 (30.71)	73.37 (44.34)
	AC	59.74 (49.36)	32.67 (47.05)
	DA	15.58 (36.50)	39.59 (49.07)
D (Control-Text)	MP	98.59 (11.86)	91.54 (27.91)
	MT	73.23 (44.58)	70.42 (45.80)
	AC	65.71 (47.80)	27.46 (44.79)
	DA	9.85 (30.02)	46.80 (50.07)

Note: *M* and *SD* represent means and standard deviation, respectively.

Students who participated in a control condition (Condition C or D) also show an increase of DA reasoning scores between the pre-test and the post-test and a decrease in AC reasoning scores. The two definite forms, MP and MT do not vary across the time of pre- and post-test.

Looking at each context separately, the following findings are observed. In particular, GLMM analysis about students' conditional reasoning scores in the mathematical context by time (pre-post), specific condition and logical form, revealed that students who took part in control conditions C and D developed more positively than those who participated in experimental Conditions A and B, but this applies only for the definite forms (figure 7). Estimated marginal means and confidence intervals are available in figure 8.

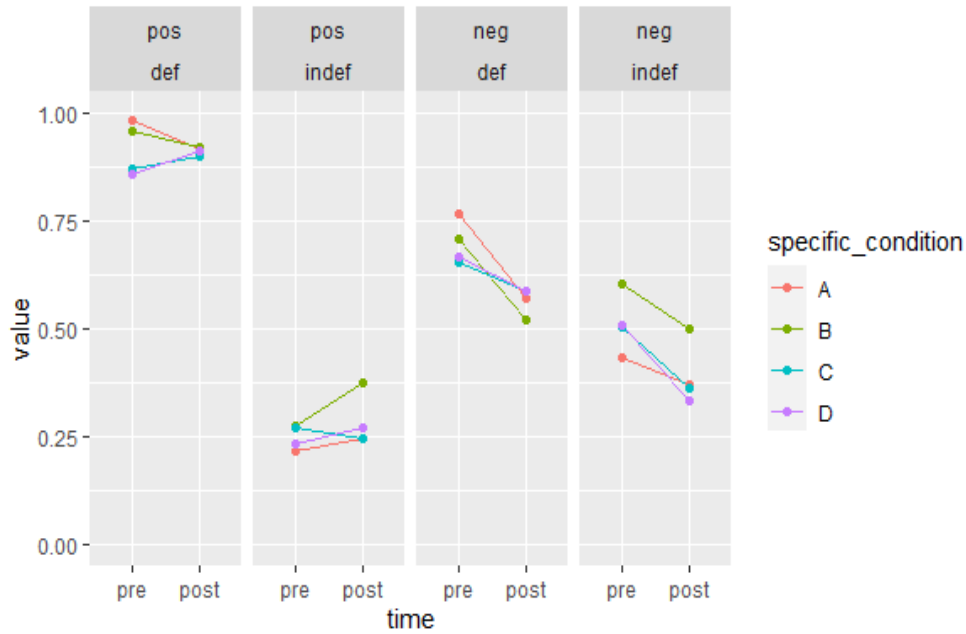


Figure 7: Solution rates and standard error of estimated marginal means (transformed to the 0-to-1 scale for solution rates) of students' conditional reasoning scores in the mathematical context by time (pre-post), specific condition (A, B, C, D) and logical form (positive, negative; definite, indefinite).

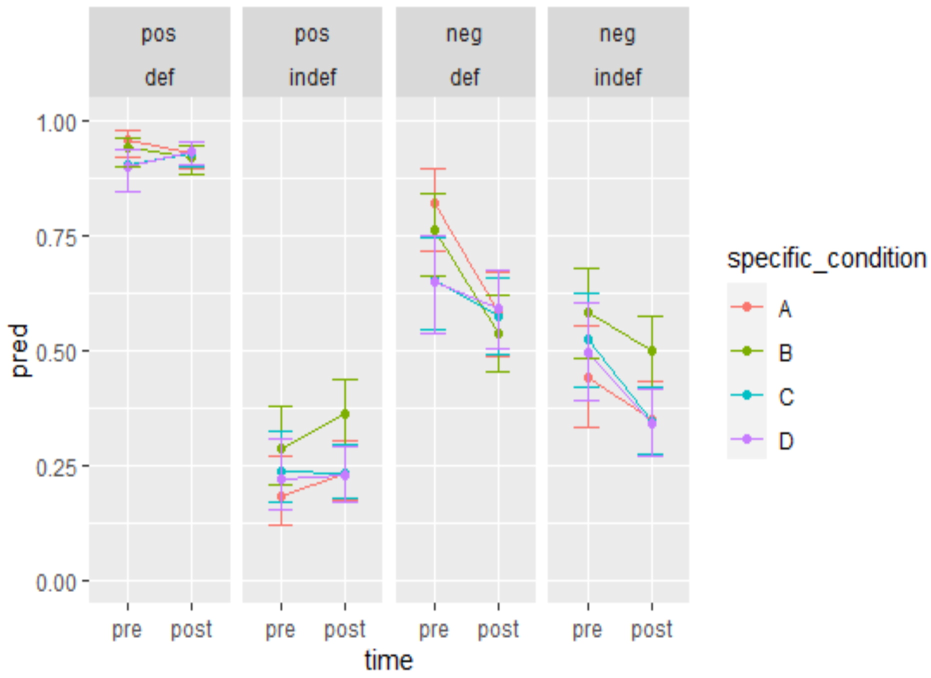


Figure 8: Estimated solution rates and standard error of estimated marginal means of the selected model (transformed to the 0-to-1 scale for solution rates) of students' conditional reasoning scores in the mathematical context by time (pre-post), specific condition (A, B, C, D) and logical form (positive, negative; definite, indefinite).

As for the everyday context and students' solution rates by time (pre and post) and logical form, the groups B (experimental) and D (control) seem to develop better than A (experimental) and C, but this again applies only for definite forms (figure 9). Estimated marginal means and confidence intervals are available in figure 10.

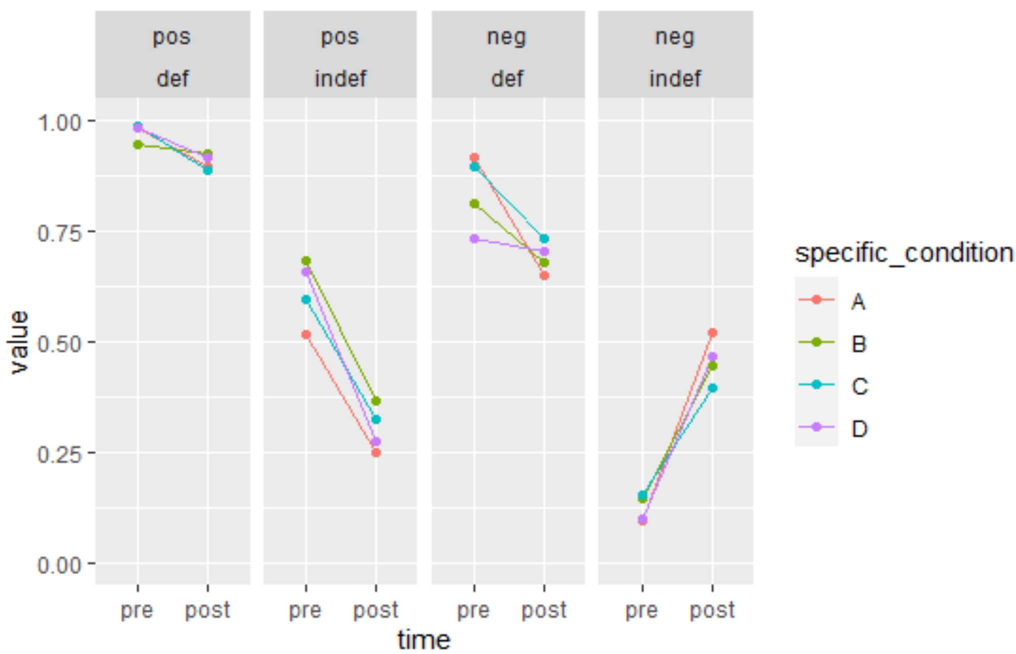


Figure 9: Solution rates and standard error of estimated marginal means (transformed to the 0-to-1 scale for solution rates) of students' conditional reasoning scores in the everyday context by time (pre-post), specific condition (A, B, C, D) and logical form (positive, negative; definite, indefinite).

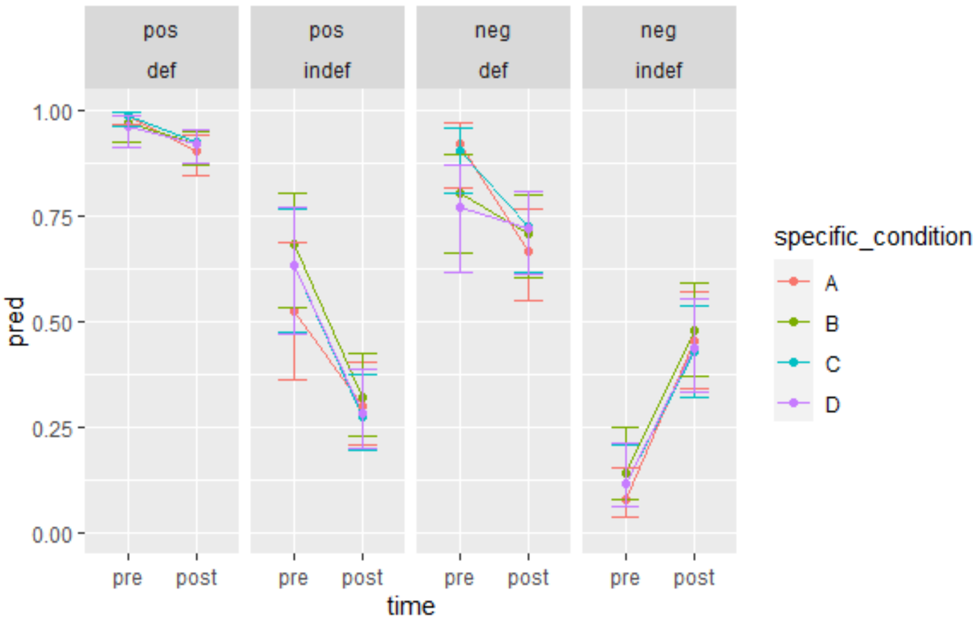


Figure 10: Estimated solution rates and standard error of estimated marginal means for the selected model (transformed to the 0-to-1 scale for solution rates(of students' conditional reasoning scores in the everyday context by time (pre-post), specific condition (A, B, C, D) and logical form (positive, negative; definite, indefinite).

In other words, students who participated in the experimental condition within mathematics, showed better results in MP and MT reasoning (definite forms) after the training. The same pattern followed students who were allocated in control condition about literature text reading.

5.4 DISCUSSION

The fact that MP and MT scores decrease after the alternative generation training is in line with our predictions that students seem to not accept the necessity of the major rule, after the training and a possible wrong interpretation of uncertainty. So, even if according to Klaczynski and Narasimham (1998) MP and MT inferences are found well-developed by early adolescents, the aforementioned finding about the disablers' activation after the alternatives generation priming has been observed among adolescents (Daniel, & Klaczynski, 2006) as well as adults (Byrne, Espino, & Santamaria, 1999) in previous studies. This probably further explains the findings of our GLMM analysis in the mathematical context and the positive development in MP and MT reasoning (definite forms) by students who participated in the control conditions (C and D). According to the latter studies, students in experimental groups after the training might become more suspicious, concerning the necessity of the truth of major rule, approaching probably probabilistic interpretations.

The findings concerning the increased DA reasoning in post-tests compared to the pre-testings, were expected, as the aim of each of the experimental conditions was the practice of alternative generation process and the right interpretation of uncertainty, which correct AC and DA reasoning is based on (e.g., Cummins et al., 1991; Markovits & Vachon, 1990). However, the finding that AC decreases in the mathematical context after the training cannot be explained accordingly. We assume that students probably considered that an uncertain answer comes with a given uncertain antecedent (DA form) and that a specific consequent (AC form) comes with the specific antecedent provided by the major rule.

As for the results by context separately and the fact that only control conditions (C and D) developed positively in the mathematical context concerning the definite forms, it is in line with prior research about disablers' generation after practicing alternative generation processes (Daniel,

& Klaczynski, 2006; Byrne, Espino, & Santamaria, 1999). Regarding the everyday context, the results are mixed with the mathematical divergent thinking priming to lead to improved MP and MT reasoning and this finding replicates prior research by Christophorides, Spanoudis, & Demetriou, (2016). The finding that students who participated in the control condition of text reading (D), showed improved MP and MT reasoning at the post-testing, probably arises due to the fact that students dealt with this type of tasks in the pre-test, in conjunction with developmental pattern by Barroulliet and Lecas (1999) according to which, students at these ages are already at the biconditional phase hence their MP and MT reasoning can be already well-developed.

In conclusion, the results of this study are mixed and even if most of the current findings can be explained by the literature, none of the trainings seemed to totally function as scaffolds for students' conditional reasoning skills in each context and logical form. However, this study requires further statistical analysis and interpretation. The analysis of qualitative data of students' justification as well as the process data will shed some light on the reasoning procedure that underlies these reasoning outcomes. A limitation of this study is the short duration of each training and the absence of a follow-up post testing. Another limitation arising from this study is that alternatives generation trainings addressed only tasks with given consequents, for which students had to create as many possible antecedents, as possible. Future experimental studies might also combine the practice of conditional reasoning through multiple solution tasks with the development of students' mathematical knowledge before engaging with basic mathematical deductions. Apart from the limitations, this study shows the importance of the development of mathematical logical reasoning in primary school through the instruction of creativity- realistic tasks with multiple solutions. This study adds some necessary evidence in this research area which is not well-documented so far.

6. GENERAL DISCUSSION

6.1 SUMMARY OF STUDIES

6.1.1 Pilot Study

The pilot study indicated that the developed instrument is accessible to elementary school students. Moreover, the questionnaire survey results as well as one-to-one interviews replicated early conditional reasoning skills reported in studies from developmental psychology. The results were consistent with this assumption and MMTs: only for the mathematical context did grade 6 students outperform their peers from grade 4, but results were comparable for the everyday context. This opens a discussion about the important role of specific knowledge for conditional reasoning in mathematical contexts. At the same time results indicate that the differences between the two contexts are not only due to task understanding: while the four logical forms were of similar difficulty for the everyday context, the logical forms requiring to analyze alternative models (AC, AD, MT) were significantly harder than MP in the mathematical context. Finally, results indicate that the 6th graders deeper knowledge of mathematical concepts does not show equally for all logical forms, but primarily for MP and AC items.

6.1.2 Study 1

Study 1 extends results concerning the gradual development of primary students' conditional reasoning with everyday concepts to reasoning with mathematical concepts adding to our understanding about the link between mathematics and conditional reasoning in primary school. Consistent with previous findings, even 2nd graders were able to make correct inferences on some logical forms. Controlling for WM, there were significant effects of grade and logical form, with stronger growth on MP and AC than on MT and DA. The main effect of context was not significant, but context interacted significantly with logical form and grade level. The pattern of results was not consistent with the predictions of MMT. Based on analyses of students' chosen responses support we propose an alternative mechanism explaining the specific pattern of results. The study indicates that deductive reasoning skills arise from a combination of knowledge of domain-general principles and domain-specific knowledge.

Our results go beyond previous reports on conditional reasoning with everyday concepts and show, that even elementary students are able to make valid deductions for some logical forms when reasoning about mathematical concepts. Our results provide new perspectives on the role of some knowledge about the concepts involved in the statements used for conditional reasoning. This is in line with findings that students do not use general, abstract reasoning rules at this age. The proposed mechanism describes how knowledge about the conditional contents and more general conditional reasoning skills could interact and develop over elementary school age. In this account, weak mathematical knowledge might inhibit reasoning in forms (e.g., MT), which are at least partially mastered in more familiar contexts according to the literature but also lead to more correct answers on other logical forms (e.g., DA). Our study does not find indications that with increasing familiarity with the concepts in higher grades, performance in mathematical concepts would exceed performance in everyday contexts. All in all, our results are still in line with a model that puts both mathematical knowledge and conditional reasoning strategies as necessary and mutually non-compensating prerequisites of conditional reasoning with mathematical concepts.

The analysis of the relation to alternatives generation points to possible differences in the reasoning process. For example, it might be that, in spite of early conditional reasoning skills in the everyday context, students are not able to activate the corresponding strategies in the mathematical context, due to restricted ability to represent the conditionals' meanings in mental models.

6.1.3 Study 2

Study 2 aimed to investigate the effect of different alternatives generation trainings on students' conditional reasoning scores. The results of this study were mixed and even if most of the current findings can be explained by the literature, none of the trainings seemed to totally function as scaffolds for students' conditional reasoning skills in each context and logical form. In detail, MP and MT scores decreased after the alternative generation training which was in line with our predictions that students seem to not accept the necessity of the major rule, after the training and a possible wrong interpretation of uncertainty. This can probably also explain the finding that there was a positive development in MP and MT reasoning (definite forms) in the mathematical context by the students who participated in the control conditions. The findings concerning the increased DA reasoning in post-tests compared to the pre-testings, were expected as the aim of each of the experimental conditions was the practice of alternative generation process and the right interpretation of uncertainty, which correct AC and DA reasoning is based on. However, the finding that AC decreases in the mathematical context after the training cannot be explained accordingly. Regarding the everyday context, the results were mixed with the mathematical divergent thinking priming to lead to improved MP and MT reasoning replicating prior research. Students who participated in the control condition of text reading (D), shown improved MP and MT reasoning at the post.-testing, probably arises due to the fact that students dealt with this type of tasks in the pre-test, and since they are already at the biconditional phase, their MP and MT reasoning might be already well-developed. However, further statistical analysis is required; so the above discussion cannot be generalized.

6.1.4 Overall Conclusions

Overall, the results of study 1 indicate that reasoning with mathematical conditionals is, overall, not substantially harder or easier than reasoning with everyday conditionals. Summarising, the role of conditional reasoning in mathematics can hardly be denied. Thus, the findings of the aforementioned studies imply the instructional necessity to include and practice conditional reasoning tasks in primary school within the context of mathematical statements by providing opportunities to students to interpret and discuss mathematical conditionals, as well as generate alternative antecedents for these conditionals. Even though open question remains, the study extends evidence, that knowledge of mathematical concepts and being able to reason about them (with conditionals) are strongly related.

6.2 IMPLICATIONS

This section presented a summary of the practical implications of this thesis which mainly concern researchers who wish to evaluate and add evidence in the area of conditional reasoning within mathematics in primary schools as well as teachers and curriculum developers in the area of mathematics. Pilot study emphasized the feasibility of mathematical conditional reasoning tasks even for early ages. From study 1 and 2 some implications arose too.

If indeed problem representation turns out as the primary problem, this implies the necessity not only to practice conditional reasoning, but also to carefully consider students' mathematical knowledge before engaging with basic deductions about mathematical concepts in mathematics instruction. However, reflecting deductions and the meaning of conditionals about mathematical concepts during classroom instruction might also help to build up this prerequisite knowledge.
If we agree that valid reasoning with mathematical concepts is a goal of classroom instruction, this implies the necessity to practice simple deductions on all logical forms when dealing with mathematical concepts. Moreover, it will be necessary to discuss why certain conclusions can or cannot be drawn – particularly for the more complex logical forms such as AC, DA, and MT (cf.

Schroyens, Schaeken, & d'Ydewalle, 2001) – with a focus on the models underlying these inferences.

Furthermore in an attempt to investigate which instructional content might function as scaffold for students' conditional reasoning in mathematics, study 2 demonstrated the importance of the development of mathematical logical reasoning in primary school through the instruction of creativity- realistic tasks with multiple solutions. This study gains some necessary evidence in this research area which is not well-documented so far.

6.3 LIMITATIONS

Undoubtedly the findings of the aforementioned studies should be interpreted critically and in light of the following limitations. The first limitation refers to the pilot study and most importantly, the small sample size – even though tackled statistically by analyzing on item level – forbids taking far-reaching conclusions apart from backing up the feasibility of the measurement and raising hypotheses for further research.

Study 1 has to be considered in light of a set of limitations caused by its specific design. Firstly, we used specific tasks to study conditional reasoning with mathematical concepts, which do not reflect deductive reasoning within a mathematical theory. We considered reasoning about mathematical structures embedded in meaningful situations to be more appropriate to study the role of mathematical concepts in conditional reasoning. Extending the results to deductive reasoning in mathematical theories, as it occurs in later years of education, however, is not straightforward. Secondly, our study has of course limited statistical power to identify small effects. Insignificant findings cannot be taken as evidence for parallel developments or null effects. On the other hand, the identified differences do offer support for accounts that argue for a role of knowledge about the conditionals' contents in conditional reasoning. Thirdly, given the cross-sectional design we cannot draw inferences on the individual development of primary school students' development in conditional reasoning. Beyond the grade-level contrasts investigated in our study, future research should also focus on individual developmental trajectories for both reasoning contents, possibly in interaction with the development of mathematical knowledge and skills. Relatedly, our study focused on the primary school age from grade 2 to 6, which is a key phase for the development of everyday conditional reasoning with causal premises. However, content effects might arise at earlier (e.g., for MP), or at later ages, when AC and DA reasoning become more secure. The latter would also correspond to complaints about problems in conditional reasoning in secondary school students (Küchemann, & Hoyles, 2002). Future research could extend the current findings beyond primary school age. As for earlier ages, the availability of the required knowledge about mathematical concepts would have to be taken into account carefully, since these concepts are usually not introduced before grades 1 and 2 of primary school.

Another limitation arising from study 1 and study 2 is that alternatives generation tasks addressed only questions with given consequents, for which students had to create as many possible antecedents, as possible. Future studies might separately measure the generation of an initial mental representation of the conditional and alternatives generation. Lastly another limitation of study 2 is the short duration of each training and the absence of a follow-up post testing.

6.4 SUGGESTIONS FOR FUTURE RESEARCH

Future research will have to collect really longitudinal data and include explicit measures of mathematical knowledge to support the interpretations. This future research might also benefit from the developed mathematical tasks to describe the role of mathematical knowledge in conditional reasoning with mathematical concepts in more detail.

Further research will be necessary to describe the negation effects in more detail, in particular focusing on the role of mathematical knowledge. The negative relation between alternatives generation and MT reasoning, moreover, might be explained through investigating the generation of disablers in future research. However, alternative antecedents have been considered more central to conditional reasoning of young students than disablers (Janveau-Brennan & Markovits, 1999). Thus, this study provides first insights for the relation between alternatives generation and conditional reasoning with mathematical concepts, which can be extended in further research.

Study 2 requires further future analysis and interpretation. The analysis of qualitative data of students' justification as well as the process data will shed some light on the reasoning procedure that underlies these reasoning outcomes. Future experimental studies might also combine the practice of conditional reasoning through multiple solution tasks with the development of students' mathematical knowledge before engaging with basic mathematical deductions. A combined training (contary-to-fact and mathematical) might also be interesting to be tested.

REFERENCES

Almuna Salgado, F. (2016). Developing a Theoretical Framework for Classifying Levels of Context Use for Mathematical Problems. *Mathematics Education Research Group of Australasia*.

Attridge, N., & Inglis, M. (2013). Advanced mathematical study and the development of conditional reasoning skills. *PloS one, 8*(7).

Barrouillet, P., & Lecas, J. F. (1999). Mental models in conditional reasoning and working memory. *Thinking & Reasoning, 5*(4), 289–302.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2018). Parsimonious mixed models. arXiv preprint arXiv:1506.04967, *https://arxiv.org/pdf/1506.04967.pdf*, last accessed May, 15th 2020.

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software, 67*(1), 1–48.

Byrnes, J. P., & Overton, W. F. (1988). Reasoning about logical connectives: A developmental analysis. *Journal of Experimental Child Psychology*, 46(2), 194-218.

Byrne, R. M., Espino, O., & Santamaria, C. (1999). Counterexamples and the suppression of inferences. *Journal of Memory and Language*, *40*(3), 347-373.

Braine, M. D., & O'Brien, D. P. (1991). A theory of if: A lexical entry, reasoning program, and pragmatic principles. *Psychological review*, 98(2), 182.

Chao, S. J., & Cheng, P. W. (2000). The emergence of inferential rules: The use of pragmatic reasoning schemas by preschoolers. *Cognitive Development, 15*(1), 39-62.

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. Memory & cognition, 19(3), 274-282.

Cummins, D. D. (1995). Naive theories and causal deduction. Memory & Cognition, 23, 646–658.

Cummins, D. (1996b). Evidence of deontic reasoning in 3- and 4-year-old children. *Memory & Cognition, 24,* 823–829.

Christoforides, M., Spanoudis, G., & Demetriou, A. (2016). Coping with logical fallacies: A developmental training program for learning to reason. *Child development*, *87*(6), 1856-1876.

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory and Cognition, 19*, 274-282.

Cliatt, M. J. P., Shaw, J. M., & Sherwood, J. M. (1980). Effects of training on the divergent-thinking abilities of kindergarten children. *Child Development*, 1061-1064.

Datsogianni, A., Ufer, S. & Sodian, B. (2018). Elementary students' conditional reasoning skills in mathematical and everyday contexts. Bergqvist, E., Österholm, M., Granberg, C., & Sumpter, L. (Eds.). (2018). *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 331-338). Umeå, Sweden: PME.

De Chantal, P. L., & Markovits, H. (2017). The capacity to generate alternative ideas is more than inhibition for logical reasoning in preschool-age children. *Memory & cognition, 45*(2), 208-220.

De Neys, W., Schaeken, W., & D'ydewalle, G. (2002). Causal conditional reasoning and semantic memory retrieval: A test of the semantic memory framework. *Memory & cognition, 30*(6), 908-920.

De Neys, W., & Everaerts, D. (2008). Developmental trends in everyday conditional reasoning: The retrieval and inhibition interplay. *Journal of Experimental Child Psychology, 100*(4), 252-263.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics, 53*(1), 5-34.

English, L. D., & Halford, G. S. (1995). Mathematics education: Models and processes. Mahwah, New Jersey: Lawrence Eribaum Associates.

Evans, J. St. B. T., & Over, D. E. (2004). If. Oxford, England: Oxford University Press.

Evans, J. S. B., & Over, D. E. (1996). Rationality in the selection task: Epistemic utility versus uncertainty reduction.

Garnham, A., & Oakhill, J. (1996). The mental models theory of language comprehension. Models of understanding text, 313-339.

Foltz, C., Overton, W. F., & Ricco, R. B. (1995). Proof construction: Adolescent development from inductive to deductive problem-solving strategies. *Journal of Experimental Child Psychology, 59*(2), 179-195.

Gauffroy, C., & Barrouillet, P. (2009). Heuristic and analytic processes in mental models for conditionals: An integrative developmental theory. *Developmental Review, 29*(4), 249-282.

Greatorex, J. (2014). Context in Mathematics questions. *Research Matters, 17*, 18-23.

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. S. B. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking & Reasoning, 10*(2), 175-195.

Holmes, J., & Adams, J. W. (2006). Working memory and children’s mathematical skills: Implications for mathematical development and mathematics curricula. *Educational Psychology, 26*(3), 339-366.

Inglis, M., & Simpson, A. (2008). Conditional inference and advanced mathematical study. *Educational Studies in Mathematics, 67*(3), 187-204.

Inglis, M., & Simpson, A. (2009). Conditional inference and advanced mathematical study: Further evidence. *Educational Studies in Mathematics, 72*(2), 185-198.

Janveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology, 35*(4), 904–911.

Johnson-Laird, P. N. (2001). Mental models and deduction. *Trends in cognitive sciences, 5*(10), 434-442.

Johnson-Laird, P. N., & Byrne, R. M. (2002). Conditionals: a theory of meaning, pragmatics, and inference. *Psychological review, 109*(4), 646.

Johnson-Laird, P. N. (1999). Deductive reasoning. Annual review of psychology, 50(1), 109-135.

Janveau-Brennan, G., & Markovits, H. (1999). The development of reasoning with causal conditionals. *Developmental Psychology, 35*(4), 904.

Klaczynski, P. A., & Narasimham, G. (1998). Representations as mediators of adolescent deductive reasoning. *Developmental Psychology, 34*(5), 865.

Klauer, K. C., Beller, S., & Hütter, M. (2010). Conditional reasoning in context: A dual-source model of probabilistic inference. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 36*(2), 298-323.

Klaczynski, P. A., Schuneman, M. J., & Daniel, D. B. (2004). Theories of conditional reasoning: a developmental examination of competing hypotheses. *Developmental Psychology, 40*(4), 559.

Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., & Yotive, W. (1988). The development of scientific thinking skills. Academic Press.

Küchemann, D., & Hoyles, C. (2002). Students' understanding of a logical implication and its converse. In A. Cockburn (Ed.): *Proceedings 26th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 3–241). Norwich, UK: PME.

Leikin, R., & Lev, M. (2007). Multiple solution tasks as a magnifying glass for observation of mathematical creativity. In J. Woo, H. Lew, K. Park, D. Seo (Eds.): *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 3, pp. 161-168). Seoul, Korea: PME.

Leikin, R., & Levav-Waynberg, A. (2008). Solution spaces of multiple-solution connecting tasks as a mirror of the development of mathematics teachers' knowledge. *Canadian Journal of Science, Mathematics, and Technology Education, 8*(3), 233-251.

Leikin, R., & Lev, M. (2007, July). Multiple solution tasks as a magnifying glass for observation of mathematical creativity. *In Proceedings of the 31st international conference for the psychology of mathematics education (Vol. 3, pp. 161-168).* Seoul, Korea: The Korea Society of Educational Studies in Mathematics.

Litkowski, E. C., Duncan, R. J., Logan, J. A., & Purpura, D. J. (2020). When do preschoolers learn specific mathematics skills? Mapping the development of early numeracy knowledge. *Journal of Experimental Child Psychology*, 195, 104846.

Liu, H., Ludu, M., & Holton, D. (2015). Can K-12 Math Teachers Train Students to Make Valid Logical Reasoning? In X. Ge, D. Ifenthaler, & M. Spector (Eds.): *Emerging Technologies for STEAM Education* (pp. 331–353). Heidelberg: Springer.

Markovits, H. (2000). A mental model analysis of young children's conditional reasoning with meaningful premises. *Thinking & Reasoning, 6*(4), 335-347.

Markovits, H. (2014). Conditional reasoning and semantic memory retrieval. In A. Feeney, V. Thompson (Eds.): Reasoning as Memory (pp. 67-84). London & New York: Psychology Press.

Markovits, H. (2017). In the beginning stages: Conditional reasoning with category based and causal premises in 8-to 10-year olds. *Cognitive Development, 41*, 1-9.

Markovits, H., & Barrouillet, P. (2002). The development of conditional reasoning: A mental model account. *Developmental Review, 22*(1), 5-36.

Markovits, H., & Brunet, M. L. (2012). Priming divergent thinking promotes logical reasoning in 6- to 8-year olds: But more for high than low SES students. *Journal of Cognitive Psychology, 24*(8), 991-1001.

Markovits, H., Fleury, M. L., Quinn, S., & Venet, M. (1998). The development of conditional reasoning and the structure of semantic memory. *Child development, 69*(3), 742–755.

Markovits, H., & Thompson, V. (2008). Different developmental patterns of simple deductive and probabilistic inferential reasoning. *Memory & Cognition, 36*(6), 1066–1078.

Markovits, H., & Quinn, S. (2002). Efficiency of retrieval correlates with “logical” reasoning from causal conditional premises. *Memory & Cognition, 30*(5), 696-706.

Markovits, H., & Lortie−Forgues, H. (2011). Conditional reasoning with false premises facilitates the transition between familiar and abstract reasoning. *Child Development, 82*(2), 646-660.

Markovits, H., & Vachon, R. (1990). Conditional reasoning, representation, and level of abstraction. *Developmental Psychology, 26*(6), 942.

Markovits, H., Venet, M., Janveau−Brennan, G., Malfait, N., Pion, N., & Vadeboncoeur, I. (1996). Reasoning in young children: Fantasy and information retrieval. *Child Development, 67*(6), 2857-2872.

Markovits, H., & Doyon, C. (2004). Information processing and reasoning with premises that are empirically false: Interference, working memory, and processing speed. *Memory & cognition, 32*, 592-601.

Morsanyi, K., McCormack, T., & O'Mahony, E. (2018). The link between deductive reasoning and mathematics. *Thinking & Reasoning*, *24*(2), 234-257.

Morsanyi, K., Kahl, T., & Rooney, R. (2017). The link between math and logic in adolescence: The effect of argument form. In M. E. Toplak & J. Weller (Eds.), *Individual differences in judgment and decision making from a developmental context* (pp. 166– 185). Hove: Psychology Press.

Morsanyi, K., McCormack, T., & O'Mahony, E. (2018). The link between deductive reasoning and mathematics. *Thinking & Reasoning, 24*(2), 234-257.

Morsanyi, K., & Szücs, D. (2014). The link between mathematics and logical reasoning. The Routledge international handbook of dyscalculia and mathematical learning difficulties, 101.

Moshman, D. (1990). The development of metalogical understanding. Reasoning, necessity, and logic: Developmental perspectives, 205-225.

Moshman, D. (2011). Adolescent Rationality and Development: Cognition, Morality, and Identity. Taylor & Francis.

Muller, F. A. (2001). Sets, classes, and categories. *The British journal for the philosophy of science, 52*(3), 539-573.

Nakamichi, K. (2007). Young children's conditional reasoning, working memory, and inhibitory control. *Japanese Journal of Educational Psychology, 55*(3), 347-358.

Nakamichi, K. (2011). Age differences in the relationship among conditional inference, working memory and prepotent response inhibition. *Psychologia, 54*(1), 52-66.

National Governors Association Center for Best Practices, Council of Chief State School Officers (2010). Common Core State Standards (Mathematics). Washington D.C.

Nickerson, R. (2015). Conditional Reasoning: The Unruly Syntactics, Semantics, Thematics, and Pragmatics of "if". Oxford University Press.

Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *The Quarterly Journal of Experimental Psychology Section A, 57*(1), 33-60.

Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology, 25*(1), 147-166.

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. Journal of Experimental Psychology: *Learning, Memory, and Cognition, 26*(4), 883.

Paulus, C. (2009). Die 'Bücheraufgabe' zur Bestimmung des kulturellen Kapitals bei Grundschülern [The 'book task' as a measure of the cultural capital with primary pupils]. http://hdl.handle.net/20.500.11780/3344, last retrieved May, 15th 2020.

Piaget, J. (1987). *Possibility and necessity* (Vol. 1). U of Minnesota Press.

Reid, D. A. (2002). Describing Young Children's Deductive Reasoning.

Robinson, K. M., Price, J. A., & Demyen, B. (2018). Understanding arithmetic concepts: Does operation matter? *Journal of Experimental Child Psychology, 166*, 421-436.

Ricco, R. B. (2010). Development of Deductive Reasoning across the Life Span. R.M. Lerner, M.E. Lamb and A.M. Freund (Eds.): The Handbook of Life−Span Development.

Schroyens, W. J., Schaeken, W., & d'Ydewalle, G. (2001). The processing of negations in conditional reasoning: A meta-analytic case study in mental model and/or mental logic theory. *Thinking & reasoning, 7*(2), 121-172.

Silver, E. A. (1997). Fostering creativity through instruction rich in mathematical problem solving and problem posing. *Zdm, 29*(3), 75-80.

Stylianides, A. J., Stylianides, G. J., & Philippou, G. N. (2004). Undergraduate students' understanding of the contraposition equivalence rule in symbolic and verbal contexts. *Educational Studies in Mathematics, 55*(1-3), 133-162.

Stylianides, A. J., & Stylianides, G. J. (2007). The mental models theory of deductive reasoning: Implications for proof instruction. In D. *Pitta-Pantazi & G. Philippou (Eds.): Proceedings of the 5th CERME Conference* (pp. 665-674), Larnaca, Cyprus: CERME.

Stylianides, A. J. (2007). Proof and proving in school mathematics. *Journal for research in Mathematics Education*, 289-321.

Stylianides, G. J., & Stylianides, A. J. (2008). Proof in school mathematics: Insights from psychological research into students' ability for deductive reasoning. *Mathematical thinking and learning, 10*(2), 103-133.

Sullivan, P., Bourke, D., & Scott, A. (1997). Learning mathematics through exploration of open-ended tasks: Describing the activity of classroom participants. In E. Pehkonen (Ed.): *Use of open-ended problems in mathematics classroom* (pp. 88-105). Univ. of Helsinki.

Thompson, V. A. (2000). The task-specific nature of domain-general reasoning. *Cognition, 76*(3), 209-268.

Zack, V. (1997). You have to prove us wrong": Proof at the elementary school level. In *Proceedings of the 21st Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 291-298). Lahti: University of Helsinki.

APPENDIX

PILOT STUDY- Interview protocol

Introduction

The purpose of this interview is to gather responses from students about their understanding of conditional statements in different contexts and their ability to draw logically valid conclusions from each of them. To understand the conditional statements does not demand high prior knowledge. However a basic understanding of everyday contexts and mathematical concepts is necessary to draw valid conclusions, as long as reasoning is not based on abstract logical rules, but on mental models. The students in this pilot study are intended to draw their conclusions, having each time on their hands one rule, one premise, and three answer alternatives. Moreover, the interview survey students' mathematical knowledge in terms of their ability to generate alternative causes for a given conclusion. The verbal structure of the tasks in general and of the four logical forms is parallel for all tasks in the interview. The sequence of the two contexts will be mixed among the students (two sequences: first mathematical, then everyday context and vice versa).
The main focus of this study is to understand the difficulties that the children may encounter during logical reasoning tasks in two different contexts and minimize the external difficulties that may arise because of the instruments' design. Lego figures and pictures are planned just to support the understanding of the instructions among the 2nd and 4th graders.

Input from this pilot study will be shared with the supervisors Prof. Dr. Ufer and Prof. Dr. Sodian. The results are going to support the further development of the tasks, and can provide important contribution to our understanding of the development of students' logical reasoning skills.

Part 1: Rapport Building, Background information & Instructions

1. General Rapport Building

Good morning, my name is Anastasia. Thank you for coming.

Let's go around the room. Please tell me your first name and what grade you are in.

- *What do you like to do with your free time?*
- *What are your favourite subjects in school?*

2. Control questions

- *Could you tell me when your birthday is?*
- *So, how old are you?*
- *Which language/s do you speak at home with your family?*
- *Do you know approximately the whole number of the books that are in your house?* (Given pictures/ visualizations of a full bookshelf, a complete bookcase and two complete bookcases).

3. Instructions

My purpose in talking with you today is to solve some tasks that I am sure that you will really enjoy. For each task I read, there are four questions that I would like for you to give a try and figure out an answer.

We are interested in how you think about the tasks. There are no right or wrong answers, or desirable or undesirable answers. I would like you to feel comfortable saying what you really think and how you figure out the tasks. If it's okay with you, I will be video-recording our conversation, specifically only your hands, since it is hard for me to write down everything while simultaneously talking with you. Everything you say will remain confidential, meaning that only my teammates and I will be aware of your answers.

- *Do you have any questions before we begin?*
- *Are you willing to begin together?*

4. Example-Training of the answer alternatives

We are starting with a short game:

(The researcher holds up a red marble that the child can see)

- *Can you say <u>for sure</u> that there is a red marble on my hand?*

Possible answers: Yes, there is/ No, there is not/ I cannot say for sure whether there is or not.

(The researcher holds up only a green marble that the child can see)

- *Can you say <u>for sure</u> that there is a red marble on my hand?*

Possible answers: Yes, there is/ No, there is not/ I cannot say for sure whether there is or not.

(The researcher is hiding one marble of the two on her hand)

- *Can you say <u>for sure</u> that there is a red marble on my hand?*

Possible answers: Yes, there is/ No, there is not/ I cannot say for sure whether there is or not.

Part 2: Mathematical tasks

2.1. Exploring the pirate boat

First of all, we should get to know a person who reveals some truths in our stories. That person is Peter, a Little Scientist! We should accept what he tells as real in the sense of the stories. He is always curious! Let's try to help him to answer some more questions!

Let's begin with the first task, we can see here five different children/ little explorers (Lego figures) *who just found some treasure boxes in a pirate boat. We know that the boxes contain small gold objects like* ▲ *or/and* ●.

Peter has found a papyrus in which the value of each object in gold coins is indicated (picture of a papyrus which is on the desk during the task below):

The object:

▲ **is worth 3 gold coins**

● **is worth 2 gold coins**

Example (control question):

C.Q.2.1.1. *For instance, in a treasure box there are a triangle and two circles. How many gold coins is this worth?*

So, Peter afterwards found out something (additional picture which is on the desk during the task):

If the box contains the objects ▲▲●●● , then the objects in the box are worth 12 gold coins.
It is certain that this is really true.

Q.2.1.1.*The second little explorer called Maria has found a box that contains* ***exactly 2 ▲ and 3 ●****.*

Based on what he knows, can Peter conclude that Maria's box contains objects that are worth 12 gold coins?

The child chooses between three possible conclusions below:

a) YES, the objects in the box **are worth** 12 gold coins.

b) NO, the objects in the box **are not worth** 12 gold coins

c) UNSURE, he **cannot say** whether the objects in the box are worth 12 gold coins or not.

Q.2.1.2. *Another little explorer called Stelios has found a box.* ***The objects in his box are <u>not</u> worth 12 gold coins.***

Based on what he knows, can Peter conclude that Stelios box does contain exactly 2 ▲ and 3 ● ?

(Three possible answers)

Q.2.1.3. *The little explorer Helen has found* ***a box that is <u>not</u> one that contains exactly 2 ▲ and 3 ●.***

Based on what he knows, can Peter conclude that the objects in Helens box are worth 12 gold coins?

(Three possible answers)

Q.2.1.4. *The little explorer Charis has found also a box.* ***The objects in his box are worth 12 gold coins.***

Based on what he knows, can Peter conclude that Charis box contains exactly 2 ▲ and 3 ● ?

(Three possible answers)

2.2. Building dwarf houses (approx. 7')

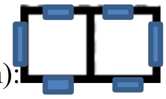
Let's continue with our second task, which is about dwarf houses!

Dwarfs build their houses so that there are rooms which all have this form: (Lego figures of Dwarfs and 3D room-cube representing a house that consists of 1 room; the dwarf houses have only one floor)

The houses always have one or more rows of rooms which are all equally long. For example one possible house that dwarfs may have is:

—> the floor plan

Whenever possible, they make one window on each exterior wall.

Example (the floor plan):  —> 6 windows

Example (control question):

C.Q.2.2.1 *For instance, a dwarf house has 2 rows with 2 rooms each. Does it have 5, 6, 7 or eight windows?*

Peter tries to find out as much as possible about houses of Dwarfs! So Peter has found out something (additional picture which is on the desk during the task below):

If a dwarf house has exactly 2 rows of 4 rooms each, then it has 12 windows.

I*t is certain that this is really true.*

Q.2.2.1. *We know that the family "bashful dwarfs" has also made a house.* ***Their house is <u>not</u> one that has exactly 2 rows with 4 rooms in each row.***

Based on what he knows, can Peter conclude that the "bashful dwarfs" house has exactly 12 windows? (Three possible answers)

Q.2.2.2. *The family "happy dwarfs" has also made a house.* ***Their house is one that has exactly 12 windows.***

Based on what he knows, can Peter conclude that the "happy dwarfs" house has exactly 2 rows of 4 rooms each?

(Three possible answers)

Q.2.2.3.*The family of "sneezy dwarfs" has also made a house.* ***Their house has exactly 2 rows of 4 rooms each.***

Based on what he knows, can Peter conclude that the "sneezy dwarfs" house has exactly 12 windows?

(Three possible answers)

Q.2.2.4. *The family of "grumpy dwarfs" has also made house.* ***Their house is <u>not</u> one that has exactly 12 windows.***

Based on what he knows, can Peter conclude that the "grumpy dwarfs" house has exactly 2 rows of 4 rooms each?

(Three possible answers)

Part 3: Everyday context

Task 3.1: Dropping the glass

We will be dealing with stories from the real world now. Peter is again the Little Scientist in our tasks. We should accept what he tells as real in the sense of the stories. He is always curious! Let's try to help him again to answer some questions!

Let's get started and think that Peter is in his house. At his home, there are some glasses in the kitchen. Sometimes it happens that a glass is dropped to the ground in the kitchen, but never outside the kitchen. Peter has found out that (additional picture which is on the desk during the task below):

"If a glass is dropped on the ground in the kitchen, then there is a sound".

It is certain that this is really true.

Q3.1.1. *Peter tries to get some rest in his bedroom.* ***There is a sound now.***

Based on what he knows, can Peter conclude that a glass was dropped on the ground?

(Three possible answers)

Q3.1.2. *Peter is in the kitchen, observing.* ***No glass is dropped on the ground now.***

Based on what he knows, can Peter conclude that there is a sound?

(Three possible answers)

Q3.1.3. *Peter tries to get some rest in his bedroom.* ***There is no sound now****.*

Based on what he knows, can Peter conclude that a glass was dropped on the ground in the kitchen?

(Three possible answers)

Q3.1.4. *Peter is in the kitchen, observing.* ***A glass is dropped on the ground in the kitchen now.***

Based on what he knows, can Peter conclude that there is a sound?

(Three possible answers)

Task 3.2. Having high fever (approx. 7')

Our last task refers to the high fever that we may have sometimes. We know that there is that child, called George who is a Peter's friend. Peter has found something about his friend (additional picture which is on the desk during the task below):

"If George has high fever, then he lies in bed".

So, remember it is certain that is really true.

Q3.2.1. *Peter is in George's room in order to see him.* ***George does <u>not</u> lie in bed.***

Based on what he knows, can Peter conclude that George has high fever?

(Three possible answers)

Q3.2.2. *Peter has been told, that* ***George has high fever.***

Based on what he knows, can Peter conclude that George lies in bed?

(Three possible answers)

Q3.2.3. *Peter is in George's room in order to see him.* **George lies in bed***.*

Based on what he knows, can Peter conclude that George has high fever?

(Three possible answers)

Q3.2.4. *Peter has been told that* ***George does <u>not</u> have high fever.***

Based on what he knows, can Peter conclude that George lies in bed?

(Three possible answers)

Probe: *How did you figure it out/ make this conclusion? What are you thinking?*

Your ideas were very interesting! We will now turn to two different tasks.

Alternatives generation tasks (approx.7')

Now there is a pirate, called Marcus, who has also gold objects in his treasure box. Remember the papyrus that Peter found:

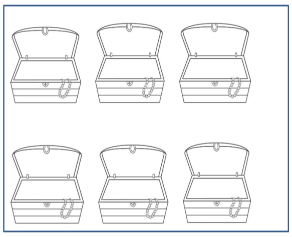
The object:

▲ is worth 3 gold coins

● is worth 2 gold coins

The objects in Marcus' treasure box are worth 18 gold coins. How many ▲ and ● could be in his box? Draw all possible combinations into treasure boxes. Draw as many different combinations as you can. The objects must be worth 18 gold coins.

The child is intended to draw the potential combinations in the provided answer sheet:



Now we look at the house of the humpy dwarf family. Remember that Dwarfs build their houses so that there are rooms which all have this form □. Whenever possible, they make one window on each exterior wall.

The house of the humpy dwarf family has 20 windows. How could the house of the humpy dwarf family look like? Draw as many possible houses in this answer sheet as you can.

The child is intended to draw the different possibilities in the above grid paper.

We just finished! I would kindly ask you not to talk to your friends about the tasks this week, so that they can still enjoy the tasks and be surprised! Thank you very much for coming this morning. Your time is very much appreciated and your ideas have been really interesting!

STUDY 1- INTERVIEW PROTOCOL

The purpose of this interview is to gather responses from students about their understanding of conditional statements in different contexts and their ability to draw logically valid conclusions from each of them. To understand the conditional statements does not demand high prior knowledge. However a basic understanding of everyday contexts and mathematical concepts may be helpful to draw valid conclusions, assuming that reasoning is not based on abstract logical rules, but on mental models. The students in this study are intended to draw their conclusions, having each time on their hands one rule, one premise, and three answer alternatives. Moreover, the interview surveys students' mathematical knowledge in terms of their ability to generate alternative causes for a given conclusion. The verbal structure of the tasks in general and of the four logical forms is parallel for all tasks in the interview. The sequence of the two contexts will be randomized among the students (two sequences: first mathematical, then everyday context and vice versa). Also the sequence of tasks within each context, and the sequence of the reasoning questions within the tasks will be randomized. Input from this study will be shared with the supervisors Prof. Dr. Ufer and Prof. Dr. Sodian. The results can provide important contribution to our understanding of the development of students' logical reasoning skills.

Part 1: Rapport Building, Background information & Instructions

1.1. General Rapport Building

Good morning, my name is Anastasia. Thank you for coming.

Please tell me your first name and what grade you are in. I am going to ask you some short questions about you.

- What do you like to do with your free time?
- What are your favourite subjects in school?

1.2. Background and demographic questions

- Could you tell me when your birthday is? So, how old are you? Which language or languages do you speak at home with your family? Do you know approximately the whole number of the books that are in your house? (visualizations of a full bookshelf, a complete bookcase and two complete bookcases are shown and students select one of the options).

1.3. Instructions

My purpose in talking with you today is to solve some tasks that I am sure that you will really enjoy. For each task, there are four questions that I would like for you to give a try and figure out an answer.

We are interested in how you think about the tasks. There are no right or wrong answers, or desirable or undesirable answers. I would like you to feel comfortable saying what you really think and how you figure out the tasks. We ensure the anonymity of your replies, meaning that your name will not be filled anywhere in this study and noone will know it. If it's okay with you, I will be video-recording our conversation, specifically only your hands, since it is hard for me to write down everything while simultaneously talking with you. Everything you say will remain confidential, meaning that only my teammates and I will be aware of your answers. Please remember that your participation is voluntary and you can stop your participation at any time without of course any negative consequences.

- Do you have any questions before we begin? Are you willing to begin together?

1.4. Example-Training of the alternative answers

We are starting with a short game:

One red and one green marble are in a small open box from the start, both visible for the child. The researcher points to the red marble when explaining the game stating that all questions will be about the red marble, because she/he likes it most.
I will show you something with these marbles and ask some questions. This card can help you regarding the possible answers (the researcher shows the relevant visualisation) The researcher holds up a red marble that the child can see and then hides it in her/his right hand.

What can you say for sure about the red marble?

Possible answers (the researcher points the three alternative answers while reading them)

It is certain that the red marble is in this hand.
It is certain that the red marble is not in this hand.
It is uncertain. One cannot say for sure whether the red marble is in this hand or not.

If the child answers "is in this hand", the researcher points the "check" of the visualization (in case that the child did not already do that).
Otherwise: You saw that I took the red marble with my hand, so you can say for sure that the red marble is in this hand. Researcher points to the "yes, certain" option (check).

Then the researcher holds up only the green marble that the child can see and then hides it in her/his right hand.Then the unused red marble remains in the container.
What can you say for sure about the red marble?

Possible answers (the researcher points the three alternative answers while reading them)

It is certain that the red marble is in this hand.
It is certain that the red marble is not in this hand.
It is uncertain. One cannot say for sure whether the red marble is in this hand or not.

If the child answers "is not in this hand", the researcher points the "cross" of the visualization (in case that the child did not already do that).
Otherwise: You saw that I took the green marble with my hand, so you can say for sure that the red marble is not in this hand. Researcher points to the "no, certain" option (cross).

Then the researcher puts both marbles into her right hand, and then shuffles them between both hands, so that one is in the right and one is in the left hand. The researcher holds closed the right hand and asks the

following question:
What can you say for sure about the red marble?

Possible answers (the researcher points the three alternative answers while reading them)à

It is certain that the red marble is on this hand.
It is certain that the red marble is not this hand.
It is uncertain. One cannot say for sure whether the red marble is in this hand or not.

If the answer is right, the researcher points the "question mark" of the visualization (in case that the child did not already do that).
If the child tries to guess which marble is hidden and gives a random wrong reply, the researcher asks whether the child can be really sure about its answer or not and gives the explicit answer. Can you be really sure about that? In some cases you cannot be really sure about the answer and this is true. The researcher points the "question mark" of the visualization.

Part 2: Mathematical tasks

Introducing Peter

First of all, we should get to know a person who reveals some truths in our stories. That person is Peter, a Little Scientist! In our stories, we assume that Peter always tells the truth. He is always curious! Let's try to help him to answer some more questions!

Mathematical tasks-Introduction

Peter is on a discovery journey with some little explorers to a small and hidden Island now which is inhabited by dwarfs

Exploring the pirate boat

Pirate boat- Instruction

Peter is walking with the little explorers and they just found some treasure boxes. We know that the boxes contain some blue  and red  diamonds.

Peter has found a message in which the value of each diamond in gold coins is indicated

(picture of the message which is shown during the task below):

 is worth 3 gold coins

 is worth 2 gold coins

Example (control question):

C.Q.2.1.1. For instance, in a treasure box there is 1 and 2 diamonds. How many gold coins is this worth?

(Procedure)

-Correct answer (7): continue with the normal process

-Wrong answer: repeat the instructions and pose the control question for a second and last time

-If the answer is wrong again, continue with the normal process without mentioning the right answer

Task 2.1

So, Peter afterwards found out something (additional picture which is shown during the task):

If the box contains exactly 2 and 3 diamonds, then the diamonds in the box are worth 12 gold coins.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q.2.1.1.This is Maria. Her box contains exactly 2 and 3 diamonds. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) The diamonds in the box are worth 12 gold coins.
b) The diamonds in the box are not worth 12 gold coins
c) He cannot be sure whether the diamonds in the box are worth 12 gold coins or not.

Q.2.1.2. This is Stelios. The diamonds in his box are not worth 12 gold coins.

Based on what he knows, what can Peter say for sure? (3 possible answers)

a) The box contains exactly 2 and 3 diamonds.

b) The box does not contain exactly 2 and 3 diamonds .

c) He cannot be sure whether the box contains exactly 2 and 3 diamonds or not.

Q.2.1.3.The little explorer Helen has found a box that is not one that contains exactly 2 and 3 diamonds. Based on what he knows, what can Peter say for sure?

(3 possible answers)

Q.2.1.4. This is Charis. The diamonds in his box are worth 12 gold coins. Based on what he knows, what can Peter say for sure?

(3 possible answers)

Task 2.2.

Peter has just found out something about the treasure boxes. Let's see what he found out:

If a box contains 4  and 1  diamonds, then the diamonds in the box are worth 14 gold coins.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q.2.2.1. This is Charis. The diamonds in his second box are not worth 14 gold coins.

Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) The box contains exactly 4  and 1  diamonds .

b) The box does not contain exactly 4  and  diamonds .

c) He cannot be sure whether the box contains exactly 4  and 1  diamonds or not.

Q.2.2.2. This is Stelios. The diamonds in his second box are worth 14 gold coins. Based on what he knows, what can Peter say for sure?

(3 possible answers)

Q.2.2.3. This is Helen. Her second box contains exactly 4  and 1  diamonds. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

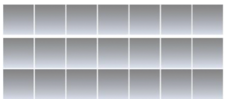
a) The diamonds in the box are worth 14 gold coins.
b) The diamonds in the box are not worth 14 gold coins
c) He cannot be sure whether the diamonds in the box are worth 14 gold coins or not.

Q.2.2.4. This is Maria. Her second box does not contain exactly 4  and 1  diamond. Based on what he knows, what can Peter say for sure?
(3 possible answers)

2.3. Exploring dwarf houses I
Exploring dwarf houses-Instruction
Peter is at this island and observes dwarfs' houses.

Dwarfs build their houses so that there are rooms which all have this form:
The houses always have one or more rows of rooms which are all equally long. For example dwarfs may have a house like this:

Example (the floor plan):



They also make one window on each exterior wall of each room.

Example  6 windows

(control question): C.Q.2.3.1 A dwarf house has 2 rows with 2 rooms each. How many windows does it have?

(Procedure)

- Correct answer (8): continue with the normal process

-Wrong answer: repeat the instructions and pose the control question for a second and last time.

-If the answer is wrong again, continue with the normal process without mentioning the right answer
Peter tries to find out as much as possible about houses of Dwarfs! So Peter has found out something (additional picture which is shown during the task below):

If a dwarf house has exactly 2 rows of 5 rooms each, then it has 14 windows.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q.2.3.1. This is one of the "bashful dwarfs". Their house is not one that has exactly 2

rows with 5 rooms in each row.

Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) The house has 14 windows.
b) The house does not have 14 windows.
c) He cannot be sure whether the house has 14 windows or not.

Q.2.3.2. This is one of the "happy dwarfs". Their house is one that has exactly 14 windows. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) The house has exactly 2 rows with 5 rooms in each row.
b) The house does not have exactly 2 rows with 5 rooms in each row.
c) He cannot be sure whether the house 2 rows with 5 rooms in each row or not.

Q.2.3.3.This is one of the "sneezy dwarfs". Their house has exactly 2 rows of 5 rooms each.Based on what he knows, what can Peter say for sure?

(Three possible answers)

Q.2.3.4. This is one of the "grumpy dwarfs". Their house is not one that has exactly 14 windows. Based on what he knows, what can Peter say for sure?

(Three possible answers)

Task 2.4. Exploring dwarf houses II

Peter is at this island and observes dwarfs' houses.

Control question: There is a dwarfs' house which has 2 rows of 3 rooms each, how many rooms does it have?

(Procedure)

-Correct answer (6): continue with the normal process

- Wrong answer: repeat the instructions and pose the control question for a second and last time

-If the answer is wrong again, continue with the normal process without mentioning the right answer

Peter has found out a second rule about the dwarfs houses in that village:

If a dwarf house has exactly 2 rows of 6 rooms each, then it has 12 rooms.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q.2.4.1. This is one of the “sleepy dwarfs”. Their house has exactly 2 rows of 6 rooms each. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) The house has exactly 12 rooms.
b) The house does not have exactly 12 rooms.
c) He cannot be sure whether the house has exactly 12 rooms or not.

Q.2.4.2. This is one of the “wise dwarfs”.Their house is not one that has exactly 2 rows with 6 rooms in each row. Based on what he knows, what can Peter say for sure?

(3 possible answers)
Q.2.4.3. This is one of the “dopey dwarfs”. Their house is not one that has exactly 12 rooms. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) The house has exactly 2 rows with 6 rooms in each row.
b) The house does not have exactly 2 rows with 6 rooms in each row.
c) He cannot be sure whether the house has exactly 2 rows with 6 rooms in each row
or not.

Q.2.4.4. This is one of the “kind dwarfs”. Their house has exactly 12 rooms. Based on what he knows, what can Peter say for sure?
(3 possible answers)

Part 3: Everyday context

Peter is in his hometown, spending time with relatives and friends.

Task 3.1: Dropping the glass

Peter is in his house now. At his home, there are some glasses in the kitchen. Sometimes it happens that a glass is dropped to the ground in the kitchen. Peter has found out that (additional picture which is shown during the task below):

“If a glass is dropped on the ground in the kitchen, then there is a sound”.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q3.1.1. Peter tries to get some rest in his bedroom. There is a sound now. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) A glass is dropped on the ground in the kitchen.

b) No glass is dropped on the ground in the kitchen
c) He cannot be sure whether a glass is dropped on the ground in the kitchen or not.

Q3.1.2. Peter is in the kitchen, observing. No glass is dropped on the ground in the kitchen now. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) There is a sound now.
b) There is no sound now.
c) He cannot be sure whether there is a sound or not.

Q3.1.3. Peter tries to get some rest in his bedroom. There is no sound now. Based on what he knows, what can Peter say for sure?

(Three possible answers)

Q3.1.4. Peter is in the kitchen, observing. A glass is dropped on the ground in the kitchen now. Based on what he knows, what can Peter say for sure?

(Three possible answers)

Task 3.2

Peter sometimes helps his parents during the cooking. For example, he cuts fruits with a knife. Peter has found out something (additional picture which is shown during the task below):

If someone's finger is cut deeply while cooking, then it bleeds.

It is certain that this is really true.

I will make some questions; please think very well before answering them.

Q3.2.1 Peter visited his friend, George. George's finger is not bleeding. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) George's finger has just been cut deeply while cooking.
b) George's finger has not just been cut deeply while cooking.
c) He cannot be sure whether George's finger has just been cut deeply while cooking or not.

Q3.2.2 Peter has been told that Cathrin's finger has just been cut deeply while cooking. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) Cathrin's finger is bleeding.
b) Cathrin's finger is not bleeding.
c) He cannot be sure whether Cathrin's finger is bleeding or not.

Q3.2.3 Peter visited his friend, Despina. Despina's finger is bleeding. Based on what he knows, what can Peter say for sure?

(3 possible answers)

Q3.2.4 Peter has been told that Kostas' finger has not just been cut deeply while cooking. Based on what he knows, what can Peter say for sure?

(3 possible answers) Task 3.3

Peter really likes swimming pools as he loves swimming. So, Peter has found out something (additional picture which is shown during the task below):

If someone jumps into a swimming pool, then she or he gets wet.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q3.3.1 A few days later, Peter just arrived at the swimming pool. His friend Vivi is wet now. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) Vivi has jumped into the swimming pool.
b) Vivi has not jumped into the swimming pool.
c) He cannot be sure whether Vivi has jumped into the swimming pool or not.

Q3.3.2 Peter just arrived at the swimming pool. His friend Sakis is not wet now. Based on what he knows, what can Peter say for sure?

(3 possible answers)

Q3.3.3 Peter has been told that his friend, Mary, has just jumped into the swimming pool. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) Mary is wet now.
b) Mary is not wet now.
c) He cannot be sure whether Mary is wet or not.

Q3.3.4 Peter has been told that his friend, Dimitris, has not just jumped into the swimming pool. Based on what he knows, what can Peter say for sure?

(3 possible answers) Task 3.4

Peter has found out that: (additional picture which is shown during the task below):

If someone breaks his arm, then she or he hurts.

It is certain that this is really true.

I will ask you some questions about this; please think very well before answering them.

Q3.4.1 Peter has been told that Chris has not just broken his arm. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) Chris hurts now.
b) Chris does not hurt now.
c) He cannot be sure whether Chris hurts now or not.

Q3.4.2 Peter visited his friend, Ntina in order to see her. Ntina hurts now. Based on what he knows, what can Peter say for sure?

The researcher points the three alternative answers while reading them:

a) Ntina has just broken her arm.
b) Ntina has not just broken her arm.
c) He cannot be sure whether Ntina has just broken her arm or not.

Q3.4.3 Peter visited his friend, Stella. Stella does not hurt now.Based on what he knows, what can Peter say for sure?
(3 possible answers)

Q3.4.4 Peter has been told that his friend, Nektaria, has just broken her arm. Based on what he knows, what can Peter say for sure?

(3 possible answers)

Probe: How did you figure it out/ make this conclusion? What are you thinking? Your ideas were very interesting! We will now turn to some different tasks.

Part 4: Alternatives generation tasks

Do you remember the stories we have heard from Peter's expedition to the dwarfs? I have some more questions about what Peter found out about these dwarfs

4.1 Alternatives generation – treasure task

There is a little explorer, called Marcus, who has also found a treasure box at the dwarfs'

village. We already know that the boxes contain some blue  and red  gems. Remember that Peter has found a message in which the value of each gem in gold coins is indicated (picture of the message which is shown during the task below):

the  is worth 3 gold coins

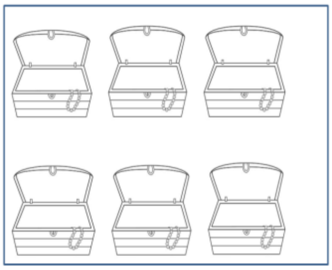
the  is worth 2 gold coins

The diamonds in Marcus' treasure box are worth 18 gold coins.

Remember what Peter found out before. If the box contains exactly 2  and 3  diamonds, then the diamonds in the box are worth 12 gold coins.

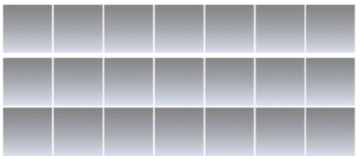
Q4.1.1 How many blue  and red  diamonds could be in his box? Create all possible combinations into treasure boxes. Draw as many different combinations as you can. The diamonds must be worth 18 gold coins.

The researcher points the respective part on the screen which should be filled in. The child is intended to draw the potential combinations in the empty treasure boxes on the iPad screen. The researcher repeats once during the process (for all children) that diamonds in Marcus‘ box are worth 18 gold coins.No feedback/hints concerning the whole task.

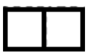


4.2 Alternatives generation – Windows task

Remember that dwarfs build their houses so that there are rooms which all have this form: The houses always have one or more rows of rooms which are all equally long. For example dwarfs may have a house like this:



We already know that they also make one window on each exterior wall of each room. Example (the floor plan):

 6 windows

Remember what Peter found out before.If a dwarf house has exactly 2 rows of 5 rooms each, then it has 14 windows

Now we look at the house of the ‘moody’ dwarf family. The house of the ‘moody’ dwarf family has 20 windows.

Q4.2.1 How could the house of the ’‘moody” dwarf family look like? Draw as many possible houses in this box/ frame as you can.

The researcher points the respective part on the screen that should be filled in. The child is intended to draw the different possibilities on the screen by using the iPad pen.

4.3 Alternatives generation – Room number task

Remember what else Peter found out before. If a dwarf house has exactly 2 rows of 6 rooms each, then it has 12 rooms.

Q4.3.1 How could a dwarf house with 12 rooms look like? Draw as many possible houses in this box as you can!

The researcher points the respective part on the screen that should be filled in. The child is intended to draw the different possibilities on the screen by using the iPad pen

4.4. Alternatives generation – Everyday contexts

Now, we will talk about what happens in Peter's daily life again. I would like to ask you to think about how some things can happen in the real world. Please find as many possibilities as you can.

- Remember what Peter found out before. If a glass is dropped on the ground in the kitchen, then there is a sound. Peter is at home and hears a sound in his kitchen. Find as many reasons why a sound in the kitchen may occur, as you can.

- Remember what else Peter found out before; if someone's finger is cut deeply while cooking, then it bleeds.Peter meets Kostas and sees that her finger is bleeding. Find as many reasons for someone's finger bleeding, as you can.

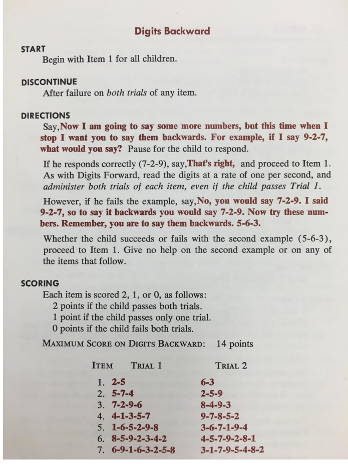
- Peter found out also before that if someone jumps into a swimming pool, then she or he gets wet. Peter goes to the swimming pool and sees that Chris is completely wet. Find as many ways for someone getting wet, as you can.

- Finally he found out that if someone breaks his arm, then she or he hurts. Peter meets Sakis and hears that he has a pain in his arm. Find as many reasons for someone feeling pain in his/her arm, as you can.

4.5. Fill in the blanks appropriately, so that the calculations are correct:

3 x 8 = □	8 x 5 = □ x 10
23 = 14 + □	□ x 6= □ x 4
3 x 12 = □	32 = □ x 8
6 = □ x □	6 x 7 = □
7 x □ = 28	14 = □ + □

4.6. Working memory test



Digits Backward

START
Begin with Item 1 for all children.

DISCONTINUE
After failure on *both trials* of any item.

DIRECTIONS
Say, **Now I am going to say some more numbers, but this time when I stop I want you to say them backwards. For example, if I say 9-2-7, what would you say?** Pause for the child to respond.

If he responds correctly (7-2-9), say, **That's right,** and proceed to Item 1. As with Digits Forward, read the digits at a rate of one per second, and *administer both trials of each item, even if the child passes Trial 1.*

However, if he fails the example, say, **No, you would say 7-2-9. I said 9-2-7, so to say it backwards you would say 7-2-9. Now try these numbers. Remember, you are to say them backwards. 5-6-3.**

Whether the child succeeds or fails with the second example (5-6-3), proceed to Item 1. Give no help on the second example or on any of the items that follow.

SCORING
Each item is scored 2, 1, or 0, as follows:
2 points if the child passes both trials.
1 point if the child passes only one trial.
0 points if the child fails both trials.

MAXIMUM SCORE ON DIGITS BACKWARD: 14 points

Item	Trial 1	Trial 2
1.	2-5	6-3
2.	5-7-4	2-5-9
3.	7-2-9-6	8-4-9-3
4.	4-1-3-5-7	9-7-8-5-2
5.	1-6-5-2-9-8	3-6-7-1-9-4
6.	8-5-9-2-3-4-2	4-5-7-9-2-8-1
7.	6-9-1-6-3-2-5-8	3-1-7-9-5-4-8-2

5. Final information

We just finished! I would kindly ask you not to talk to your friends about the tasks this week, so that they can still enjoy the tasks and be surprised! Thank you very much for coming this morning. Your time is very much appreciated and your ideas have been really interesting!

STUDY 2

CONDITION A

Students' Booklets (training tasks)

Imagine that you have gone to a planet where everything happens very differently to here on Earth.

Answer the following questions.

On this planet, if you clean a sweater with ketchup, then the sweater will become clean.

Can you imagine other ways of cleaning a sweater on this planet?

Give as many responses as you can.

On this planet, if you jump into a lake with freezing water, then you will warm up.

Can you imagine other ways of warming up on this planet?

Give as many responses as you can.

On this planet, if you eat French fries, then this is good for your health.

Can you imagine other things that you can do that are good for your health on this planet?

Give as many responses as you can.

Researcher's Script (Feedback part)

Very well; so time is over; now let's discuss together about these tasks. Please leave your pencils and pens down and do not fill anything else in your booklets. Please turn the page and look at the slides here (the researcher points out the board while it is indicated 'Listen to Anastasia now' on the next page of students' booklets).

Often when approaching a problem, it is good not to use the first and only solution that comes into your mind. Often it is better to wait a bit and think. Could it be completely different? Are there other possibilities? This often helps us to avoid false answers and false decisions.

As for the task about imagining other ways of cleaning a sweater on that planet here are some possible answers (these answers are depicted through illustrations on ppt slides and the researcher read them out) *that someone could give. Probably you already wrote one or more similar answers in your booklet or some different ones and that would be great! The more different ones you created, the better!*

As for task about imagining other ways of warming up on this planet, these are some possible answers (these answers are depicted through illustrations on ppt slides and the researcher read them out) that someone could give. Probably you already wrote one or more similar answers in your booklet or some different ones and that would be great! The more different ones you created, the better!

Regarding task about imagining other things that you can do that are good for your health on that planet, here are some possible answers that someone could give (these answers are depicted through illustrations on ppt slides and the researcher read them out). Probably you already wrote one or more similar answers in your booklet or some different ones and that would be also great! The more different ones you created, the better!

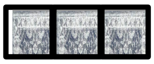
As we see there are so many and different solutions for every single question or task and all of them can be right. It is really important to think if there any other possibilities for each task because this makes us being creative! Being creative is not only helpful for our everyday life but also for mathematics and other school subjects! In the next tasks try to do the same: Be creative, think about different possibilities that did not come to your mind in the first moment.

CONDITION B

Students' Booklet (training tasks)

A friend gives you a riddle: He says that yesterday in the park, there were some dogs and some children, but no other animals. Altogether, there were 600 legs in the park. How many children and dogs could have been in the park, which have together 600 legs?

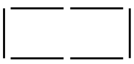
Here we know that in Japan some families use carpets to cover the whole floor of their living-room. The carpets always have this form  . For example this is one possible living- room covered fully by 3 carpets:



If a family used 24 carpets, how could their living-room be like?

Remember that they used all of those carpets to cover their living room without any gaps or overlaps. Find as many and different solutions as you can!

A farmer has 40 pieces of fencing which are equally long. He wants to fence a field to keep his sheep in by using all of those fencing pieces. He can do this on a large grass plain, but he wants the area to be rectangular. For example, one possible rectangular field in which 6 pieces of fencing are used would be the following:



How long and wide could his field be?

Find as many and different solutions as you can.

Researcher's Script (Feedback part)

Very well; so time is over; now let's discuss together about these tasks. Please leave your pencils and pens down and do not fill anything else in your booklets.

Please turn the page and look at the slides here (the researcher points out the board while it is indicated 'Listen to Anastasia now' on the next page of students' booklets).

Often when approaching a problem, it is good not to use the first and only solution that comes into your mind. Often it is better to wait a bit and think. Could it be completely different? Are there other possibilities? This often helps us to avoid false answers and false decisions.

Regarding the task about how many children and dogs could have been in the park, which have together 600 leg, these are some possible answers (these solutions are depicted through illustrations on ppt slides and the researcher read them out*) that someone could give. Probably you already wrote one or more similar answers in your booklet or some different ones and that would be also great! The more different ones you created, the better!*

As for the task about the family owing 24 carpets and how their living room could be like, these are some possible solutions (these solutions are depicted through illustrations on ppt slides and the researcher read them out) *that someone could give. Probably you already wrote one or more similar answers in your booklet or some different ones and that would be great! The more different ones you created, the better!*

Regarding the task about long and wide could be the farmer's field, these are some possible solutions (these solutions are depicted through illustrations on ppt slides and the researcher read them out) *that someone could give. Probably you already drew some similar solutions in your booklet or some different ones and that would be also great! The more different ones you created, the better!*

As we see there are so many and different solutions for every single question or task and all of them can be right. It is really important to think if there any other possibilities for each task because this makes us being creative! Being creative is not only helpful for our everyday life but also for mathematics and other school subjects! In the next tasks try to do the same: Be creative, think about different possibilities that did not come to your mind in the first moment.

CONDITION C

Students' Booklet (mathematical single solution tasks)

Read the instructions very carefully and complete your answers in the lines provided. Remember to solve correctly as many tasks you can. It is not mandatory to solve all of the tasks below if the time does not allow you to do so. Try to solve correctly the ones you fill in though.

A friend gives you a riddle: He says that yesterday in the park, there were some dogs and some children, but no other animals. Now you will investigate how many dogs, children or how many legs they have altogether in each case.

Remember that each dog has 4 legs and each child has 2 legs.

a) If there are only 4 dogs and 3 children, how many legs do they have altogether?

b) If there are only 12 dogs and 20 children, how many legs do they have altogether?

c) If there are only 9 dogs and 9 children, how many legs do they have altogether?

d) If there are only 17 dogs and some children and altogether have 72 legs, how many children are there?

e) If there are only 12 children and some dogs and altogether have 36 legs, how many dogs are there?

f) If there are only 5 children and 8 dogs, how many legs do they have altogether?

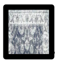
g) If there are only 17 children and 2 dogs, how many legs do they have altogether?

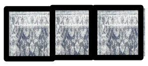
h) If there are only 17 children and some dogs and altogether have 54 legs, how many dogs are there?

i) If there are only 1 child and some dogs and altogether have 20 legs, how many dogs are there?

j) If there are only 4 dogs and some children and altogether have 20 legs, how many children are there?

Read the instructions very carefully and write down your answers in the lines provided.

Here we know that in Japan some families use carpets to cover the whole floor of their living-room. The carpets always have this form:  For example this is one possible living- room covered fully by 3 carpets:



There is a family in Japan and now you will investigate more about the floor of their living-room as well as the carpets they need to cover the whole floor of this room in each case!

a) If the family used 5 carpets and 13 carpets along each wall of their rectangular living-room respectively, what would be the overall number of carpets they should use in order to cover the whole living-room floor without any gaps or overlaps?

b) The family used an overall number of 48 carpets in order to cover their whole rectangular living-room floor. Along one of the room walls there were 8 carpets. How many carpets should be along the other room wall so that there are no gaps or overlaps on the floor?

c) If the family used 8 carpets and 14 carpets along each wall of their rectangular living-room respectively, what would be the overall number of carpets they should use in order to cover the whole living-room floor without any gaps or overlaps?

d) The family used an overall number of 72 carpets in order to cover their whole rectangular living-room room. Along one of the room walls there were 9 carpets. How many carpets should be along the other room wall so that there are no gaps or overlaps on the floor?

e) If the family used 3 carpets and 12 carpets along each wall of their rectangular living-room respectively, what would be the overall number of carpets they should use in order to cover the whole living-room floor without any gaps or overlaps?

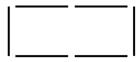
f) The family used an overall number of 64 carpets in order to cover their whole rectangular living-room floor. Along one of the room walls there were 8 carpets. How many carpets should be along the other room wall so that there are no gaps or overlaps on the floor?

g) If the family used 6 carpets and 12 carpets along each wall of their rectangular living-room respectively, what would be the overall number of carpets they should use in order to cover the whole living-room floor without any gaps or overlaps?

h) The family used an overall number of 24 carpets in order to cover their whole rectangular living-room floor. Along one of the room walls there were 6 carpets. How many carpets should be along the other room wall so that there are no gaps or overlaps on the floor?

i) If the family used 9 carpets and 7 carpets along each wall of their rectangular living-room respectively, what would be the overall number of carpets they should use in order to cover the whole living-room floor without any gaps or overlaps?

j) The family used an overall number of 44 carpets in order to cover their whole rectangular living-room floor. Along one of the room walls there were 4 carpets. How many carpets should be along the other room wall so that there are no gaps or overlaps on the floor?

Read the instructions very carefully and write down your answers in the lines provided. A farmer has some pieces of fencing which are equally long. He wants to fence a field to keep his sheep in by using all of those fencing pieces. He can do this on a large grass plain, but he wants the area to be rectangular. For example, one possible rectangular field in which 6 pieces of fencing are used would be the following: 

Now you will investigate more about the sides of the area he is going to fence as well as how much fencing he would need in each case!

a) The farmer used 10 pieces of fencing to fence one side of the rectangular area of land and 15 pieces for the other side. How many pieces of fencing would the farmer need in order to fence the whole area of his farm?

b) The farmer used exactly 38 fencing pieces to fence a rectangular area of land while he used 12 pieces for one side. How many fencing pieces would the farmer need in order to fence the other side of this rectangular area?

c) The farmer used 10 pieces of fencing to fence one side of the square area of land. How many pieces of fencing would the farmer need in order to fence the whole area of his farm?

d) The farmer used exactly 34 fencing pieces to fence a rectangular area of land while he used 12 pieces for one side. How many fencing pieces would the farmer need in order to fence the other side of this rectangular area?

e) The farmer used 7 pieces of fencing to fence one side of the rectangular area of land and 8 pieces for the other side. How many pieces of fencing would the farmer need in order to fence the whole area of his farm?

f) The farmer used exactly 24 fencing pieces to fence a rectangular area of land while he used 7 pieces for one side. How many fencing pieces would the farmer need in order to fence the other side of this rectangular area?

g) The farmer used 11 pieces of fencing to fence one side of the rectangular area of land and 21 pieces for the other side. How many pieces of fencing would the farmer need in order to fence the whole area of his farm?

h) The farmer used exactly 46 fencing pieces to fence a rectangular area of land while he used 12 pieces for one side. How many fencing pieces would the farmer need in order to fence the other side of this rectangular area?

i) The farmer used 16 pieces of fencing to fence one side of the square area of land. How many pieces of fencing would the farmer need in order to fence the whole area of his farm?

j) The farmer used exactly 42 fencing pieces to fence a rectangular area of land while he used 11 pieces for one side. How many fencing pieces would the farmer need in order to fence the other side of this rectangular area?

Researcher's Script (Feedback part)

Very well; so time is over; now let's discuss together about these tasks. Please leave your pencils and pens down and do not fill or change anything else in your booklets. Please turn the page and look at the slides here (the researcher points out the board while it is indicated 'Listen to Anastasia now' on the next page of students' booklets).

As for the tasks you just dealt with, these are the correct answers (these solutions are depicted through illustrations on ppt slides and the researcher read them out). *You can check your answers but please remember you should not make any changes in your booklet.*

Probably you already found some correct answers in your booklet or no. It is important in these tasks to be really sure that you got the right operation, and calculated it correctly. Checking your solution again is also helpful for the subsequent tasks.

CONDITION D

Students' Booklet (text reading)

Oscar Wilde

The Devoted Friend

Little Hans had a great many friends, but the most devoted friend of all was big Hugh the Miller. Indeed, so devoted was the rich Miller to little Hans, that he would never go by his garden without leaning over the wall and plucking a large nosegay, or a handful of sweet herbs, or filling his pockets with plums and cherries if it was the fruit season.

«Real friends should have everything in common,» the Miller used to say, and little Hans nodded and smiled, and felt very proud of having a friend with such noble ideas.

Sometimes, indeed, the neighbours thought it strange that the rich Miller never gave little Hans anything in return, though he had a hundred sacks of flour stored away in his mill, and six milch cows, and a large flock of woolly sheep; but Hans never troubled his head about these things, and nothing gave him greater pleasure than to listen to all the wonderful things the Miller used to say about the unselfishness of true friendship.

So little Hans worked away in his garden.(...) Early the next morning the Miller came down to get the money for his sack of flour, but little Hans was so tired that he was still in bed.

«Upon my word,» said the Miller, «you are very lazy. Really, considering that I am going to give you my wheelbarrow, I think you might work harder (...)Anybody can say charming things and try to please and to flatter, but a true friend always says unpleasant things, and does not mind giving pain. Indeed, if he is a really true friend he prefers it, for he knows that then he is doing good.»

«I am very sorry,» said little Hans (...),«but I was so tired that I thought I would lie in bed for a little time, and listen to the birds singing (...)

«Well, I am glad of that,» said the Miller, clapping little Hans on the back, «for I want you to come up to the mill as soon as you are dressed, and mend my barn-roof for me.»

Poor little Hans was very anxious to go and work in his garden, for his flowers had not been watered for two days, but he did not like to refuse the Miller, as he was such a good friend to him.

«Do you think it would be unfriendly of me if I said I was busy?» he inquired in a shy and timid voice.

«Well, really,» answered the Miller, «I do not think it is much to ask of you, considering that I am going to give you my wheelbarrow; but of course if you refuse I will go and do it myself.»

«Oh! on no account,» cried little Hans; and he jumped out of bed, and dressed himself and went up to the barn.

«He worked there all day long, till sunset, and at sunset the Miller came to see how he was getting on. «Have you mended the hole in the roof yet, little Hans?» cried the Miller in a cheery voice.

«It is quite mended,» answered little Hans, coming down the ladder.

«Ah!» said the Miller, «there is no work so delightful as the work one does for others.»

«It is certainly a great privilege to hear you talk,» answered little Hans, sitting down and wiping his forehead, «a very great privilege. But I am afraid I shall never have such beautiful ideas as you have.»

«Oh! they will come to you,» said the Miller, «but you must take more pains. At present you have only the practice of friendship; some day you will have the theory also.»

«Do you really think I shall?» asked little Hans.

«I have no doubt of it,» answered the Miller; «but now that you have mended the roof you had better go home and rest, for I want you to drive my sheep to the mountain tomorrow». (...)

So little Hans worked away for the Miller, and the Miller said all kinds of beautiful things about friendship, which Hans took down in a note-book, and used to read over at night, for he was a very good scholar.

Now it happened that one evening little Hans was sitting by his fireside when a loud rap came at the door. (...)

«There stood the Miller with a lantern in one hand and a big stick in the other.

«Dear little Hans,» cried the Miller, «I am in great trouble. My little boy has fallen off a ladder and hurt himself and I am going for the Doctor. But he lives so far away, and it is such a bad night, that it has just occurred to me that it would be much better if you went instead of me. You know I am going to give you my wheelbarrow, and so it is only fair that you should do something for me in return.»

«Certainly,» cried little Hans, «I take it quite as a compliment your coming to me, and I will start off at once. But you must lend me your lantern, as the night is so dark that I am afraid I might fall into the ditch.»

«I am very sorry,» answered the Miller, «but it is my new lantern, and it would be a great loss to me if anything happened to it.»

«Well, never mind, I will do without it» cried little Hans. (...)

Please answer the questions below. It is not mandatory to reply to all of them if the time does not allow you to do so.

1. What feelings does the title cause to you before and after reading the text?

2. What do you believe about the character of rich Miller? Keep in mind his general attitude towards little Hans.

3. How do you think a really devoted friend would behave?

Researcher's Script (Feedback part)

Very well; so time is over; Please leave your pencils and pens down and do not fill or change anything else in your booklets. Please turn the page and look at the slides here (the researcher points out the board while it is indicated 'Listen to Anastasia now' on the next page of students' booklets). *Now let's discuss together about this text (the researcher discusses briefly the 3 questions with the students).*