QUEEN'S UNIVERSITY BELFAST

The link between deductive reasoning and maths

Morsanyi, K., McCormack, T., & O'Mahoney, E. (2017). The link between deductive reasoning and maths. *Thinking & Reasoning*. Advance online publication. https://doi.org/10.1080/13546783.2017.1384760

Published in:
Thinking & Reasoning

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2017 Informa UK Limited, trading as Taylor & Francis Group.
This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Open Access
This research has been made openly available by Queen's academics and its Open Research team. We would love to hear how access to this research benefits you. – Share your feedback with us: http://go.qub.ac.uk/oa-feedback

Download date:06. May. 2024

Title: The link between deductive reasoning and mathematics

Running head: *Deductive reasoning and maths*

Kinga Morsanyi, Teresa McCormack & Eileen O'Mahony

School of Psychology, Queen's University Belfast

Corresponding author: Kinga Morsanyi

School of Psychology, Queen's University Belfast, Belfast, BT7 1NN, Northern Ireland, UK
Telephone: +44 (0)28 9097 4326; Fax: +44 (0) 28 9097 4524; email: k.morsanyi@qub.ac.uk

ORCID ID: 0000-0002-7550-1454

Please cite paper as:
Morsanyi, K., McCormack, T. & O'Mahony, E. (2018). The link between deductive reasoning and mathematics. *Thinking & Reasoning*. *(in press)* DOI:10.1080/13546783.2017.1384760

Recent studies have shown that deductive reasoning (including transitive and conditional inferences) are related to mathematical abilities. Nevertheless, so far the links between mathematical abilities and these two forms of deductive inference have not been investigated in a single study. It is also unclear whether these inference forms are related to both basic maths skills and mathematical reasoning, and whether these relationships still hold if the effects of fluid intelligence are controlled. We conducted a study with eighty-seven adult participants. The results showed that transitive reasoning skills were related to performance on a number line task, and conditional inferences were related to arithmetic skills. Additionally, both types of deductive inference were related to mathematical reasoning skills, although transitive and conditional reasoning ability were unrelated. Our results also highlighted the important role that ordering abilities play in mathematical reasoning, extending findings regarding the role of ordering abilities in basic maths skills. These results have implications for the theories of mathematical and deductive reasoning, and they could inspire the development of novel educational interventions.

Keywords: arithmetic skills; conditional reasoning; deductive reasoning; number line task; order processing; transitive inference

Theorists, including Piaget (1952) and Russell (1919), have long considered that mathematics and logical reasoning skills are closely related. Thus, it might come as a surprise that relatively few studies have investigated the links between the two skills, and even fewer studies have tried to explain the nature of this link. In the following, we will review this literature, focusing on two particular forms of logical/deductive inference: transitive and conditional reasoning, both of which were found to be related to mathematical skills by previous studies. We will also seek to answer the question of what types of mathematical skills are linked to each form of deductive inference, and why this might be the case. Then we will discuss the implications of these findings.

Transitive inferences involve comparisons between items on the basis of a certain property. For example, *if A is darker than B, and C is darker than A*, then we can infer that *C is darker than B*. Thus, transitive reasoning requires representing the relative position of items along a single continuum. Although information about item positions could be represented either verbally or spatially, there is behavioural evidence that transitive inferences utilize spatial representations (e.g., Goodwin & Johnson-Laird, 2005, 2008; Prado, Van der Henst & Noveck, 2008; Vandierendonck & De Vooght, 1997). Neuroimaging studies have also shown associations between the activation of the spatial regions of the parietal cortex and transitive reasoning (see Prado, Chadha & Booth, 2011 for a meta-analysis). The proposed mental representations that underlie transitive inferences are remarkably similar to the concept of the "mental number line" (Moyer & Landauer, 1967; Restle, 1970), a spatial representation of the number sequence. Performance on number line tasks is closely related to arithmetic skills in children (e.g., Booth & Siegler, 2006, 2008; Link, Nuerk & Moeller, 2014; Siegler & Booth, 2004), suggesting that this type of representation is important for arithmetic skills. Thus, a plausible hypothesis is that transitive inferences and mathematics will be linked because they utilize similar spatial representations of item positions.

So far the link between transitive reasoning and maths skills has only been investigated in the case of children (Handley, Capon, Beveridge & Dennis, 2004; Morsanyi, Devine, Nobes & Szucs, 2013) and adolescents (Morsanyi, Kahl & Rooney, 2017). These studies have found a relationship between mathematical skills and transitive reasoning ability in typical populations (Handley et al., 2004; Morsanyi et al., 2017), as well as when comparing groups of children with exceptionally low, average and high mathematical ability (Morsanyi et al., 2013). In the study by Morsanyi et al. (2013) the children with low and average maths ability were matched on IQ, verbal working memory and reading skills. Thus, the link between transitive reasoning and maths skills could not be attributed to any of these factors. Nevertheless, none of the existing studies investigated the question of exactly what types of mathematical ability are linked to transitive reasoning skills. In fact, Handley et al. (2004) used a reasoning measure that combined transitive and conditional inference problems, and they did not consider the two reasoning measures separately.

Conditional inferences require the ability to reason on the basis of “if p then q”- type statements. Examples of the basic inference types, modus ponens (MP), modus tollens (MT), affirmation of the consequent (AC), and denial of the antecedent (DA), are presented in Table 1. Conditional reasoning skills develop slowly (e.g., De Neys & Everaerts, 2008; Gauffroy & Barrouillet, 2009; Klaczynski, Schuneman & Daniel, 2004; Markovits & Barrouillet, 2002), reasoning performance depends strongly on problem content, and even adults perform relatively poorly on these problems.

One factor that has been known to influence the ability to reason about conditionals with everyday content is the availability of counterexamples (e.g., Cummins, 1995; De Neys, Schaeken & d'Ydewalle, 2005; Quinn & Markovits, 1998; Thompson, 1994). For example, consider the following problem: *If the radio is turned on, then you will hear music. The radio is not turned on. Is it necessary that you will not hear music?* Answering yes to this question would be an example of an invalid DA inference. However, when we consider this problem, we might find it easy to think of situations where the conclusion would not be true (e.g., we can hear music, because there is a music programme on TV, or our neighbour is listening to loud music). Thus, counterexamples can help us to reject invalid conclusions. However, when it is hard to think of counterexamples (i.e., when the availability of counterexamples is low; see examples in Table 1), people often accept the invalid AC and DA inferences. In the case of the valid MP and MT inferences, high availability of counterexamples can lead to the opposite effect: an incorrect rejection of the conclusions.

In a number of studies, Inglis and colleagues have investigated the relationship between conditional reasoning ability and maths skills (Attridge & Inglis, 2013; Inglis & Simpson, 2008; 2009). In these studies, the reasoning skills of students studying post-compulsory mathematics were compared to the reasoning skills of arts students. Inglis and Simpson (2008) found better conditional reasoning performance among mathematics students than arts students, and Inglis and Simpson (2009) replicated these findings in a new sample of maths and arts students, who were matched on their level of intelligence. In another study, Attridge and Inglis (2013) investigated the changes in the reasoning skills of maths and arts students during the first year of their post-compulsory studies in maths/arts. Attridge and Inglis (2013) found that the conditional reasoning performance of arts students did not change between the start and the end of the academic year. By contrast, maths students improved in their ability to accept the valid MP inference, and reject the invalid AC and DA inferences. However, their reasoning performance declined in one respect. They were more likely to incorrectly reject MT inferences at the end of their first year of post-compulsory maths education. Although this finding regarding the MT inferences might seem counterintuitive, it is in line with Newstead, Handley, Harley, Wright and Farelly (2004) who reported that, whereas correct reasoning about the MP, DA and AC inferences was positively related to intelligence in an adult sample, there was a non-significant trend for a negative relationship between correct MT inferences and intelligence. Additionally, correct MT inferences were negatively related to correct reasoning about the DA and AC inferences. Thus, it is possible that the positive relationship between maths and conditional reasoning will be restricted to the MP, DA and AC inferences, whereas there might be a negative relationship (or no relationship) between maths and the MT inferences.

Insert Table 1 about here.

The Current Study

Although existing studies have provided some evidence of a link between both transitive and conditional inferences and maths abilities, there are a number of important unanswered questions. First, we do not know whether transitive reasoning and conditional reasoning are differentially related to maths skills, or whether both are related to maths skills because they share a common processing demand. Second, none of the previous studies have explored the question of exactly which aspects of maths ability are linked to reasoning performance. Third, the fact that reasoning and maths skills have been found to be related in both children and adults hints at the possibility that both basic and advanced maths skills are related to reasoning performance. However, it is not known whether this is the case. Finally, recent studies (starting with Lyons & Beilock, 2011) have highlighted the important role that ordering abilities play in mathematics performance. Ordering abilities might be important for deductive inferences as well, although this prediction has not been tested before. We now consider each of these four issues in more detail.

(i) Differential relations or a general deductive ability?

As we have pointed out, both forms of deductive inference have been found to be related to mathematics skills. However, no study so far investigated these two types of reasoning skills and maths abilities in a single study[1]. In fact, conditional and transitive inferences are also typically investigated independently. Thus, the question remains whether a general deductive reasoning ability is responsible for the link between transitive and conditional reasoning and maths, or the links are specific to the form of deductive inference. On the basis of a previous study (Morsanyi et al., 2017) that compared the links between transitive inferences, categorical syllogisms and maths, as well as on the basis of neuroscience evidence (Prado et al., 2011) that showed dissociations between different inference forms, we expected that the links should be specific to the form of deductive inference. In the current study, we used both types of reasoning tasks, and examined their relation to maths skills, exploring whether there was evidence for a shared processing demand. We also investigated different types of conditional inferences separately. In particular, we expected that MT inferences might show different patterns of relationships with maths than the MP, DA and AC inferences. In addition, we considered whether reasoning abilities and basic mathematics skills share any variance once the effect of intelligence is taken into account.

(ii) The relation between reasoning and specific basic maths skills

Exploring in more detail the relation between performance on the reasoning tasks and performance on tasks tapping different maths skills might shed light on why maths and reasoning are linked. One of the basic maths tests that we used was a measure of arithmetic

[1] Although Handley et al. (2004) did this, they combined the scores of the two types of reasoning task, and did not report the individual results.

skills. Groen and Parkman (1972) proposed that arithmetic operations rely both on the activation and retrieval of solutions from long-term memory. This includes not just exact numerical solutions, but also procedures, such as counting skills and transformations (see Ashcraft & Guillaume 2009 for a recent review on strategies in arithmetic computations). In addition to retrieval processes, the inhibition of irrelevant information also plays a role in arithmetic operations (e.g., Campbell, 1990; Passolunghi & Siegel, 2001). Given that retrieval and inhibition processes are also important for conditional reasoning (e.g., De Neys et al., 2005), we expected that conditional reasoning and arithmetic skills might be related.

Another basic maths task that we used in this study was the number line task (e.g., Siegler & Opfer, 2003). In this task, participants have to indicate the approximate position of a number on a line where the end points are labelled by the corresponding numbers. Although the main aim of this task is to assess the nature of participants' mental representation of the number sequence (i.e., the mental number line), the task also relies on proportional reasoning and the strategic use of information about the end points of the line to find the position of the target numbers (e.g., Link et al., 2014). We expected that performance on the number line task might be related to transitive reasoning, due to the similarity of the underlying representations of item positions (see above).

(iii) Basic maths and mathematical reasoning

In addition to assessing some basic maths skills, we included two tasks assessing mathematical reasoning about word problems, to examine whether reasoning performance is similarly related to these various types of maths skills. It could be, for example, that particularly strong relations are found between performance on reasoning tasks and maths word problems, because the latter tasks require more domain-general, higher-level cognitive processes than tasks assessing basic numeracy skills. Alternatively, it may be, for the reasons outlined in the previous sub-section, that there are quite specific relations between particular aspects of basic maths skills and particular types of deductive reasoning.

The two tasks that assessed mathematical reasoning about word problems were the Cognitive Reflection Test-Long (CRT-Long; Primi, Morsanyi, Chiesi, Donati & Hamilton, 2016) and the Probabilistic Reasoning Scale (PRS; Primi, Morsanyi, Donati, Galli & Chiesi, 2017). The CRT-Long is an extended, 6-item version of the CRT (Frederick, 2005), and it measures the ability to resist intuitively compelling, but incorrect responses, and to rely on effortful processing instead. All problems include numbers, and performance on the CRT has been found to be related to numerical ability (e.g., Campitelli & Gerrans, 2014; Liberali, Reyna, Furlan, Stein, & Pardo, 2012; Morsanyi, Busdraghi & Primi, 2014), as well as cognitive abilities (e.g., Frederick, 2005). Previous studies have found that performance is also related to deductive reasoning, including syllogistic reasoning (Campitelli & Gerrans, 2014; Toplak, West & Stanovich, 2011), and conditional and transitive inferences (Primi et al., 2016). Thus, we expected that performance on the CRT-Long would be related to maths skills, as well as reasoning performance.

The PRS has also been found to be related to numerical skills and cognitive abilities, and it also showed a moderate correlation with cognitive reflection (Primi et al., 2017), although no previous studies have examined the relation with deductive reasoning. Given that some items assess conditional probability reasoning, we expected a link between the PRS and conditional reasoning. We also looked in more detail at the interplay between numerical and reasoning skills in predicting performance on both of these mathematical reasoning tasks (i.e., whether these explain unique variance – see e.g., Campitelli & Gerrans, 2014) or if there is a mediational relationship.

(iv) The role of ordering abilities in maths and reasoning skills

A basic property of the number system that affects numerical processing at multiple levels is ordinality. Numbers follow each other in a set order in the count list, the meaning of multi-digit numbers depends on the order of numerals, and complex arithmetic operations that have multiple components have to be performed in a particular order. A very simple task that has been developed to measure order processing skills is the number ordering task (i.e., three numerals are presented, and participants have to say quickly and accurately if the numerals are in the correct, ascending order). This task has been found to be a very powerful predictor of numerical skills in the case of both children (e.g., Lyons, Price, Vaessen, Blomert & Ansari, 2015) and adults (e.g., Lyons & Beilock, 2011). Nevertheless, ordering abilities are not only relevant for the domain of numbers. Indeed, our mental representation of time also relies on ordered sequences (e.g., the days of the week and the months of the calendar year), which might be very similar to the way we represent numbers (e.g., Bonato, Zorzi & Umiltà, 2012).

We note that whereas the role of ordering skills in mathematical abilities has been the focus of much research attention recently, the role of ordering abilities in reasoning performance has not been investigated so far. However, it is plausible that ordering skills may be important for deductive reasoning as well. In the case of transitive inferences, an ordered representation of the items might underlie judgements regarding the relationships between those items. In the case of conditional inferences, the ordering of the terms plays an important role in determining logical validity. Specifically, to be able to draw correct inferences, it is important to understand that *if p then q* does not necessarily imply *if q then p* (i.e., instead of a simple association between the *p* and *q* terms, where *p* and *q* would always happen together, it is only the case for *p* that it always implies *q*, but the argument does not necessarily hold the other way round – e.g., Barrouillet, Grosset, & Lecas, 2000; Gauffroy & Barrouillet, 2011). For these reasons, we hypothesized that ordering skills might play a role in both transitive and conditional reasoning performance.

Additionally, if ordering abilities are important for both basic maths and mathematical reasoning (and, possibly, for deductive reasoning skills as well), it is of interest whether ordering skills mediate the relationship between basic maths and maths reasoning skills. Indeed, as mathematical reasoning can be expected to draw more heavily on domain-general cognitive resources than basic maths skills, ordering ability might play an especially prominent role in these skills.

In summary, the current study was the first to investigate the links between both conditional and transitive reasoning skills and mathematical abilities in a single study. We used both basic (arithmetic skills, and number line performance) and complex (the CRT-Long and the PRS) measures of mathematical performance. Additionally, we assessed some general cognitive skills (fluid intelligence and ordering abilities) that could potentially explain the links between reasoning and maths skills.

Methods

Participants

The participants were 87 undergraduate psychology students (69 females) aged between 18 and 56 years (*M* = 21 years 8 months, *SD* = 6.12)[2]. Most of these participants (66 students; 54 females) also participated in an additional testing session (see procedure section). All participants provided written consent, and the study received ethical approval from the School of Psychology Ethics Committee. The students received ungraded course credit for their participation.

Materials

Deductive reasoning tasks.

[2] The results regarding the links between ordering abilities and basic mathematics skills were already reported in Morsanyi, O'Mahony & McCormack (2017). However, the main focus of that paper was on a detailed analysis of number and month ordering performance and the domain-specificity of the link between ordering abilities and mathematics skills.

The **conditional reasoning task** consisted of 16 problems (see examples in Table 1). Four problems were included, which corresponded to each of the following inference forms: MP, MT, AC, DA. All the problems had familiar, everyday content, and the availability of counterexamples was manipulated (i.e., for half of the problems within each category counterexamples were easy/difficult to retrieve)[3]. The participants were provided with detailed instructions and were asked to imagine that the first two statements were always true, before deciding if the conclusion necessarily followed. Before completing the task, the participants were also presented with a practice problem to familiarize them with the presentation format[4].

The **transitive inference task** included 12 problems. Four problems had believable conclusions, four had unbelievable conclusions and the remaining four had belief-neutral conclusions. Additionally, half of the problems within each category were valid (i.e. the conclusion followed from the premises) and half were invalid. Problems from the various categories were mixed together and were presented in the same order to all participants. Detailed instructions accompanied the problems, asking participants to accept the first two statements to be true, even if they were not true in real life. Participants were then asked to determine if the third statement logically followed from the first two statements, or if it did not necessarily follow. Cronbach's alpha was .74[5].

Tests of basic maths skills.

The **math fluency** subtest of the Woodcock-Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001) was used as a measure of arithmetic skills. This test assesses the ability to solve simple addition, subtraction, and multiplication problems. Participants had to

[3] The availability of counterexamples was established on the basis of the findings of McKenzie, Evans and Handley (2010).

[4] See the Results section for analyses regarding reliability.

[5] We have considered the possibility of analysing the results separately for different types of problems. However, when we computed the reliability of the task including all problems, we found that reliability was the highest when all problems were included together.

work through the problems as quickly and accurately as possible within a three-minute time limit. The total number of correct items was calculated to provide a maths fluency score.

Number line task. This task was based on the number-to-position problems used by Siegler and Opfer (2003) to examine numerical estimation abilities. In each of the 10 problems, participants were presented with a number and asked to estimate its position on a 0-1000 number line. The lines were positioned on the pages, so that the starting points were not directly beneath one another, preventing the participants from using previous number lines to aid their estimations. Additionally, the numbers were chosen, so that their positions could not be easily estimated based on prior knowledge (e.g., 500 should be in the middle). Performance on this task was measured by calculating the total root-mean-square error. This is the square root of the average squared difference (in millimetres) from the position selected by the participant to the actual position of the number on the number line. Greater errors reflect poorer performance. Estimation errors tend to be larger in the case of larger numbers. On this basis, we divided the trials into two halves, which were expected to have roughly equal difficulty. The Spearman-Brown corrected split-half reliability of our measure was .63.

Mathematical reasoning about word problems. The **Cognitive Reflection Test-Long** (CRT-Long; Primi et al., 2016) is an extended version of the CRT (Frederick, 2005). The test includes 6 open-ended mathematical word problems, and it measures the ability to resist intuitive response tendencies. The problems are designed in such a way that there is a typical incorrect response that comes easily to mind. It is assumed that individuals who give the correct response have to suppress an initial tendency to give the typical incorrect heuristic response (e.g., Travers, Rolison & Feeney, 2016). An example item is the following: “If it takes 5 minutes for five machines to make five widgets, how long would it take for 100 machines to make 100 widgets?” The typical heuristic response to this problem is 100 minutes, whereas the correct response is 5 minutes. The number of correct responses was summed to obtain a total score. Cronbach's alpha was .67.

The **Probabilistic Reasoning Scale** (PRS; Primi et al., 2017) is a 16-item multiple choice questionnaire that provides a comprehensive assessment of basic aspects of probabilistic reasoning, including basic and conditional probabilities presented in text and tables, reasoning about random sequences of events, and the ability to resist some typical fallacies and biases. Similar to the CRT-Long, the PRS consists of word problems, but the participants have to select the correct response out of 3 options, instead of generating a response. An example item is the following: "60% of the population in a city are men and 40% are women. 50% of the men and 30% of the women smoke. We select a person from the city at random. What is the probability that this person is a smoker?" Response options: a) 42%; b) 50%; c) 85%. (The correct response is 42%.) The total number of correct responses was summed to obtain a total score. Cronbach's alpha in the current sample was .66.

Tests of general cognitive skills

Raven's Advanced Progressive Matrices (short form). A short form of the Raven's Advanced Progressive Matrices (Arthur & Day, 1994) was used as a measure of fluid intelligence. This test consisted of 12 items plus three practice items from the Raven's Standard Progressive Matrices (Raven, 1938). This short form has been shown to have adequate psychometric properties, and it is a valid and reliable instrument (Chiesi, Ciancaleoni, Galli, Morsanyi, & Primi, 2012). In our sample, Cronbach's alpha was .69.

Ordering abilities. Two computer-based tasks were used to measure numerical (based on Lyons & Beilock, 2011) and temporal ordering ability. The participants were presented with triads of numbers (e.g., 2, 4, 1) in the number ordering task, and triads of months in the month ordering task, and they were asked to decide whether the items within each triad were in the correct order by pressing a yes/no button on the keyboard. All numbers were between 1 and 9, and the months were between January and September. Eight practice trials were presented for each task, followed by 48 experimental trials. Both accuracy and reaction times were recorded for each task, and were combined using the formula developed by Lyons et al. (2014). Then, a combined z-score was created, based on the two tasks. Cronbach's alpha for the combined score was .91.

Procedure

The participants (with the exception of 18 people who did not participate in the second testing session) were involved in two testing sessions, which lasted approximately 50 and 25 minutes, respectively. In the first session the participants were tested in groups of 8-10, and the second session was conducted in a big lecture theatre with all participants completing the tests together. The order in which the tasks were presented was the same for all participants. In session 1, the participants worked through the tasks in the following order: the maths fluency test, month ordering, the number line task, the Raven's APM, number ordering and the transitive and conditional reasoning tasks. In the second session the participants completed the PRS and the CRT-Long. The ordering tasks were computer-based. All other tasks were administered in a paper-and-pencil format.

Results

Descriptive statistics for each measure are reported in Table 2. For each task, with the exception of the number line and ordering tasks, the total number of correct responses is reported.

Insert Table 2 about here.

Our first research question was whether mathematics performance was related to a general deductive reasoning ability, or if the links were specific to certain types of inferences. In order to investigate this question, we first analysed performance on the reasoning tasks in more detail. To investigate transitive reasoning, we ran a 2x3 ANOVA with validity (valid/invalid) and believability (believable/unbelievable/belief-neutral) as within-subjects factors on endorsement rates (i.e., whether participants accepted the conclusions as correct). There was a significant effect of validity [$F(1,172) = 786.61$, $p < .001$, $\eta_p^2 = .90$], but no effect of believability, and no interaction between validity and believability. The participants endorsed 91% of the valid conclusions (*SD* = 18), but only judged 6% (*SD* = 13) of the invalid conclusions to be correct. Thus, performance was close to ceiling, and it was not affected by the content of the problems.

Concerning the conditional inferences, we ran a 2x4 ANOVA to analyse the effects of argument form (MP/MT/DA/AC) and the availability of counterexamples (high/low) on endorsement rates (see Figure 1). There was a significant effect of argument form [$F(3,258) = 41.70$, $p < .001$, $\eta_p^2 = .33$], a significant effect of the availability of counterexamples [$F(1,258) = 57.85$, $p < .001$, $\eta_p^2 = .40$], and a significant interaction between argument form and the availability of counterexamples [$F(3,258) = 32.76$, $p < .001$, $\eta_p^2 = .28$]. We ran a series of within-subjects *t* tests as follow-up analyses to investigate the effect of argument form, and the availability of counterexamples on performance on each type of problem. Endorsement rates for the MP inferences were significantly higher than for all other inference forms. Additionally, endorsement rates were significantly higher for MT than for AC arguments. The effect of counterexamples was not significant in the case of the valid MP ($p = .567$) and MT ($p = .132$) inferences, but there was a significant effect of counterexample availability in the case of the invalid inference forms ($t(86) = 6.52$, $p < .001$ for DA, and $t(86)$

= 8.78, $p < .001$ for AC). That is, when counterexamples easily came to mind, participants were more likely to correctly reject the invalid conclusions.

We also investigated the relationships between the 8 different types of conditionals (i.e., MP, MT, DA and AC with high/low availability of counterexamples). In these analyses, we considered correct responses (i.e., "yes" for MP and MT, and "no, it's not necessary" for DA and AC). Performance on the MP inferences was very high (97% correct), and the correlations with the other inference forms were inconsistent in direction and weak. For these reasons (i.e., because the MP inferences did not discriminate well between participants), we did not consider these results further[6].

Regarding the MT arguments, correct responses were negatively related to correct responding on both the DA and AC arguments (the correlation coefficients ranged from -.12 to -.36). Although the correlation between performance on the MT problems with low and high availability of counterexamples was not particularly strong [$r(85) = .28$ $p = .009$], we considered the results for these problems together (i.e., we created an MT composite score). Finally, there were significant, medium-to-strong, positive correlations between the DA and AC arguments with high/low availability of counterexamples (*r*s ranged from .34-.76 $p<.001$). For this reason, we analysed these results together (Cronbach's alpha for this measure was .79).

____

Insert Figure 1 about here.

____

[6] On a theoretical basis, combining the scores from the MP, AC and DA inferences would seem appropriate. Nevertheless, we decided against this on the basis of our empirical findings, which showed that performance on the MP problems was at ceiling, and that including these scores would have somewhat reduced the reliability of our conditional reasoning measure.

Next we considered the relationships between the reasoning tasks and basic maths skills (i.e., arithmetic skills and performance on the number line task), taking into account the potential effect of more general skills, such as fluid intelligence and ordering ability (see Table 3 for raw correlations[7]). Transitive reasoning was related to performance on the number line task, as well as fluid intelligence. The relationship between transitive inferences and the number line task remained significant after controlling for fluid intelligence and ordering ability [$r(79) = .37$ $p = .001$].

Correct performance on the MT inferences was negatively related to correct performance on the DA and AC inferences, but it was not related to any other tasks. This was in contrast with performance on the DA and AC problems, which was related to a wide variety of tasks, including maths fluency and number ordering. It is also notable that performance on the transitive and conditional reasoning tasks was unrelated, although both tasks showed relationships with at least some measures of mathematical ability. To investigate the link between the DA and AC inferences and maths fluency, we ran partial correlations controlling for fluid intelligence and ordering ability. When the effect of fluid intelligence was controlled, the link between maths fluency and conditional reasoning skills remained significant. However, controlling for ordering skills rendered this relationship non-significant.

Next we investigated the link between reasoning skills and performance on the mathematical reasoning tasks, taking into account the effect of general cognitive skills, ordering abilities, and basic maths skills. Concerning the CRT-Long, performance was related to transitive and conditional reasoning ability, maths fluency, probabilistic reasoning, fluid intelligence and ordering skills. We ran a stepwise regression analysis including all

[7] Given the large number of variables, we have checked the robustness of these correlational patterns using a bootstrapping procedure with 10,000 bootstrap samples. All significant correlations remained significant when a bootstrapping procedure was used, with the exception of the relationship between the CRT-Long and fluid intelligence, and mathematical reasoning and performance on the number line task.

variables apart from the PRS to find the best basic predictors of cognitive reflection. The final model [$F(3,61) = 14.59$, $p < .001$] explained 30% of the variance in cognitive reflection, and it included ordering ability [$\beta = .32$ $p = .005$], conditional reasoning [DA and AC; $\beta = .32$ $p = .005$], and transitive reasoning [$\beta = .25$ $p = .025$]. Although fluid intelligence was not a significant predictor in this model, we re-ran this analysis, including only the significant predictors of cognitive reflection and fluid intelligence as a covariate, to check the robustness of this model when the effects of intelligence are controlled. The effect of conditional and transitive reasoning, as well as ordering ability remained significant when the effect of fluid intelligence was controlled[8].

____

Insert Table 3 about here.

____

Probabilistic reasoning performance was significantly related to conditional reasoning, maths fluency, performance on the number line task, fluid intelligence, ordering ability and CRT performance. We also ran a stepwise regression analysis to predict performance on the PRS including all variables, apart from CRT performance. The regression analysis resulted in a significant model [$F(3,64) = 14.86$, $p < .001$] that explained 39% of the variance in probabilistic reasoning. The significant predictors that were included in the final model were ordering ability [$\beta = .47$ $p < .001$], conditional reasoning (i.e., the composite of AC and DA; $\beta = .26$ $p = .014$) and transitive reasoning ($\beta = .23$ $p = .022$). Similar to our analysis regarding cognitive reflection, we checked the robustness of this model, by including fluid intelligence as an additional predictor, and by using a bootstrapping procedure with 10,000 bootstrap samples. All significant effects remained unchanged.

[8] We have also tested the robustness of the regression results by using a bootstrapping procedure with 10,000 samples. The same results were obtained as with the traditional analyses (i.e., transitive and conditional inferences, as well as ordering ability were significant predictors of cognitive reflection, but fluid intelligence was not).

It is interesting that maths fluency was not included in the final regression models predicting CRT and PRS performance, although these tasks require arithmetic operations, and maths fluency was moderately related to both tasks. Given the pattern of relationships between these tasks, the most likely explanation is that the effect of maths fluency on cognitive reflection and probabilistic reasoning was mediated by ordering ability or conditional reasoning (or both). In order to check this possibility, we tested a multiple mediation model to predict performance on the maths reasoning problems from maths fluency, with conditional inferences and ordering ability as potential mediators (see Figure 2). In this analysis, we combined the two mathematical reasoning tasks (using combined z scores), because performance on these tasks was strongly correlated, and because the regression models that we reported above showed that performance on these tasks was predicted by the same set of variables. In the mediation model, we considered ordering ability as a potential mediator between both maths fluency and mathematical reasoning, and conditional reasoning and mathematical reasoning, as ordering performance was particularly strongly related to performance on the maths reasoning tasks. Indeed, it is possible that the links between conditional reasoning ability and mathematical reasoning are mediated by ordering skills. As in the previous analyses, we also included fluid intelligence as a covariate to control for the effect of general cognitive ability. The mediation hypotheses were tested using the INDIRECT procedure for bootstrapping (with 10,000 bootstrap samples) to estimate 95% confidence intervals for the regression coefficients (CI; for more details see Preacher & Hayes, 2008). We estimated the indirect effect of maths fluency on mathematical reasoning via the following pathways. Pathway 1: maths fluency – conditional reasoning – mathematical reasoning. Pathway 2: maths fluency – conditional reasoning - ordering ability – mathematical reasoning. Pathway 3: maths fluency – ordering ability – mathematical reasoning. A bias-corrected bootstrap-confidence interval (CI) of the product of the paths within each indirect route that does not include zero provides evidence of a significant indirect effect of maths fluency on mathematical reasoning through the mediator variables (Preacher & Hayes, 2008). The INDIRECT procedure resulted in a significant model ($p < .0001$ for the total effect), and it revealed that all indirect pathways were significant. Specifically, there was a significant indirect effect of maths fluency on mathematical reasoning through indirect pathway 1 (95% *CI* = .0005 to .0147), pathway 2 (95% *CI* < .0001 to .0033) and pathway 3 (95% *CI* = .0059 to .0259). Additionally, fluid intelligence was a significant covariate in the model (p=.003).

———

Insert Figure 2 about here.

———

Discussion

This study addressed several questions regarding the relationship between transitive and conditional reasoning and mathematics skills. We were interested in whether the two types of reasoning had a similar relation to maths skills, suggesting that maths is linked to a general deductive ability, or whether each of the two types of reasoning had specific relations with particular maths skills. We found little evidence to suggest that maths is linked to a general deductive ability that is measured by both types of reasoning tasks. Performance levels on the two reasoning tasks did not correlate with each other; moreover, the two types of tasks showed different patterns of relations with the various maths tasks.

The relation between transitive and conditional reasoning and specific basic maths skills

Performance on the transitive reasoning task, but not the conditional reasoning task, was significantly related to number line performance when we controlled for intelligence. This finding is in line with previous behavioural (e.g., Goodwin & Johnson-Laird, 2005, 2008; Prado et al., 2008; Vandierendonck & De Vooght, 1997), and neuroimaging (Prado et al., 2011) evidence that transitive inferences utilize a spatial representation (which might be similar to the mental number line representation). Given that number line performance is closely related to maths abilities from childhood (e.g., Siegler & Booth, 2004), this finding might also explain why transitive reasoning skills have been found previously to have a strong association with maths performance in children (Morsanyi et al., 2013) and adolescents (Morsanyi et al., 2017).

Regarding conditional inferences, performance was moderately related to basic arithmetic skills. This relationship remained significant when the effect of fluid intelligence was controlled, but it was not significant anymore once the effect of ordering skills was controlled. The mediation analyses additionally showed that ordering abilities were important for both conditional reasoning and arithmetic skills, and the requirement of ordering skills also explained (at least partially) why arithmetic skills and conditional reasoning were linked to complex mathematical reasoning. These are novel findings that might inspire further research into the processes that underlie both conditional inferences and complex mathematical reasoning performance.

Our analyses regarding the availability of counterexamples showed (in line with previous studies – e.g., Cummins, 1995; De Neys et al., 2005; Quinn & Markovits, 1998; Thompson, 1994) that conditional reasoning performance was greatly influenced by the retrieval of counterexamples from memory. The retrieval of relevant knowledge (facts, rules and procedures) is very important for arithmetic performance as well (e.g., Campbell & Xue, 2001; Groen & Parkman, 1972; LeFevre, Sadesky & Bisanz, 1996). Nevertheless, we did not find evidence for a differential relationship between conditional reasoning and arithmetic skills, depending on the availability of counterexamples[9]. An alternative explanation that has been proposed by Markovits and Lortie-Forgues (2011) and Moshman (1990) was that "if-then" statements form the basis of all scientific thinking, including mathematical thinking. However, this does not explain why the link between conditional reasoning and maths was restricted to certain types of conditional inferences.

The relations between different types of deductive inferences

As we have noted, performance on the two types of deductive reasoning tasks (i.e., on the conditional and transitive inferences) was unrelated; moreover, we found that performance on the different categories of conditional inferences was in one instance negatively related (DA and AC were negative correlated with MT inferences). The results regarding the MT inferences are in line with Attridge and Inglis (2013) who reported that performance on MT problems changed in the opposite direction than performance on the DA and AC problems in the case of students who participated in post-compulsory maths education. Newstead et al. (2004) also reported that, whereas correct reasoning about DA and AC inferences was positively related to intelligence, there was a negative relationship between correct MT inferences and intelligence. These results, together with the differential links between maths skills and different types of deductive inferences (see also Morsanyi et al., 2017), suggest that deductive reasoning skills are supported by inference-specific processes, rather than a general deductive reasoning ability.

A potential alternative explanation regarding the negative correlation in our study between the MT and the AC and DA inferences (which is similar to the findings of Attridge

[9] The correlation coefficients for DA and AC inferences with a high availability of counterexamples were relatively low (.18 for both DA and AC), and non-significant. The correlation coefficients for DA and AC inferences with a low availability of counterexamples were stronger (.30 for DA and .29 for AC), and significant at the $p<.01$ level. This seems to suggest that the relationship with maths was stronger in the case of conditionals where the retrieval of counterexamples required more effort. Nevertheless, when we obtained confidence intervals for these correlations using a bootstrapping procedure with 10,000 samples, we found that these correlation coefficients were not significantly different.

& Inglis, 2013 and Newstead et al., 2004) could be that some participants were simply more likely to accept/reject the conclusions than others, regardless of the logical status of the conditionals. Indeed, there is a dominant response in the case of conditionals (i.e., participants tend to endorse the conclusions). Nevertheless, our results suggested that participants' responses were strongly affected by both the argument form and the availability of counterexamples. Thus, these responses clearly reflect more than just a response bias. Developmental studies (e.g., Klaczynski, Schuneman & Daniel, 2004) also suggest an increasing dissociation between responses to different types of conditionals. For these reasons, we believe that our decision to investigate these inference forms separately, rather than creating an overall index of conditional reasoning, was justified.

The conclusion that deductive reasoning relies on inference-specific processes is also consistent with a meta-analysis of brain-imaging studies (Prado et al., 2011) that found that transitive, conditional and categorical inferences were related to activations in three distinct brain subsystems. Regarding conditional inferences, evidence for a dissociation between the brain basis of MP and MT inferences has been reported (Noveck, Goel & Smith, 2004). Specifically, the brain activation patterns related to MP and MT inferences were different from each other, and they also significantly differed from a baseline condition that required participants to draw a trivial conclusion. However, for the AC and DA arguments, the brain activation patterns were not different from baseline.

Mathematical reasoning ability and its predictors

We also investigated the interplay between mathematical skills and deductive reasoning ability in shaping complex mathematical reasoning skills. Interestingly, although the problems had a mathematical content, and arithmetic skills were moderately related to performance on both the CRT-Long and the PRS, arithmetic skills did not explain further variance in performance on these tasks, once the effect of transitive and conditional inferences, as well as ordering abilities were taken into account. Ordering skills have been found to play a very important role in mathematical cognition (e.g., Lyons & Beilock, 2011; Lyons et al., 2014), and there is evidence that order memory is also important for reading (Perez, Majerus & Poncelet, 2012; 2013). On the basis of these findings, it might not be very surprising that complex mathematical reasoning tasks (which combine the requirement of text comprehension with using numerical information) are also strongly related to ordering abilities. Nevertheless, our study is the first to demonstrate this link, and further investigations into the role of ordering skills in mathematical reasoning could be important for a better understanding of this relationship.

It is also important to note that whereas the believability of conclusions has been found to strongly influence transitive reasoning performance in the case of children and adolescents (e.g., Morsanyi et al., 2013; 2017) and even in the case of adults, when the problems were presented briefly on a computer screen (Andrews, 2010), in the current study, believability had no effect on transitive reasoning performance. In fact, our participants seemed to rely on very similar strategies in the case of belief-laden and belief-neutral problems, which suggests that they probably relied on an abstract strategy that did not necessitate the processing of problem content. Such strategies might be very important for mature cognitive skills, and future studies could investigate the development of these strategies (which probably occurs around late adolescence). The fact that transitive reasoning skills were predictive of mathematical reasoning extends the previous findings that linked transitive reasoning to basic maths abilities, and it suggests that reasoning abilities might play a role in educational achievement, especially in the case of quantitative subjects.

A related issue is whether there is a causal link between the development of reasoning and maths skills. Piaget (1952) and Russell (1919) assumed that logic was a prerequisite of mathematics knowledge, whereas other approaches, such as the “theory of formal discipline” (first proposed by Plato) assumed that the link was in the other direction (i.e., training in maths improves reasoning). Attridge and Inglis (2013) found that post-compulsory education in mathematics resulted in improvements in conditional reasoning skills – a finding that is in line with the theory of formal discipline. Nevertheless, the current educational system does not provide training in reasoning skills, and, thus, it is difficult to tell whether such training would generalise to mathematics abilities. A recent study by Knoll, Fuhrmann, Sakhardande, Stamp, Speekenbrink and Blakemore (2016) reported that training in abstract relational reasoning benefited all age groups between the ages of 11-33, but late adolescents and adults showed disproportionate improvement. It is an intriguing possibility that late adolescence could be an ideal time to train other types of reasoning skills as well.

Our findings suggest that the developmental links between mathematics and reasoning might be very complex, and they exist both at the level of basic skills and mathematical reasoning. As transitive and conditional reasoning skills continue to develop into late adolescence, and they play a role in complex mathematics abilities, they could be ideal targets for future training studies. Given its central role in both reasoning ability and mathematical skills, another target for cognitive training programmes could be ordering ability. In fact, it is possible that any transfer from training in mathematics to reasoning skills, and vice versa, is at least partially the consequence of incidentally training ordering skills.

Limitations and future directions

Previous studies on the link between mathematics and deductive reasoning have not investigated transitive reasoning and conditional inference together. A methodological issue regarding a comparison between these two forms of deductive inference is that the ability to draw transitive inferences develops much earlier than conditional reasoning skills. Although even young children can draw MP inferences (e.g., Byrnes & Overton, 1986), the ability to reject the invalid AC and DA inferences only starts to emerge in adolescence (e.g., Markovits

& Vachon, 1990). In the current study, whereas performance on the transitive reasoning problems was close to ceiling, rejection rates for the AC and DA inferences were low-to moderate (i.e., these inferences were challenging even for our sample of educated adults). It is possible that we underestimated the strength of the relationship between transitive reasoning and maths (relative to the link between conditional reasoning and maths), because our measure had limited sensitivity. Possible ways to address this issue in future studies would be to use a more difficult transitive inference task (for example, by including problems with more than three terms – see Andrews, 2010), to measure reaction times, or to recruit younger participants for the study. Nevertheless, despite its limitations, our transitive reasoning task had a good level of reliability, and showed moderate relationships with some other measures, which confirmed its validity.

Summary and conclusions

In summary, our study replicated previous findings (Attridge & Inglis, 2013; Handley et al., 2004; Inglis & Simpson, 2008, 2009; Morsanyi et al., 2013, 2017) that showed a link between both transitive and conditional inferences and mathematics skills. However, previous studies did not investigate the nature of these relationships in detail. We found that these inference types independently predicted mathematical reasoning, and they were also related to some basic maths abilities. Additionally, our results highlighted the important role that ordering abilities play in complex mathematical reasoning, extending findings regarding the role of ordering abilities in basic maths skills. Indeed, ordering abilities were also related to conditional reasoning ability – a novel finding that deserves further attention. Future studies could investigate the educational implications of these findings, by providing training in reasoning and ordering skills (whilst acknowledging that training in maths might also improve reasoning, and, possibly, ordering abilities). It can also be expected that the development of the representations and strategies that underlie reasoning and mathematical skills might show remarkable similarities.

References

Andrews, G. (2010). Belief-based and analytic processing in transitive inference depends on premise integration difficulty. *Memory & Cognition*, *38*, 928-940.

Arthur, W. Jr., & Day, D. V. (1994). Development of a short form for the Raven advanced progressive matrices test. *Educational and Psychological Measurement*, *54*, 394–403.

Ashcraft, M. H., & Guillaume, M. M. (2009). Mathematical cognition and the problem size effect. In B. Ross (Ed.). *The psychology of learning and motivation* (Vol. 51, pp. 121–151). Burlington: Academic Press.

Attridge, N. & Inglis, M. (2013). Advanced mathematical study and the development of conditional reasoning skills. *PLOS ONE, 8,* e69399.

Barrouillet, P., Grosset, N., & Lecas, J. F. (2000). Conditional reasoning by mental models: Chronometric and developmental evidence. *Cognition, 75,* 237–266.

Bonato, M., Zorzi, M., & Umiltà, C. (2012). When time is space: Evidence for a mental time line. *Neuroscience and Biobehavioral Reviews, 36,* 2257–2273.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 42*, 189–201.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79,* 1016–1031.

Byrnes, J. P., & Overton, W. F. (1986). Reasoning about certainty and uncertainty in concrete, causal, and propositional contexts. *Developmental Psychology, 22*, 793-799.

Campbell, J. I. (1990). Retrieval inhibition and interference in cognitive arithmetic. *Canadian Journal of Psychology, 44,* 445-464.

Campbell, J. I. D., & Xue, Q. (2001). Cognitive arithmetic across cultures. *Journal of Experimental Psychology: General, 130*, 299–315.

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition*, *42,* 434–447.

Chiesi, F., Ciancaleoni, M., Galli, S., Morsanyi, K. & Primi, C. (2012). Item Response Theory analysis and Differential Item Functioning across age, gender and country of a short form of the Advanced Progressive Matrices. *Learning and Individual Differences, 22,* 390-396.

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

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

De Neys, W., Schaeken, W., & d'Ydewalle, G. (2005). Working memory and everyday conditional reasoning: Retrieval and inhibition of stored counterexamples. *Thinking & Reasoning, 11,* 349-381.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42.

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

Gauffroy, C., & Barrouillet, P. (2011). The primacy of thinking about possibilities in the development of reasoning. *Developmental Psychology, 47*, 1000-1011.

Goodwin G.P. & Johnson-Laird, P.N. (2005). Reasoning about relations. *Psychological Review, 112*, 468–493.

Goodwin, G.P. & Johnson-Laird, P.N. (2008). Transitive and pseudotransitive inferences. *Cognition, 108*, 320–352.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review, 79*, 329–343.

Handley, S., Capon, A., Beveridge, M., Dennis, I., & Evans, J.St.B.T. (2004). Working memory, inhibitory control, and the development of children's reasoning. *Thinking & Reasoning, 10,* 175–195.

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

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

Klaczynski, P. A., Schuneman, M. J., & Daniel, D. B. (2004). Theories of conditional reasoning: A developmental examination of competing hypotheses. *Developmental Psychology, 40,* 559–571.

Knoll, L. J., Fuhrmann, D., Sakhardande, A. L., Stamp, F., Speekenbrink, M., & Blakemore, S. J. (2016). A window of opportunity for cognitive training in adolescence. *Psychological Science*, *27*, 1620– 1631.

LeFevre, J.-A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem size effect in adults. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 216–230.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, *25*, 361-381.

Link, T., Nuerk, H. C., & Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *The Quarterly Journal of Experimental Psychology*, *67,* 1597-1613.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*, 256–261.

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in Grades 1–6. *Developmental Science, 17*, 714–726.

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

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

Markovits, H., & Vachon, R. (1990). Conditional reasoning, representation, and level of abstraction. *Developmental Psychology*, *26*, 942-951.

McKenzie, R., Evans, J. S. B., & Handley, S. J. (2010). Conditional reasoning in autism: activation and integration of knowledge and belief. *Developmental Psychology*, *46*, 391-403.

Morsanyi, K., Busdraghi, C. & Primi, C. (2014). Mathematical anxiety is linked to reduced cognitive reflection: A potential road from discomfort in the mathematics classroom to susceptibility to biases. *Behavioral and Brain Functions* 2014, 10:31.

Morsanyi, K., Devine, A., Nobes, A. & Szucs, D. (2013). The link between logic, mathematics and imagination. Evidence from children with developmental dyscalculia and mathematically gifted children. *Developmental Science*, *16*, 542-553.

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., O'Mahony, E. & McCormack, T. (2017). Number comparison and number ordering as predictors of arithmetic performance in adults: Exploring the link between the two skills, and investigating the question of domain-specificity. *The Quarterly Journal of Experimental Psychology, 70*, 2497-2517.

Moshman, D. (1990). The development of metalogical understanding.In W.F. Overton (Ed.), *Reasoning, necessity, and logic: developmental perspectives* (pp. 205–225). Hillsdale, NJ: Erlbaum.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519-1520.

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

Noveck, I. A., Goel, V., & Smith, K. W. (2004). The neural basis of conditional reasoning with arbitrary content. *Cortex, 40*, 613–622.

Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, *80*, 44-57.

Perez, T. M., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, *111*, 708-723.

Perez, T. M., Majerus, S., & Poncelet, M. (2013). Impaired short-term memory for order in adults with dyslexia. *Research in Developmental Disabilities*, *34*, 2211-2223.

Piaget, J. (1952). *The child's conception of number.* London: Routledge & Kegan Paul.

Preacher, K.J. & Hayes, A.F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavioral Research Methods, 40*, 879–891.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M.A. & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making, 29,* 453–469.

Primi, C., Morsanyi, K., Donati, M.A., Galli, S. & Chiesi, F. (2017). Measuring probabilistic reasoning: The construction of a new scale applying Item Response Theory. *Journal of Behavioral Decision Making (in press)* DOI:10.1002/bdm.2011

Quinn, S., & Markovits, H. (1998). Conditional reasoning, causality, and the structure of semantic memory: Strength of association as a predictive factor for content effects. *Cognition, 68*, 93–101.

Raven, J. C. (1938). *Progressive matrices: A perceptual test of intelligence*. London: H. K. Lewis.

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83*, 274–278.

Russell, B. (1919). Introduction to Mathematical Philosophy. London: George Allen & Unwin, Ltd.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75*, 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237-243.

Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition*, *22,* 742-758.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The Cognitive Reflection Test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39,* 1275-1289.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning*, *20,* 147-168.

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the Cognitive Reflection Test. *Cognition*, *150,* 109-118.

Vandierendonck, A., & deVooght, G. (1997). Working memory constraints on linear reasoning with spatial and temporal contents. *The Quarterly Journal of Experimental Psychology, 50A*, 803–820.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.

Table 1. Examples of conditional inferences (with low availability of counterexamples) that were used in the study

Modus Ponens (MP)	Modus Tollens (MT)
Imagine that the following is always true: If butter is heated, then it will melt. **Now imagine that this is also true:** The butter is heated. **Is it necessary that:** The butter will melt? *Yes.* * *No, it's not necessary.* No, it's not necessary.	**Imagine that the following is always true:** If there is a power cut, then the lights will go out. **Now imagine that this is also true:** The lights don't go out. **Is it necessary that:** There is no power cut? *Yes.** *No, it's not necessary.*
Affirmation of the consequent (AC)	**Denial of the antecedent (DA)**
Imagine that the following is always true: If the trigger is pulled, then the gun will fire. **Now imagine that this is also true:** The gun fires. **Is it necessary that:** The trigger was pulled? *Yes. No, it's not necessary.**	**Imagine that the following is always true:** If a paper is burnt, then it will become ash. **Now imagine that this is also true:** The paper is not burnt. **Is it necessary that:** The paper doesn't become ash? *Yes. No, it's not necessary.**

Correct responses are marked with an asterisk.

Table 2. *Descriptive statistics for the measures used in the study.*

	Mean	SD	Minimum	Maximum
Transitive reasoning	11.11	1.70	4	12
MP inferences	3.89	.36	2	4
MT inferences	3.29	.94	0	4
DA and AC inferences	5.47	2.43	0	8
Maths fluency	110.15	21.24	67	158
Number line task	3.57	1.68	1.21	9.61
Raven's matrices	5.06	2.30	0	11
Ordering (accuracy)	.92	.06	.70	1.00
Ordering (RT)	1485.89	335.70	821.77	2518.08
CRT-Long	2.03	1.44	0	6
Probabilistic Reasoning Scale	12.72	2.24	7	16

Table 3. *Relationships between the tasks measuring reasoning skills, basic maths skills, mathematical reasoning, fluid intelligence and ordering skills*

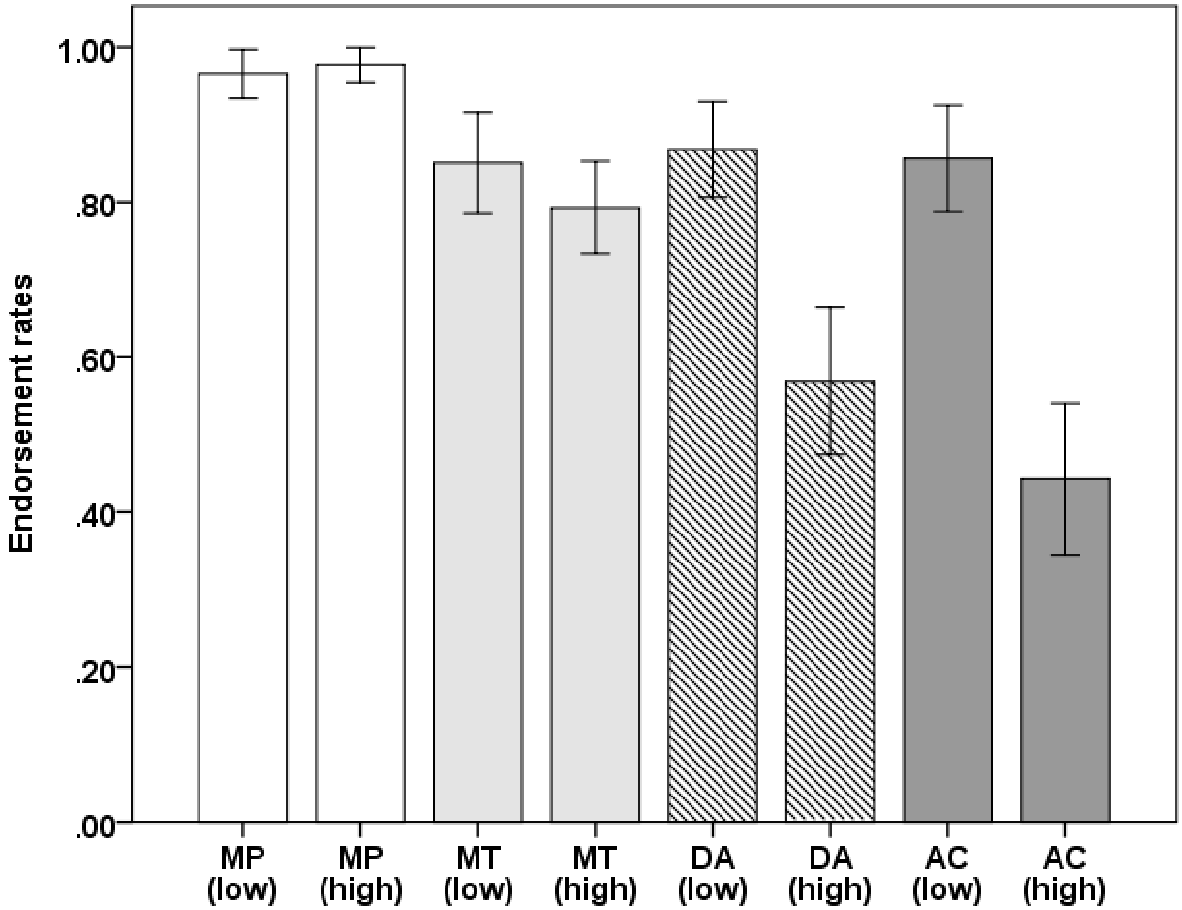
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.
1. Transitive	--									
2. MT	.07	--								
3. DA and AC	-.06	-.41**	--							
4. Maths fluency	.04	-.19	.28**	--						
5. Number line	.37**	.18	.10	.18	--					
6. Raven's APM	.23*	.15	.08	.04	.27*	--				
7. Ordering	.17	-.01	.23*	.51**	.38**	.35**	--			
8. CRT-Long	.26*	-.22	.38**	.40**	.16	.25*	.45**	--		
9. PRS	.21	.02	.36**	.33**	.39**	.36**	.56**	.50**	--	
10. Maths reasoning[1]	.30*	-.12	.43**	.42**	.32*	.35**	.58**	.87**	.87**	--

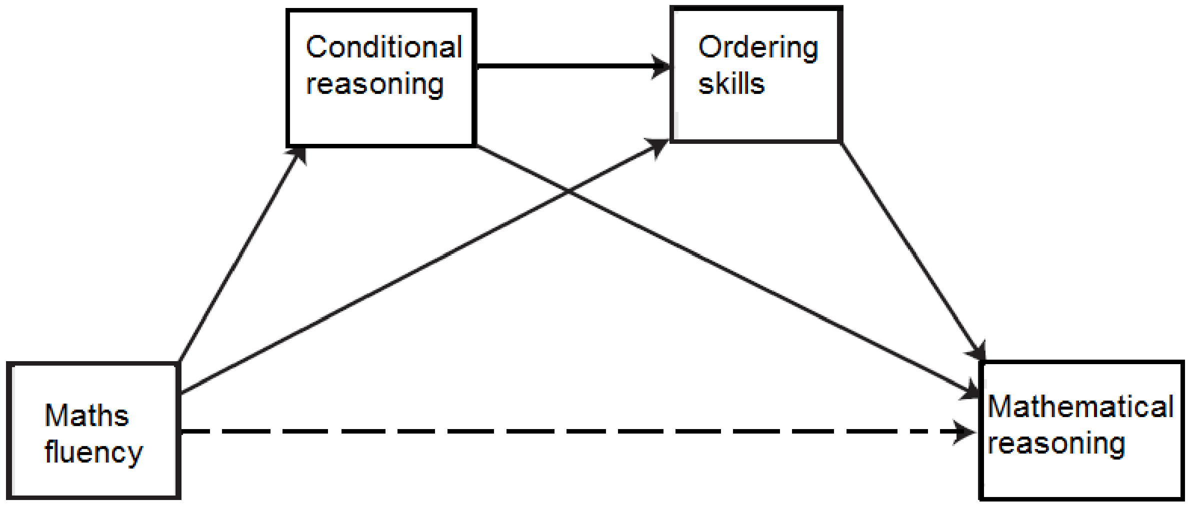
[1]This measure was based on the combined *z* scores of the CRT-Long and the PRS

Figure captions

Figure 1. Endorsement rates (i.e., the proportion of “yes” responses) for the MP, MT, DA and AC inferences, as a function of the availability of counterexamples.

Figure 2. Mediation model to predict mathematical reasoning from maths fluency with conditional reasoning ability and ordering skills as mediators. (Solid lines represent significant links. Fluid intelligence was included in the model as a covariate.)


Endorsement rates
1.00
.80
.60
.40
.20
.00
MP (low)
MP (high)
MT (low)
MT (high)
DA (low)
DA (high)
AC (low)
AC (high)


Conditional reasoning
Ordering skills
Maths fluency
Mathematical reasoning