Investigating the role of attention in the identification of associativity shortcuts using a microgenetic measure of implicit shortcut use

Journal:	*Quarterly Journal of Experimental Psychology*
Manuscript ID	QJE-STD-19-007.R2
Manuscript Type:	Standard Article
Date Submitted by the Author:	07-Jul-2019
Complete List of Authors:	Eaves, Joanne; Loughborough University Gilmore, Camilla; Loughborough University Attridge, Nina; Loughborough University
Keywords:	Associativity, attention, conceptual understanding, arithmetic, measurement, strategy

DOI: 10.1177/1747021820905739

Title

Investigating the role of attention in the identification of associativity shortcuts using a microgenetic measure of implicit shortcut use

Running head

Attention and arithmetic strategy identification

Keywords

Associativity; attention; conceptual understanding; arithmetic; strategy; measurement

Author names and affiliations

Joanne Eaves[a]

[a]Mathematics Education Centre, Loughborough University, LE11 3TU

Camilla Gilmore[a]

[a]Mathematics Education Centre, Loughborough University, LE11 3TU

Nina Attridge[a]

[a]Mathematics Education Centre, Loughborough University, LE11 3TU

Corresponding author

Joanne Eaves[a]

Email: J.Eaves@lboro.ac.uk

[a]Mathematics Education Centre, Loughborough University, LE11 3TU

Acknowledgements

This work was supported by a PhD Studentship from Loughborough University Doctoral college. C.G. is supported by a Royal Society Dorothy Hodgkin Fellowship.

DOI: 10.1177/1747021820905739

Abstract

Many mathematics problems can be solved in different ways or by using different strategies. Good knowledge of arithmetic principles is important for identifying and using strategies that are more sophisticated. For example, the problem '6 + 38 − 35' can be solved through a shortcut strategy where the subtraction '38 − 35' is performed before the addition '3 + 6 = 9', a strategy which is derived from the arithmetic principle of associativity. However, both children and adults make infrequent use of this shortcut and the reasons for this are currently unknown. To uncover these reasons, new sensitive measures of strategy identification and use must first be developed, which was one goal of our research. We built a novel method to detect the time-point when individuals first identify an arithmetic strategy, based on a trial-by-trial response time data. Our second goal was to use this measure to investigate the contribution of one particular factor, attention, in the identification of the associativity shortcut. In two studies, we found that manipulating visual attention made no difference to the number of people who identified the shortcut, the trial number on which they first identified it, or their accuracy and response time for solving shortcut problems. We discuss the theoretical and methodological contribution of our findings, and argue that the origin of people's difficulty with associativity shortcuts may lie beyond attention.

DOI: 10.1177/1747021820905739

Introduction

The importance of conceptual knowledge

Arithmetic knowledge can be crudely divided into procedural and conceptual (Rittle-Johnson & Siegler, 1998), two broad categories of understanding that, whilst related, may be understood by an individual to different extents (Gilmore & Papadatou-Pastou, 2009). Procedural skills refer to the action sequences that are used to solve problems, for example, counting and decomposition. Conceptual knowledge refers to the understanding of principles and relationships that underlie a domain (Hiebert & Lefevre, 1986), for example, knowing that addition and subtraction have an opposite relation (the principle of inversion). Seven arithmetic principles are often discussed: identity, negation, complementarity, commutativity, inversion, equivalence and associativity (Kilpatrick, Swafford, & Findell, 2002). A solid grasp of these principles is key for success in mathematics. For example, conceptual knowledge enables children to progress from using less sophisticated arithmetic strategies such as counting, to those that are more efficient, such as decomposition (Bryant, Christie, & Rendu, 1999).

Our research focused on associativity, a principle thought to be important for aiding the transition from arithmetic to algebra and that may predict educational and employment success (Ladson-Billings, 1997; Kilpatrick et al., 2002). Associativity is the principle that allows problems to be solved by first decomposing, and then recombining their problem sets (Canobi, Reeve, & Pattison, 1998), for example, solving ‘a + b − c’ by first performing ‘b − c’ and then adding the result to ‘a’. In other words, for some problems, the answer will be the same regardless of which pair of numbers is dealt with first. Different forms of the principle exist, such as addition only, ‘a + b + c = c + b + a’ (Canobi, 2005), addition-subtraction, ‘a + b − c = b − c + a’, and multiplication-division, ‘a × b ÷ c = b ÷ c × a’ (Robinson & Ninowski, 2003). Problems with opposing operations (addition-subtraction, multiplication-division) are a dominant paradigm used to investigate how well an individual

DOI: 10.1177/1747021820905739

understands associativity (Robinson & Dube, 2017). We adhered to this by focusing specifically on addition-subtraction.

Methodological issues in measuring conceptual understanding

Methods for measuring conceptual knowledge sparks animated debate (Schneider & Stern, 2010). Associativity is no exception. One reason for this is that almost any mathematical task involves a combination of procedural and conceptual knowledge and therefore pure measures of conceptual understanding are difficult to develop. Existing methods broadly divide into explicit and implicit, both of which normally infer conceptual knowledge from the strategies used to solve unfamiliar problems. Explicit methods require individuals to explain or justify why strategies are appropriate, and implicit approaches infer strategy use from accuracy and solution latencies to different problems. For associativity, three-term (a + b − c) problems are commonly used (Klein & Bisanz, 2000). Verbal reports of solving 'a + b − c' by performing the subtraction first and adding the result to 'a' allow the researcher to assume that the individual has applied some knowledge of associativity. That is, they have used their understanding of the principle to execute a 'right-to-left' strategy (hereafter a 'shortcut'). For implicit measures, accuracy and response time can be compared between problems that are conducive and non-conducive to a shortcut (Edwards, 2013). Conducive problems such as '16 + 47 − 45' encourage shortcut use by the subtraction being easier than the addition. Non-conducive problems do not encourage a shortcut, for example '36 + 27 − 45', because solving the subtraction first offers less advantage. If conducive and non-conducive problems are equally challenging when solved through a left-to-right strategy, differences in accuracy and response time between the two problem types may indicate that an individual is using different strategies to solve them. If accuracy and response time are substantially better for conducive than non-conducive problems, it may be inferred that the individual has used the shortcut and has applied their knowledge of associativity.

DOI: 10.1177/1747021820905739

Neither of these methods (explicit, implicit) are perfect, and researchers have widely documented concerns with each (Prather & Alibali, 2009; Crooks & Alibali, 2014; Schneider & Stern, 2010; Rittle-Johnson & Schneider, 2015; Faulkenberry, 2013). For example, explicit techniques rely on participants being aware of, remembering, and accurately reporting their strategies, which may not always give reliable results (Posner & Gertzog, 1982) and may not be suitable for some individuals (e.g. children). Implicit techniques of accuracy and response time imply but do not guarantee strategy use, but they can capture elements of conceptual understanding that an individual may be unaware of or unable to express (Bryant et al., 1999). As a result, researchers are encouraged to use multiple methods for measuring strategy use (Schneider & Stern, 2010), to provide greater justification for the tasks that they use (Crooks & Alibali, 2014), and to develop new measures where appropriate (Rittle-Johnson & Schneider, 2015). While some progress has been made (e.g. Thevenot & Oakhill, 2005; Thevenot & Oakhill, 2006), few have focused on developing new methods for measuring conceptually-derived strategy use on digit-based problems. Our research contributed to these goals by 1) developing a new tool to measure conceptual shortcut use, 2) incorporating both self-report and implicit measures, and 3) ensuring that the measure we used closely matched the specific aspect of the associativity shortcut strategy that we wished to test. The specific aspect that we wished to test was the time-point when an individual first used the strategy within a set of problems, hereafter ‘identification’.

Barriers to using conceptually-derived shortcuts

Knowledge of associativity is often compared to inversion, the simpler principle that addition-subtraction and multiplication-division have opposite relations (Baroody, 2003). Inversion is often measured through simpler shortcut problems of the form ‘a + b − b’. Individuals who understand inversion know that the addition and subtraction cancel out, and that they can simply pick ‘a’ (Starkey & Gelman, 1989; Bisanz & LeFevre, 1990). Around 35 – 60% of children use inversion shortcuts by the age of 6 – 10 years (Watchorn et al., 2014; Robinson & Dubé, 2012, 2013) and

DOI: 10.1177/1747021820905739

match adults' frequency of use by 14 years (Dubé, 2014). In comparison, the use of associativity shortcuts lags behind (Robinson & Dube, 2017). For children aged 6 – 10 years, use of the associativity shortcut is between 15 – 25%, a rate that remains low (approximately 30%) in early adolescence (11 – 13 years) and reaches only approximately 50% in adulthood (Dubé, 2014; Dubé & Robinson, 2010). Education practitioners have called for this situation to change (National Mathematics Advisory Panel, 2008). To do so, we first need to understand the reasons why associativity shortcut use is low, a topic that we now address.

There are many reasons why an individual may not use an associativity shortcut. These may be domain-specific, referring to the skills that are required only in arithmetic, or domain-general, referring to the skills required on a range of tasks (Fuchs et al., 2010). From a domain-specific perspective, it may be that some individuals have a poor understanding of associativity, or do not understand the principle at all. In other words, they may not understand that because some operations are related (e.g. addition and subtraction), they can be solved in different orders. Alternatively, it may be that they apply the principle in more grounded contexts, such as with words or concrete objects (Gilmore & Bryant, 2006), but not in more abstract contexts with digits. Furthermore, even with an understanding in both contexts, an individual may still choose to operate left-to-right, for example, if they are highly proficient in calculating (Newton, Star, & Lynch, 2010), or dislike the process of re-ordering operations (Robinson & Dubé, 2012). Such domain-specific factors may therefore hinder shortcut use.

From a domain-general perspective, attention, working memory, switching and inhibition are likely to be important for shortcut use. When an individual encounters a novel problem such as 'a + b − c', they may initially begin using a left-to-right strategy, a strategy that must be inhibited in order to use the shortcut (Robinson & Dubé, 2013). Separately, if the problem that follows has a different structure (e.g. 'a − b + c'), the individual must then switch to a strategy that they did not use before.

DOI: 10.1177/1747021820905739

In everyday settings outside of the classroom, the arithmetic problems that individuals encounter are likely to be varied. Likewise, in research studies different arithmetic problems are usually presented in an intermixed manner. The result is that the most efficient strategy for solving the current problem could be completely different to the most efficient strategy for the previous problem, and therefore require switching skills to identify and execute it. Thus, in everyday life and in experimental studies, individuals are often required to hold multiple strategies in mind, inhibit default procedures, and to switch between them (Lemaire & Lecacheur, 2010; 2011). A failure in any one of these domain-general skills may therefore prevent shortcut use.

Attention

Our research focused on the domain-general construct of attention because more than any other, it has frequently been suggested as important for identifying and executing conceptually-derived strategies. Attention consists of multiple components (Petersen & Posner, 2012; Robertson, Ward, Ridgeway, & Nimmo-Smith, 1996). Selective and spatial attention are relevant here because they are both likely to be required for using associativity shortcuts on visually-presented digit problems such as “26 + 48 − 45”. Selective attention refers to the prioritised processing of certain stimuli (Zentall, 2005), such as a target word embedded among distractors (Johnston & Dark, 1986). Spatial attention refers to the prioritised processing of information at a relevant location (Kim & Cave, 1995) such as looking to the left or right in response to a sound being presented from that direction. Our goal was not to distinguish between selective and spatial attention as we would expect them to have a similar role in shortcut use, i.e. directing attention to the right-hand side and selecting ‘b − c’. Instead, our goal was to investigate the question of whether visual attention, as a global construct, is involved in associativity shortcut use.

Both theoretical and empirical work provide some indication that attention may be required to identify the associativity shortcut. Theoretically, models in the strategy literature (Lovett &

DOI: 10.1177/1747021820905739

Anderson, 1996; Payne, Bettman, & Johnson, 1993; Rieskamp & Otto, 2006) highlight the cognitive processes that might be required for using different arithmetic strategies. Of particular relevance is SCADS*, the Strategy Choice and Discovery Simulation Model* (Shrager & Siegler, 1998; Siegler & Araya, 2005), which was designed to predict the discovery of inversion shortcuts ('a + b − b'). SCADS* suggests that six processes are required for discovering the shortcut, the first of which is the deployment of attention to the right-hand side to encode 'b − b'. After discovery, the shortcut is primed for use such that subsequent trials require less attention to identify it. SCADS* has since been extrapolated to associativity (Robinson & LeFevre, 2012, p413), suggesting that the same mechanisms, including attention, may apply.

Empirically, three strands of research provide some, preliminary evidence that attention could be important for identifying associativity shortcuts (Landy & Goldstone, 2007a; Dubé & Robinson, 2010; Eaves, Attridge, & Gilmore, 2019). In the experimental studies by Landy & Goldstone (2007a, 2007b, 2010) adults validated the equivalence of multi-term problems, and solved multi-term problems such as '2 + 3 × 4' in conditions where the spacing within and between the operations was manipulated. When the operation with precedence (in this case multiplication) had narrow spacing (i.e. '2   +   3 × 4'), individuals solved problems more accurately, quickly, and made fewer precedence errors, compared to a condition with wide spacing (i.e. 2 + 3   ×   4). Gestalt principles (Wertheimer, 1923) were used to explain their findings: items that are close in proximity are more likely to be grouped together, and selective attention directed to them as a whole 'object'. Items further apart are more likely to be perceived as separate units, and not attended to as a whole object. Perceptually-driven biases of attention may therefore influence the order in which arithmetic operations are performed.

Second, in a classroom intervention study with adults, Eaves et al. (2019) found that individuals who solved 'a + b − b' inversion problems were more likely than individuals who solved 'a + b − a'

DOI: 10.1177/1747021820905739

inversion problems to subsequently use the associativity shortcut on ‘a + b − c’ problems. They proposed two mechanisms through which ‘a + b − b’ inversion problems helped individuals to identify the associativity shortcut. One was an attention mechanism, where ‘b − b’ directed attention to the location of the associativity shortcut, and one was a strategy validation mechanism, where performing ‘b − b’ first implicitly communicated that a right-to-left strategy was a valid approach on ‘a + b − c’ associativity problems. However, they could not determine from their data which (if either) of these mechanisms was more likely.

Finally, the experiment by Dubé & Robinson (2010) is the most relevant, because they directly manipulated spatial attention while adults solved inversion and associativity problems. In a between-subjects design, individuals were either primed to look to the left or right of the problem. In the left-prime condition, the left-most digits (e.g. 7 × 9 in the problem 7 × 9 ÷ 3) appeared on the screen 250ms before the whole problem and in the right-prime condition the right-most digits (e.g. 9 ÷ 3 in the problem 7 × 9 ÷ 3) appeared on the screen 250ms before the whole problem. They found that for inversion problems, accuracy and response time were better for those in the right-prime condition than the left-prime condition, but there was no difference between the conditions on associativity problems. They suggested that attention was involved in identifying inversion shortcuts, but that it was difficult to interpret the results on associativity problems because the shortcut was infrequently used (p.64).

Further research is therefore warranted to investigate the role of attention in associativity shortcut use. While there is a theoretical rationale for the role of attention, the empirical evidence to date is preliminary and inconclusive. Our studies are the first thorough investigation into whether attention enables the identification of the conceptual associativity strategy on ‘a + b − c’ problems.

DOI: 10.1177/1747021820905739

How to measure the role of attention in shortcut use

Implicit in the theories and studies above is an important distinction between shortcut identification and shortcut use. Identification we refer to as the processes involved in using the shortcut for the first time on a task or in a situation. Shortcut use we refer to as the processes involved in executing the shortcut after it has been identified in that task or situation. We argue that attention is important for shortcut identification. After the shortcut has been identified, the demand for attention on problems that are conducive to a shortcut may be less because the strategy can be executed in a predictable manner. In other words, attention may play an important role in the time leading up to the identification of the associativity shortcut, less so thereafter.

As previously mentioned, there is a need to develop new measures of conceptual understanding and strategy use. For our research, we wanted to investigate the role of attention specifically in shortcut identification, and to do so, we needed a method that could separate identification from use. In two of the above studies (Landy & Goldstone, 2007; Dubé & Robinson, 2010), like most others in the literature, performance was measured by averaging data over many trials. This approach captures the processes involved in both identification and use; an individual's accuracy and RT across all trials will reflect several factors including whether they identified a shortcut, when they identified it, and how well they executed the strategies both before and after identification. It may therefore be that attention is involved in shortcut identification, but that because average accuracy and RT measure both shortcut identification and use, they are not sensitive enough to detect its role. For inversion problems (Dubé & Robinson, 2010) this is less of an issue, as the two processes (identification and use) are more closely related. Once an individual has identified an inversion shortcut, there is nothing left to compute (i.e. there is nothing to 'use'). For associativity problems however, there is a clearer distinction because applying the shortcut involves more steps; an individual must first identify the shortcut and then use it, and the former does not guarantee the latter.

DOI: 10.1177/1747021820905739

We therefore built a new tool to measure identification, hereafter the 'identification analytic'. The tool follows a microgenetic approach, an approach used for studying rapid changes in development by conducting high density observations in a narrow time-period (Siegler & Svetina, 2006). Microgenetic approaches afford the advantage of capturing precisely how and when change occurs (Siegler, 1995), which is what we wanted to achieve.

We also wanted to capture the time-point of identification without asking individuals to describe the strategy that they had used after every single trial. As some have suggested (Haider, Gaschler, Vaterrodt, & Frensch, 2014; Watchorn et al., 2014), repeatedly asking an individual about how they were solving a problem may provide a hint that an alternative, more sophisticated strategy exists, which encourages them to identify it. Meta-cognitive studies provide some support for this (Flavell, 1979), where encouraging individuals to reflect upon, and be consciously aware of the cognitive processes they are performing benefits their strategy performance (Carr & Jessup, 1995; Schoenfeld, 1985; Ghatala, Levin, Pressley, & Goodwin, 1986; Mevarech & Fridkin, 2006; Mevarech & Kramarski, 1997; Dewolf, Van Dooren, Ev Cimen, & Verschaffel, 2014; Babai, Shalev, & Stavy, 2015). We therefore wanted to capture identification as it naturally occurred, i.e. without encouraging or inducing discovery through subtle hints or self-reflection.

The identification analytic is based on a technique used in the insight literature that was designed to capture if and when individuals identified a simpler way for solving 'number reduction' problems (Haider & Rose, 2007). Haider & Rose (2007) recorded the time participants took to complete problems where sequences of digits were presented one at a time, and participants were asked to identify what digit should come next based on 'rules' they had been taught prior to the experiment. What participants were not told was that some positions in the sequence were predictable and did not require using the rules to deduce the next number, while other positions were unpredictable and did require using the rules. If participants identified the underlying regularity, they could

DOI: 10.1177/1747021820905739

respond efficiently at the predictable points in the sequence, i.e. without thinking about the rules they were given. Haider & Rose (2007) compared participants' response times at the predictable and unpredictable positions in the sequence in live-time. If and when median response times at predictable positions fell below the confidence interval of the mean response time at unpredictable positions, participants were assessed to have discovered a regularity in the number sequence.

We applied this logic to associativity problems to create a novel method that captures strategy identification using implicit, response time data, which is then corroborated by an explicit self-report at the end of the task. During the task, response time is compared between two types of problems, those that encourage shortcuts (conducive) and those that do not encourage shortcuts (non-conducive), on a trial-by-trial basis. These two types of problems should be solved with similar speed if they are solved through the same strategy. However, if a shortcut is identified and used, conducive problems should be solved much more quickly than non-conducive problems. By comparing RT on a trial-by-trial basis, the identification analytic detects when a difference between the problem types emerges, and thus, when an individual first identifies and begins to use the shortcut.

The present studies

We aimed to investigate whether attention was involved in identifying associativity shortcuts by developing and implementing a new measure of conceptual-shortcut identification. More specifically, we investigated whether the promotion of attention to the right-hand side of associativity problems (the location of the 'b − c' shortcut) could influence the number of individuals who identified it and the time-point at which it was identified. Our measure more cleanly separates identification from use than the measures employed by Dubé & Robinson (2010) and thus, our studies are a stronger test of the hypothesis that attention is important in the identification of the associativity shortcut.

DOI: 10.1177/1747021820905739

Experiment 1

Method

Both studies were approved by the (name removed for blind review) University's Ethics Approvals (Human Participants) Sub-committee (reference numbers C17-42, C17-70). Before the data were collected the hypotheses, designs, sample sizes, exclusion criteria and analysis plans were pre-registered at https://aspredicted.org. The pre-registrations are available at http://aspredicted.org/blind.php?x=jm8vs9 and http://aspredicted.org/blind.php?x=8q6zz7. The scripts to run the experiments can be found at (https://doi.org/ 10.17028/rd.lboro.7533770 and https://doi.org/10.17028/rd.lboro.7533794).

Participants

108 adults aged 18 – 59 years (71 female, 37 male, *M* = 28.71 years, *SD* = 10.89) participated. This sample size provides 80.37% power to detect a medium-sized effect in a chi-square analysis of three conditions. All participants were proficient in English and were not studying for, or had not studied for, a mathematics degree.

Participants were categorised into two groups based on how long they had studied mathematics: In the UK, all individuals study mathematics up to the age of 16 years (GCSE qualification), and some choose to study for one or two years more (A-level qualification). For consistency across different qualification systems, individuals who studied mathematics up to and including the age of 16 years were classed as GCSE achievers and those who studied mathematics beyond age 16 were classed as A-level achievers. All participants were naïve to the purpose of the experiment and were reimbursed for their time.

DOI: 10.1177/1747021820905739

Design

A between-subjects design was used: participants completed one of three conditions, left-prime, right-prime or control (no prime) conditions. Participants were assigned to the conditions through blocked random assignment: Two lists (one for A-level, one for GCSE-level achievers) of the numbers 1, 2 and 3 (representing the three conditions) were created, and within each list the numbers were ordered randomly within blocks of three. This ensured that the number of participants, and the proportion of A-level to GCSE-level achievers was equal in each condition.

Materials and procedure

Our experiment was based on the method described by Dubé & Robinson (2010). Each trial began with a central fixation cross (500ms) followed by a three-term arithmetic problem ('a + b − c'). In the left-prime condition, the 'a + b' operation was presented for 250ms, followed by the whole problem. In the right-prime condition, 'b − c' was presented for 250ms, followed by the whole problem. In the control condition there was no prime and the entire problem was presented at the same time (Figure 1). The experiment was run on a 15" laptop, and responses were made using the in-built keyboard and an external USB-keypad. Audio was presented to the experimenter through headphones (Sony MDRZX310) and a USB-dictaphone was used to record participants' verbal responses to interview questions that were asked at the end of the experiment.

(Insert Figure 1 here)

Participants were told that they would be presented with approximately 40 mathematics problems, and that their task was to solve each problem in their head and to say their answer out loud. They were asked to press the spacebar at the same time as vocalising their answer, and afterwards the experimenter would enter their response via the USB-keypad. They were verbally told that both

DOI: 10.1177/1747021820905739

accuracy and time were recorded, but not to worry or panic, and to just try their best. Strategies were not mentioned in the instructions and feedback was not given.

Each problem remained on the screen until participants responded. Participants completed three practice trials with non-conducive stimuli before commencing the experiment, to familiarise themselves with the equipment, procedure and task. The non-conducive stimuli increased in difficulty as the practice trials progressed (‘9 + 2 − 5’, ‘3 + 19 − 8’ and ‘39 + 14 − 27’) and primes were not presented.

At the end of the experiment, participants were interviewed about the strategies they used to solve the problems (https://doi.org/10.17028/rd.lboro.7533764)). First, they were asked “Can you tell me about how you were solving the problems?”, to which they responded unprompted until they had finished. Most people could be categorised from this response. If their description was insufficient or unclear, they were asked to “Describe in more detail”, or to think of an example to help describe what they were saying. Participants were then asked questions about their strategy preference; they were shown the written problem “46 + 38 − 35” and the left-to-right and right-to-left strategy were described in writing and verbally by the experimenter. They were asked to select the strategy they preferred and explain why. They were then asked questions about their reason(s) for use or non-use of the shortcut including “Did you ever use a right-to-left strategy?”, “If not, why not?”, “If yes, approximately when did you start to use the strategy? Finally, they were asked whether they had been aware of the prime and/or spacing manipulation.

Stimuli

Three-term arithmetic problems of the form ‘a + b − c’ were created, half of which were conducive to an associativity shortcut and half of which were non-conducive. 46 problems were selected for use in both studies (see https://doi.org/10.17028/rd.lboro.7533764) that met pre-defined criteria (below). The problems were presented in the same order for all participants.

DOI: 10.1177/1747021820905739

Conducive problems

Conducive stimuli (e.g. ‘43 + 38 − 35’) were designed to be similar in difficulty to each other, and to have variety. To ensure that the stimuli were similar in difficulty, the following criteria were imposed:

- ‘a’, ‘b’ and ‘c’ were all double-digits.
- The right-hand side, ‘b − c’, resulted in a small positive integer (2 to 6), which was always smaller than the smallest addend (‘a’ or ‘b’).
- ‘b − c’ did not involve a decade boundary cross or a borrow operation.
- The left-hand side, ‘a + b’, resulted in a double-digit number whose calculation involved a decade boundary cross and a carry operation.

Non-conducive problems

For each conducive stimulus, a non-conducive stimulus was created (e.g. ‘58 + 23 − 35’). Stimuli were therefore made in pairs, to ensure that they were of similar difficulty assuming a left-to-right procedure[1]. For example, the counterpart for the conducive stimulus ‘44 + 38 − 33’ was ‘58 + 24 − 33’. Non-conducive stimuli were defined as follows:

- The sum of the interim addition (‘a + b’) and the value of the subtrahend (‘c’) matched conducive stimuli.
- ‘a + b’ involved a decade boundary cross and a carry operation.
- ‘b − c’ involved a decade boundary cross.
- The size of ‘b − c’ ranged between -7 to -19 and between +19 to +24.
- The size of ‘a − c’ ranged between -7 to -18 and between +18 to +25.

For both the conducive and non-conducive stimuli, none of the operands contained a zero or 1, and none were identical. None of the answers to the problems, or interim solutions (‘a + b’ and ‘b − c’)

[1] Post-hoc analyses of accuracy and RT confirmed that the conducive and non-conducive problems were of statistically equivalent difficulty if solved through a left-to-right strategy.

DOI: 10.1177/1747021820905739

equalled 0 or a decade boundary. The sum of the interim addition (a + b) ranged from 61 − 84, and 'c' ranged from 23 − 47. The sum of the interim addition and the value of the subtraction were not paired together more than once, and no more than two problems had the same answer.

Building a measure of identification

We built a tool, the 'identification analytic' in PsychoPy (Peirce et al., 2019) to detect when individuals first switched from solving the problems using a left-to-right procedure to using an associativity shortcut. This switch is referred to as the 'Identification Point' (IP), depicted in Figure 2. The analytic was based on techniques used in implicit learning studies (Haider & Rose, 2007), where changes in behaviour have been inferred from trial-by-trial analyses of response times.

(Insert Figure 2 here)

On each trial, response time (RT) was recorded. On the non-conducive trials, rolling means, standard deviations and 99.9% confidence intervals of the non-conducive RT data were calculated. For the first conducive trial, absolute RT was recorded. For the second, the mean RT was calculated and recorded. For the third conducive trial onwards, the median RT of the three most recent conducive trials was calculated and recorded (henceforth a 'lag-3 median filter'). When the conducive median (or mean on the second stimulus) fell below the lower-endpoint of the confidence interval of the non-conducive mean, a 'trigger' was generated. With a lag-3 median, triggers only occurred if two of three RTs were below the confidence interval. Three consecutive triggers defined the IP.

Pilot study

The analytic was piloted on 42 individuals, using the procedure described above. Twenty of the participants were told of the two main strategies that could be used to solve 'a + b − c' stimuli. They were instructed to use a left-to-right procedure on each trial up to a point in the trial sequence when

DOI: 10.1177/1747021820905739

the experimenter would say the word 'switch'. At this point, they were asked to use the shortcut. Twenty-two participants were uninstructed in how to solve 'a + b − c' stimuli, and they completed the trials without any information on strategy use. During the trial sequence, an audible tone played to the experimenter through headphones whenever a trigger occurred. Three consecutive triggers (the IP) signalled to the experimenter that the participant may have switched strategy and that they could interrupt them to ask questions. Upon interruption, participants were asked to describe how they had been solving the problems, whether they had changed strategy, and if so, how and when they had changed strategy. If they were not interrupted, they were asked the questions at the end of the task. Note that participants were not interrupted during Experiments 1 and 2, only during the pilot testing.

The pilot data were explored to investigate whether participants' self-reports corresponded to their IP recorded by the analytic. The analytic correctly identified a strategy switch for all 20 instructed participants, although 3 of the participants had IPs before the instruction to switch. For the naïve participants, it correctly identified a strategy switch for 14 individuals and it did not 'miss' a strategy switch (i.e. there was no-one who self-reported switching but did not have an IP). Two naïve participants had IPs much before the timepoint that they estimated their switch to have occurred, and four had IPs but self-reported that they did not identify the shortcut (hereafter 'false alarms'). It could be that these false alarms are people who used the shortcut unconsciously or did not provide an accurate self-report. However, participants were always able to describe the strategy they had used clearly, in detail, and with certainty. We therefore judge their self-reports to be reliable. It may alternatively be the case that they became much faster on conducive problems without changing strategy and this led to the analytic giving an IP despite the participants' strategy not changing.

In an attempt to reduce the number of false alarms, we refined the criteria of the IP. The IP was defined as the trial number of the first of three consecutive conducive trials with an RT median (or

DOI: 10.1177/1747021820905739

mean on the second stimulus) that was 1) below the 99.9% confidence interval of the rolling non-conducive mean, 2) at least 20% less than the rolling non-conducive mean, and 3) at least 20% less than the final non-conducive mean (the mean after all the trials had been presented). With these criteria, there were 20 opportunities for an IP to be identified, with the earliest possible IP on the second conducive stimulus and the latest possible IP on the 21st conducive stimulus (participants who identify on the first conducive stimulus would therefore be detected by the analytic and assigned an IP on the 2nd conducive stimulus). After implementing the revised criteria on the pilot data, two of the false alarms (i.e. two people who self-reported not identifying the shortcut but who did have an IP) were correctly classed as non-identifiers, and all of the identifiers were still correctly categorised as identifiers (i.e. no ‘misses’ were incurred with the additional criteria).

The pilot study therefore demonstrated that the analytic correctly classified most individuals as identifiers or non-identifiers of the shortcut strategy and for the identifiers it captured the point of shortcut identification, making it suitable for use in the experiment. However, because there was a risk of some false alarms, we combined the analytic with a post-task interview about strategy use in the full experiments (Experiments 1 and 2).

Outcome measures

There were six dependent variables, 1) categorisation as an identifier or non-identifier, 2) the trial number of identification, 3) accuracy of all of the problems solved (i.e. problems both before and after the IP), 4) response time of all of the problems solved (i.e. problems both before and after the IP), 5) accuracy of the problems solved after the IP and 6) response time of the problems solved after the IP.

DOI: 10.1177/1747021820905739

Results

In this experiment, participants solved a series of "a + b − c" problems in one of three conditions. In two of the conditions, their attention was biased to the left or right using a 250ms prime ('left-prime' and 'right-prime' condition respectively). In the other condition no prime was presented ('control' condition). We begin by presenting the results of our pre-registered analyses, followed by our exploratory analyses that were not pre-registered. We then briefly discuss our results, and our rationale for Experiment 2. The data for Experiment 1 can be found at (https://doi.org/10.17028/rd.lboro.7533755.

Pre-registered analyses

It was hypothesised that 1) the number of identifiers and non-identifiers would differ between the conditions, 2) for those who identified the shortcut, there would be a difference between the conditions in the trial number on which it was identified, 3) when reaction time was averaged across all trials, there would be a difference between the conditions, and 4) when reaction time was averaged across trials after the identification point, there would be no difference between the conditions. Analyses for each of these hypotheses are now presented.

Our hypotheses derive from the fact that we thought attention enabled the *identification* of the associativity shortcut, and that attention was less important in the use of the shortcut after it had been identified. This argument forms the distinction between hypotheses 3 and 4. If attention was more important for identification than use, we would expect a difference in accuracy and RT between the conditions when we included all of the trials (hypothesis 3) because there would be more trials on which the shortcut had been used in the right-prime condition than in the control condition. We would not expect this when analysing the trials after the IP, as accuracy and RT on

DOI: 10.1177/1747021820905739

those trials just reflect use (hypothesis 4), which we would expect to be equally efficient across conditions. The analysis for hypothesis 3 also allows us to compare our results to those of Dubé & Robinson (2010), who analysed accuracy and RT across intermixed blocks of trials.

Number of identifiers

Participants were classed as identifiers if they had an IP and self-reported using the shortcut when questioned at the end of the experiment. If not, they were classed as non-identifiers. A small number of false alarms did occur: these were 15 occasions where an IP occurred, but the individual did not self-report identifying the shortcut, so they were classed as non-identifiers[2]. There was one person who self-reported identifying the shortcut on the last trial of the experiment and did not therefore have an IP; due to the criteria of the analytic, a trigger could not, by definition, occur on the last two stimuli. This participant was classed as a non-identifier.

Table 1 displays the frequencies of identifiers and non-identifiers. A 3*2 chi-square test found that the frequency of identifiers to non-identifiers was not significantly different between the conditions, X^2 (2, N = 108) = 3.92, p = 0.141, Cramers V = 0.19.

(Insert Table 1 here)

Trial numbers of the identification point

For those who were classed as identifiers, the trial number of their identification point was analysed. Table 2 displays the results.

(Insert Table 2 here)

[2] The results are unchanged if these individuals are classed as identifiers

DOI: 10.1177/1747021820905739

The trial number data in the right-prime condition were significantly positively skewed. A Kruskal-Wallis test was therefore performed, which had three levels (left-prime, right-prime, control). No significant difference was found between the conditions in the conducive trial number of identification, X^2 (2, N = 61) = 3.67, p = 0.159, or the total trial number of identification, X^2 (2, N = 61) = 3.67, p = 0.159.

Response time across all trials

For those who were classed as identifiers, median RT of the correctly solved problems across all trials were calculated (Table 3).

(Insert Table 3 here)

A 3*2 mixed ANOVA was performed on median RTs, with condition (left-prime, right-prime, control) and problem type (conducive, non-conducive) as between and within-subject factors respectively. There was a main effect of problem type, $F(1, 58) = 159.68$, $p<0.001$, $\eta_p^2 = 0.73$, where conducive problems were solved quicker than non-conducive problems. There was no main effect of condition, $F(2, 58) = 0.99$, $p = 0.377$, $\eta_p^2 = 0.03$, and no significant interaction between problem type and condition, $F(2, 58) = 1.00$, $p = 0.373$, $\eta_p^2 = 0.03$. This analysis was repeated using all participants' data, identifiers and non-identifiers, and the same result emerged. There was a main effect of problem type, $F(1, 105) = 65.74$, $p<0.001$, $\eta_p^2 = 0.39$, no main effect of condition, $F(2, 105) = 1.16$, $p = 0.318$, $\eta_p^2 = 0.02$, and no significant interaction between problem type and condition, $F(2, 105) = 0.17$, $p = 0.842$, $\eta_p^2 < 0.01$.

DOI: 10.1177/1747021820905739

Accuracy across all the trials [3]

For those who were classed as identifiers, the mean percent of correctly solved problems across all trials were calculated. Accuracy data are displayed in Table 3. A 3*2 mixed ANOVA was performed, with condition (left-prime, right-prime, control) and problem type (conducive, non-conducive) as between and within-subject factors respectively. There was a main effect of problem type, $F(1, 58) = 32.67$, $p<0.001$, $\eta_p^2 = 0.36$, where conducive problems were solved more accurately than non-conducive problems. There was no main effect of condition, $F(2, 58) = 0.41$, $p = 0.665$, $\eta_p^2 = 0.01$, and no interaction between problem type and condition, $F(2, 58) = 0.09$, $p = 0.910$, $\eta_p^2 < 0.01$. This analysis was repeated using all participants' data, identifiers and non-identifiers, and the same result emerged. There was a main effect of problem type, $F(1, 105) = 15.65$, $p<0.001$, $\eta_p^2 = 0.13$, no main effect of condition, $F(2, 105) = 0.38$, $p = 0.682$, $\eta_p^2 = 0.01$, and no significant interaction between problem type and condition, $F(2, 105) = 0.56$, $p = 0.572$, $\eta_p^2 = 0.01$.

Response time for the trials after the IP

For those who were classed as identifiers, median RT of the correctly solved trials after the IP were calculated. The data were analysed for a) statistical equivalence and b) statistical difference, between the three conditions (left-prime, right-prime, control) separately for each problem type (conducive, non-conducive). For each analysis, the alpha levels were adjusted for the number of comparisons (0.05/6 = 0.008) and the result can be found in Supplementary Material A (Table A1).

To test for equivalence, we used the two one-sided tests (TOST) procedure described by Lakens (2017). This test identifies whether the data significantly support the null hypothesis. Non-significant results using null hypothesis significance tests do not do this because a failure to accept the alternative hypothesis does not evidence that the null hypothesis is likely to be true. Equivalence tests however, are one approach that can (Harms & Lakens, 2018). The TOST procedure works by

[3] We pre-registered analysis on RT but not accuracy. However, we present the two analyses together for readability. The outcomes of both analyses are the same.

DOI: 10.1177/1747021820905739

specifying an upper and lower equivalence bound based on the smallest effect size of interest and uses two t-tests to assess whether the observed effect falls above the lower-bound and below the upper-bound. If it does, the difference between the two groups or conditions from which the effect was derived are deemed to be statistically equivalent.

In our analyses, the upper and lower confidence bounds (Cohens d) were set to ±0.8 for each pairwise comparison, to give a large range within which our results could fall, i.e. a liberal criterion of equivalence. For statistical difference, independent t-tests were performed. A summary of the output can be found in the Supplementary Material. None of the comparisons were statistically equivalent (p>0.008 for all comparisons) and none were significantly different (p>0.008 for all comparisons). Collectively, this is an 'undetermined' outcome (Lakens, 2017): The conditions were not significantly different or statistically equivalent.

Accuracy for the trials after the IP [4]

For those who were classed as identifiers, accuracy of the trials after the IP were calculated. The data were also analysed for a) statistical equivalence and b) statistical difference, in the same way that the RT data were. For each analysis, the alpha levels were adjusted for the number of comparisons (0.05/6 = 0.008). None of the comparisons were statistically equivalent (p > 0.008 for all comparisons) and none were significantly different (p>0.008 for all comparisons, see Table A2 in Supplementary Material A). As per the RT data, this is an 'undetermined' outcome (Lakens, 2017).

Validity of the analytic

To establish the validity of the analytic, median correct RTs were compared before and after the IP for participants who identified the shortcut. If the analytic correctly identifies the point where participants switched to using an associativity shortcut, we would expect median RT for the

[4] We pre-registered analysis on RT but not accuracy. However, we present the two analyses together for readability. The outcomes of both analyses are the same.

DOI: 10.1177/1747021820905739

conducive problems to significantly differ before and after the IP. We would expect this difference in RT to be significantly larger than any difference on the non-conducive problems. A 2*2 within-subjects ANOVA was performed, with problem type (conducive, non-conducive) and time (before IP, after IP) as the factors. There was a main effect of problem type, $F(1, 54) = 126.44$, $p<0.001$, $\eta_p^2 = 0.70$, where conducive problems were solved quicker than non-conducive problems. There was a main effect of time, $F(1, 54) = 60.85$, $p<0.001$, $\eta_p^2 = 0.53$, where problems were solved quicker after the IP, and there was a significant interaction between problem type and time, $F(1, 54) = 20.63$, $p<0.001$, $\eta_p^2 = 0.28$. The interaction indicates that the improvement in RT after the IP was significantly larger for conducive problems (circa 5.3 seconds) than for non-conducive problems (circa 1.8 seconds).

Exploratory analyses (not pre-registered)

To explore the extent to which the data supported our hypotheses regarding the difference between the priming conditions, all hypotheses were tested with Bayesian analyses and all showed anecdotal to strong evidence (Jeffreys, 1961) in favour of the null hypothesis (BF_{10}’s from 0.09 to 0.49).

Discussion

Participants solved three-term problems either in a condition where their attention was manipulated to the left, to the right, or in no direction. Attention was biased by presenting “a + b” or “b − c” for 250ms before the onset of the whole problem. Our theory was that attention would be important for identifying the associativity shortcut and less important for using it. In other words, we thought that attention would affect whether and when a person first identified the shortcut, but it would have little effect on the accuracy and speed with which they executed it after they had done so. We therefore measured performance in terms of whether the individual was an identifier or not,

DOI: 10.1177/1747021820905739

the trial number of identification (if they were an identifier) and the accuracy and RT with which they solved the problems. We hypothesised that there would be differences among the conditions for the variables that captured more of the processes involved in identification (number of identifiers, trial number of identification and RT for all the problems correctly solved). We hypothesised that there would be no difference among the conditions for the variables that captured the processes involved in using the shortcut (RT after the point of identification).

Contrary to our hypotheses, we found no evidence that attention influenced the number of people who identified the shortcut, the trial number on which they identified it, or the accuracy and RT for solving the problems. Thus, attention may not be as important a factor in identification as we originally thought. However, we reasoned that attention might still influence shortcut identification, and that our study might not have been powerful enough to detect it. In particular, the attention manipulation (250ms prime) was relatively subtle and may have been insufficient to change the behaviour of enough participants. Indeed, in some of the post-experiment interviews, some participants voluntarily commented that the prime had influenced the strategy they used to solve the conducive problems, while others commented that they did not notice it. Thus, the attention manipulation may have encouraged identification for some, but not enough people. In Experiment 2 we improved this by using a stronger attention manipulation.

Experiment 2

Method

Participants

108 adults aged 18 – 56 years (76 female, 32 male, *M*= 22.56, *SD* = 7.59) participated, using the same recruitment criteria as Experiment 1.

DOI: 10.1177/1747021820905739

Design

A between-subjects design was used. Participants completed one of three conditions, the left-prime, right-prime or control (no-prime) condition. As per Experiment 1, participants were assigned to the conditions through blocked random assignment, and the proportion of A-level to GCSE-level achievers was equal in each condition.

Materials and procedure

The procedure was the same as Experiment 1. Participants mentally solved the conducive and non-conducive problems, vocalised their answer out loud and pressed the spacebar at the same time. The experimenter recorded their answers using an external USB-keypad. The stimuli were presented in the same order as in Experiment 1, except for one stimulus which was moved to earlier in the trial sequence in an attempt to reduce the false alarm rate of Experiment 1.

To create stronger attentional biases than Experiment 1, the left-prime and right-prime conditions were created by manipulating three factors. First, as per Experiment 1, the temporal prime ('a + b' or 'b − c') was presented for 250ms before the onset of the whole problem in the left and right conditions respectively. Second, the spacing of the subexpressions were altered. In the left-prime condition, the spacing within 'a + b' was narrow (1 space, 6mm) and the spacing within 'b − c' was wide (3 spaces, 18mm), while in the right-prime condition, 'b − c' was narrowly spaced, and 'a + b' widely spaced. Finally, the position of the '+' and '−' were shifted: the '−' was shifted towards the 'c' in the left-prime condition (e.g. '42 + 39   − 33'), while '+' was shifted towards the 'a' in the right-prime condition (e.g. '42 +   39 − 33'). In all, these manipulations had the perceptual effect of making the 'a + b' or 'b − c' more salient. Similar to previous studies (Landy & Goldstone, 2007, 2010; Landy et al., 2008), the digits were 1cm in height and 6-7mm in width, and participants were sat approximately 55cm from the screen.

DOI: 10.1177/1747021820905739

Outcome measures

There were three pre-registered primary dependent variables, 1) categorisation as an identifier or non-identifier, 2) categorisation as an early identifier or a not an early identifier, and 3) the percent of conducive trials remaining after the IP. Accuracy and median RT of the correctly solved problems for all the trials presented were also recorded.

The early-identifier and percent use variables are new, in that we did not use them in Experiment 1. We incorporated them in Experiment 2 because we thought they might be more sensitive to the attention manipulation. For example, we thought attention might be more important for early identification rather than identification at any point during the trial sequence, where an 'early identifier' is someone who identifies the shortcut on the second or third conducive stimulus. We also thought that if we could include non-identifiers in our measure of trial number, the variable would be more informative and might be more likely to show an effect of the attention manipulation. We therefore created a different DV, where trial number was converted to the percent of trials remaining after the point of identification. This allowed for all participants to be included and made the measure continuous.

Results

We present the results of our pre-registered inferential and Bayesian analyses and then briefly discuss our results. The data for Experiment 2 can be found at https://doi.org/10.17028/rd.lboro.7533755.

Again, our rationale was that attention would be important for the identification of the associativity shortcut, less so for using the shortcut thereafter. We hypothesised that for the variables which captured more of the processes of identification, there would be a difference between the

DOI: 10.1177/1747021820905739

conditions. More specifically, 1) the frequency of identifiers and non-identifiers would differ between the conditions, 2) the number of early identifiers and non-early identifiers would differ between the conditions and 3) the percent of trials after the IP would differ between the conditions. Because accuracy and RT across all of the trials capture both processes, we also expected these to be significantly different between the conditions. We now present our analyses for each hypothesis.

Number of identifiers

Participants were classed as identifiers or non-identifiers and Table 4 displays the result.

(Insert Table 4 here)

A 3*2 chi-square test found that the frequencies of identifiers to non-identifiers was not significantly different between the conditions, X^2 (2, N = 108) = 0.52, p = 0.772, Cramers V = 0.07. Bayesian analyses indicated that there was moderate evidence in support of the null hypothesis, with a BF_{10} of 0.09.

It should be noted that the false alarm and miss rates in Experiment 2 were similar to Experiment 1, with 13 false alarms[5] and 1 miss in Experiment 2 (there were 15 false alarms and no misses in Experiment 1).

Number of early identifiers

There were five early identifiers each in the left-prime and right-prime conditions, and three in the control condition. These frequencies were not significantly different, X^2 (2, N = 108) = 0.70, p = 0.705, Cramers V = 0.08. Bayesian analyses indicated that there was moderate evidence in support of the null hypothesis, with a BF_{10} of 0.04.

[5] The results are unchanged if these individuals are classed as identifiers

DOI: 10.1177/1747021820905739

Percent of conducive trials after the IP

The data were not normally distributed, and a Kruskal-Wallis test using all participants' data (identifiers and non-identifiers) found no significant difference between the conditions, X^2 (2) = 0.38, p = 0.827. The mean percent of conducive trials after the IP was 44.07 (SD = 41.95) for the right-prime condition, 40.53 (SD = 42.58) for the left-prime condition and 36.24 (SD = 41.92) for the control condition. A one-way between-subjects Bayesian ANOVA indicated that there was moderate evidence in support of the null hypothesis, with a BF_{10} of 0.11.

RT and accuracy for all of the trials

Median RT of the correctly solved problems and mean accuracy were analysed for all of the trials presented for all of the participants (identifiers and non-identifiers). For RT, a 3*2 mixed ANOVA with condition (left-prime, right-prime, control) and problem type (conducive, non-conducive) as between and within-subject factors respectively identified a main effect of problem type, $F(1, 105)$ = 61.96, $p<0.001$, η_p^2 = 0.37, where conducive problems were solved quicker than non-conducive problems, no main effect of condition, $F(2, 105)$ = 2.16, p = 0.121, η_p^2 = 0.04, and no significant interaction between problem type and condition, $F(2, 105)$ = 0.66, p = 0.519, η_p^2 = 0.01. For accuracy, there was a main effect of problem type, $F(1, 105)$ = 20.37, $p<0.001$, η_p^2 = 0.16, where conducive problems were solved more accurately than non-conducive problems, no main effect of condition, $F(2, 105)$ = 0.66, p = 0.520, η_p^2 = 0.01, and no significant interaction between problem type and condition, $F(2, 105)$ = 0.40, p = 0.672, η_p^2 = 0.01.

Discussion

Experiment 2 used the same design and procedure as Experiment 1, but with a more powerful manipulation of attention and dependent variables that we judged to be potentially more sensitive measures of identification. We hypothesised that there would be a difference among the conditions

DOI: 10.1177/1747021820905739

in terms of the number of identifiers, the number of early identifiers, and the percent of trials that the associativity shortcut was used on (indexed by the percent of trials remaining after the IP). Accuracy and RT for all of the trials were also recorded and analysed, on the assumption that they captured some of the processes of identification and might therefore differ between the conditions.

However, Experiment 2 found no evidence that manipulating visual attention to the left or right of three-term problems influenced the identification of associativity shortcuts. The number of identifiers, the number of early identifiers, and the percent of trials remaining after the IP did not significantly differ between the conditions, and neither did accuracy or RT. Thus, in two studies we found no evidence that attention plays a role in the identification of associativity shortcuts with a moderate or larger effect size.

General discussion

We developed a new tool for measuring the application of conceptual knowledge in solving arithmetic problems. Our tool measures when an individual first identifies an arithmetic strategy on a task. In this case the strategy was an associativity shortcut and the task was one where problems that were conducive to a shortcut and problems that were non-conducive to a shortcut were presented in an intermixed order. This tool was implemented in two studies that, for the first time, captured strategy change on a trial-by-trial basis using implicit (RT) data without relying on explicit self-reports after each problem was solved. In both studies, the tool proved effective in pinpointing associativity shortcut identification, although we found no evidence that manipulations designed to draw participants' visual attention to parts of the problem changed the rates of identification. We discuss the methodological and theoretical contributions of our research below.

DOI: 10.1177/1747021820905739

Methodological contribution

A variety of explicit and implicit tasks have been used to measure the understanding of different arithmetic principles, all of which may capture subtly different facets of knowledge (Crooks & Alibali, 2014). To maximise the validity of their assessments, scholars have been encouraged to use multiple measures (Schneider & Stern, 2010), to give more thought and consideration to the method they select, and to make explicit how their tasks align with their theory and definition of the principle that they are studying (Crooks & Alibali, 2014). The new measure that we developed and implemented in the current studies is a direct response to these calls.

First, we implemented both implicit and explicit measures. Explicit measures may be restricted to those who are aware of the strategy that they used and can express that strategy accurately, while implicit methods are not. However, implicit methods can only infer strategy use and cannot guarantee it, and explicit techniques can compensate for this by providing some reassurance of strategy choice. By combining the result of the identification analytic (an implicit measure) with self-report data (an explicit measure), we drew on the benefits of both techniques.

Second, we selected dependent measures that specifically matched the mechanism through which we thought attention would operate. Our rationale was that attention was more important for determining whether and when an individual identifies the shortcut (Siegler & Araya, 2005), and less important for determining how accurately and efficiently they solved shortcut problems after they had identified it. Existing measures of shortcut use average accuracy and response time across multiple trials, which reflect the cognitive processes of both identification and use. For example, an individual's average RT over a sequence of conducive problems will reflect whether they identified the shortcut, when they identified the shortcut, and the speed with which they execute arithmetic strategies. To obtain a purer measure of identification, we therefore built a new a tool that could

DOI: 10.1177/1747021820905739

separate identification from use. As a result, we had a closer match between our outcome measures and our hypothesis.

Our tool blends the logic of the microgenetic method with theories of insight and implicit learning (Siegler, 2006; Haider & Rose, 2007), to measure the trial number on which individuals first identify and use an associativity shortcut. While a handful of microgenetic studies have tracked the development of conceptual knowledge and changes in strategy use (Rittle-Johnson, Siegler, & Alibali, 2001; Robinson & Dubé, 2009; Siegler, 1987; Geary, 1990; Siegler & Svetina, 2006), only one monitored performance with good temporal resolution, that is, on a trial-by-trial basis (Siegler & Stern, 1998). Siegler & Stern (1998) monitored the use of inversion shortcuts, based on self-reports. However, it has been suggested that providing self-reports after every problem may hint that alternative strategies exist (Haider et al., 2014; Watchorn et al., 2014) or cause participants to become doubtful of their previously used strategies (Siegler & Stern, 1998, p362). Siegler & Stern (1998) addressed this by capping the number of consecutive self-reports to three. However, to completely circumvent the issue, shortcut identification must be captured without any self-reports being given during the task. To the best of our knowledge, our tool is the first to achieve this.

Our tool could be helpful for researchers investigating the development of conceptual knowledge, strategy switching and the microgenetics of cognitive change. The analytic monitors the time taken to solve problems that can be answered primarily through two strategies, where the difference in solution time between those two strategies is not trivial. It could therefore be used to investigate other principles that lend themselves to a small number of strategies that differ in efficiency. For example, it could be used to investigate the identification of the commutativity principle on problems such as ‘7 + 19 + 3’, where a strategy of adding the first and last addend to 10 is more efficient than operating left-to-right. Or, it could be used to investigate commutativity on sequences of two-addend problems such as ‘27 + 16’ after solving ‘16 + 27’, where a strategy of thinking back to

DOI: 10.1177/1747021820905739

the previous trial negates computation on the current trial, making it substantially more efficient than calculating an answer to each problem. In any case, if solution time for a problem of interest suddenly declines, and does not immediately revert, it may be inferred that a person has switched from using one strategy to another. And if this change is specific to problems with a certain characteristic (i.e. conducive to one strategy) then the change is unlikely to be due to factors such as improved calculation speed. We also note that specifics of our metric could be tailored to researchers' individual needs. For example, the confidence interval could be calculated over different numbers of trials (e.g. over 3 trials, 10 trials, or after all trials have been presented) depending on the sensitivity and specificity of the tool that they require. Thus, our tool could be used and adapted by researchers to implicitly measure the identification of arithmetic strategies and the time-point of cognitive change.

Theoretical contribution

Our findings extend three strands of research. First, SCADS* (Siegler & Araya, 2005), a computational model intended to predict the discovery of inversion shortcuts, included attention as the first of six cognitive processes required for shortcut discovery. Our findings suggest that the claims to apply the exact same cognitive mechanisms involved on inversion problems to associativity problems may be inappropriate. Separately, multiple studies have manipulated the spacing within problems that have three or more digits, akin to our method in Experiment 2 (Jiang, Cooper, & Alibali, 2014; Landy & Goldstone, 2007; Landy et al., 2008). Those studies found that on problems such as '2 + 3 × 4', operations with narrow spaces were performed before the operations with wider spaces. Our findings suggest that similar attention manipulations do not have the same effect on 'a + b − c' problems. Finally, and most relevant to our research, is a preliminary study by Dubé & Robinson (2010), who described using an experimental manipulation that we implemented in Experiment 1. They found that priming the location of the shortcut improved RT on inversion problems but not

DOI: 10.1177/1747021820905739

associativity problems. Our findings support and extend their conclusion by being the first to rigorously test whether attention is important for identifying the associativity shortcut.

We found no evidence for the role of attention in the identification of the associativity shortcut, and our Bayesian analyses suggest that the evidence was always in favour of the null hypothesis. This null result is important to communicate for three reasons. First, the result indicates that models from the inversion literature (e.g. SCADS*) and the findings from multi-term problems more generally (Jiang et al., 2014; Landy & Goldstone, 2007; Landy et al., 2008) do not necessarily apply to 'a + b − c' associativity problems. Second, they help to answer unresolved questions in the discussion of other studies. For example, Eaves et al., (2019) proposed that attention and/or strategy validation mechanisms could explain why 'a + b − b' inversion problems promoted associativity shortcut use on 'a + b − c' problems, and the findings from this study imply that an attention mechanism is less likely. Finally, our results provide good evidence that there must be factors other than attention that act as a barrier to using the associativity shortcut (e.g. working memory, inhibition, shifting), and these factors therefore warrant further investigation.

We offer five explanations for why we found no evidence for the role of attention in the identification of the associativity shortcut. These explanations concern 1) the statistical power of the studies, 2) strength of the attention manipulation, 3) inadequate conceptual knowledge, 4) conflicting knowledge and 5) the demands of the task. These explanations are not mutually exclusive. We now discuss each of these in turn.

Statistical power

Each experiment was powered to detect a medium-size effect (W = 0.30) for a 3*2 chi-square test for the number of identifiers, which we deemed to be the smallest effect size of interest and our primary outcome measure (note that we found effect sizes of 0.19 and 0.07 in Study 1 and 2

DOI: 10.1177/1747021820905739

respectively). Each study was powered to at least 80% (Cohen, 1988) and required 108 participants, a number which we deemed to be practically achievable. One consequence is that, if the alternative hypothesis is true in the population, we had a 20% chance of failing to find that effect significant in each study (Type II error). However, by conducting two studies at the same level of power (80%), the chance of obtaining a null result in both, given that the alternative hypothesis is true in the population, declines to 4%. We therefore judge that overall our studies had enough power to detect a medium or larger effect size, and that there are better alternative explanations of our results.

Strength of the manipulation

The attention manipulations in both studies were subtle, reflected by the fact that some participants (approximately 30% in Experiment 1, and 14% in Experiment 2) did not report being aware of the manipulations when questioned at the end of the experiment. Other manipulations of attention may have a different effect. We intentionally avoided any stronger manipulation: with a longer prime duration than that used in Experiment 1, any difference between the conditions, had one been found, could have been due to the initiation of computation, rather than attention. Separately, any stronger spacing manipulation than that used in Experiment 2 could have made the stimuli look unusual, and confused participants.

We judge that our attention manipulations were sufficient for three reasons. First, most participants were aware of the manipulations (69% and 86% in Experiment 1 and 2 respectively), and re-analysis of the data excluding those who were not aware made no difference to the outcome of the primary analyses. Secondly, the same prime that we used in Study 1 did produce an effect in the study by Dubé & Robinson (2010) in the context of inversion problems (where the primes were “a + b” and “b − b”), suggesting that the attention manipulation can be effective. Finally, smaller spacing manipulations have been found to influence the order in which operations are performed (Jiang et al. 2014, p.1629), suggesting that they only need to be subtle to produce an effect. We therefore

DOI: 10.1177/1747021820905739

think that the strength of the manipulation is an unlikely explanation for why we found no difference among the conditions.

Inadequate knowledge

Our results are not directly related to the model of SCADS* (Siegler & Araya, 2005) because it was primarily developed from inversion problems, while we were concerned with associativity. Associativity and inversion are different, with the former more difficult than the latter (Robinson & Dube, 2017): individuals are much older when they start to use associativity shortcuts compared to inversion shortcuts, and are more likely to prefer using a left-to-right procedure on associativity than inversion problems (Robinson & Dubé, 2012). For a prime to facilitate shortcut identification, it seems sensible to assume that an individual must have some level of knowledge of the principle from which it is derived, because attention mechanisms alone would be unlikely to teach new strategies. Thus, attention manipulations might facilitate the identification of inversion shortcuts because adults understand the principle. In contrast, the same manipulations may not be sufficient for associativity because it is a principle that adults may not understand.

Indeed, for the one study that has included associativity problems (Dubé & Robinson, 2010), our results are entirely consistent. Dubé & Robinson (2010) found no difference in accuracy and response time on associativity problems in the domain of multiplication-division. Here, we report no difference in the domain of addition-subtraction. Using the same procedure for manipulating attention (Experiment 1), similar outcome measures (accuracy and response time), and our newly developed measure that more sensitively measures the time-point of identification, we replicated their preliminary finding that attention manipulations do not alter performance on associativity problems. It is therefore unlikely that the lack of evidence for a role of attention is due to the specific outcome measures chosen, or a Type II error in one particular study. Rather, the evidence is more consistent with the explanation that attention is less important for the identification of associativity

DOI: 10.1177/1747021820905739

shortcuts and that other factors, such as knowledge of the principle, are stronger determinants of shortcut use.

Other researchers have similarly suggested that perceptual manipulations in stimuli only help individuals with sufficient conceptual knowledge (Jiang et al., 2014). For example, in two interventions, Alibali and colleagues (Alibali, 2015; Alibali et al., 2017) aimed to improve 8 – 11 year olds' understanding of equivalence, using ink colour as a perceptual support. They found that solving equivalence problems with an equals sign presented in red ink led to better problem reconstruction and strategy generation, compared to a group without that support. However, they also noted that not all individuals benefited equally, and that perceptual support was more effective for those "close to correct performance" (Alibali et al, 2017, p.10). These individuals, they argued, required only minimal scaffolding to achieve a complete understanding of equivalence, which the perceptual manipulations provided. Thus, the individuals in our studies may have had insufficient knowledge to benefit from similar manipulations.

Conflicting knowledge

One relatively common but unexpected response to the question "Why did you not use the right-to-left strategy?" was "BODMAS" (Brackets, Order, Division, Multiplication, Addition, Subtraction), an acronym taught at around the age of 11 years in the UK to help children remember the order in which arithmetic operations should be performed. The acronym intends to highlight the precedence of multiplication and division over addition and subtraction, while not prescribing any precedence between multiplication and division, or between addition and subtraction. However, these responses indicate that some adults have a literal interpretation of BODMAS and ascribe precedence to addition over subtraction. These individuals may never use a shortcut on 'a + b − c' problems, because they believe it would not be permitted by a rule they had learnt.

DOI: 10.1177/1747021820905739

Indeed, we are not the first to identify misinterpretations with BODMAS; during interviews with prospective primary-school teachers, Zakis & Rouleau (2017) found that 55% believed that division should be performed before multiplication. Interestingly, in the United States, where the literal order of multiplication and division are reversed in the acronym PEMDAS, 38% of prospective teachers believed the opposite, that multiplication should precede division (Glidden, 2008). To avoid misinterpretations, teachers have been found to encourage a left-to-right approach on problems that contain operations at equal levels (Kirshner, 1989) and this situation is particularly problematic for addition-subtraction associativity. A literal interpretation of BODMAS, and the remedy to overcome that literal interpretation (‘go left to right’), both suggest associativity shortcuts are not allowed. Thus, for some individuals, conceptual knowledge and procedural knowledge may conflict, and attention mechanisms would be unlikely to correct this. Individuals may understand associativity, but procedures or ‘rules’ which they have learnt interfere when applying it.

Task demands

There were two features in our task that may have hindered shortcut use, 1) the demands of switching between two strategies, and 2) the validity of the primes. As in previous literature (e.g. Dubé & Robinson, 2010; Robinson & Dubé, 2013; Robinson & Ninowski, 2003) we intermixed the order in which different problems (i.e. conducive and non-conducive problems) were presented. This may have increased switching demands because efficient performance required alternating between left-to-right and right-to-left strategies, depending on the problem. Individuals have often been found reluctant to change from using one strategy to another within tasks (Lovett & Anderson, 1996; Siegler & Lemaire, 1997; Schillemans, Luwel, Onghena, & Verschaffel, 2011; Schillemans, Luwel, Ceulemans, Onghena, & Verschaffel, 2012; Verschaffel, Luwel, Torbeyns, & Van Dooren, 2009). In our studies, individuals may have been reluctant to switch from a strategy which “worked” (left-to-right) on both problem types, to one which also “worked” (right-to-left) but was more efficient for half of the problems.

DOI: 10.1177/1747021820905739

Indeed, this could explain why our findings differ from the aforementioned perceptual studies that used spacing manipulations (Jiang et al., 2014; Landy & Goldstone, 2007; Landy et al., 2008). In their studies, the stimuli contained operations that were not associative, e.g. '25 − 10 + 2 × 3', which afford only one correct strategy: multiplication, then subtraction, then addition. Efficiency and accuracy influence individuals' evaluations of different problem solving strategies (Brown, Menendez, & Alibali, 2019). Furthermore, it has been found that individuals are more likely to switch strategies if the difference between them is one of accuracy, rather than efficiency (Siegler, 2007). In other words, individuals are more willing to switch from a strategy which produces an incorrect result to one which produces a correct result; they are less willing to switch from a correct strategy to another correct strategy just to save time (Siegler, 2007). Individuals in the previous studies may have therefore been more willing to abandon a left-to-right approach than the participants in our studies because, in contrast to associativity stimuli, order *mattered.*

Finally, it may have been that the primes were perceived as invalid cues. Cue validity is "the conditional probability that an object is in a particular category, given its possession of some feature or cue" (Murphy, 1982). In psychology, this translates into the observation that the influence of a cue (e.g. an alarm) on behaviour towards a target (e.g. a fire) is proportional to the number of times that the stimuli co-occur, divided by the number of times that they do not. In other words, if the cue reliably predicts the onset of the target, the cue is informative, relevant and the individual adjusts their behaviour accordingly (e.g. to escape). If not, the cue is less relevant and may be ignored. This cue validity effect is widely documented and robust (Posner & Petersen, 1990; Eckstein, Abbey, Pham, & Shimozaki, 2004; Eckstein, Drescher, & Shimozaki, 2006; Jollie, Ivanoff, Webb, & Jamieson, 2016). In the current studies, the primes were valid on only 50% of the trials. For example, in the right-prime conditions, the prime cued the more efficient strategy on the conducive trials only. On non-conducive trials, the prime was not helpful, or even worse, *unhelpful*, to the extent that a left-

DOI: 10.1177/1747021820905739

to-right procedure was more efficient. Thus, it could be expected that some individuals ignored the primes because they were not beneficial for every trial.

Conclusion

We developed a new tool that can be used by other researchers investigating the development of conceptual understanding and strategy use in a range of domains. For the first time, we measured shortcut use on a trial-by-trial basis using implicit (response time) data without relying on explicit self-reports after each arithmetic problem was solved. This allowed us to capture the time-point of arithmetic strategy change. In two studies we used this method to investigate whether manipulating attention could encourage the identification of an associativity shortcut on 'a + b − c' problems and in both studies, we found no evidence that attention encouraged or hindered identification. We offer five explanations of our results and suggest that the most likely reasons for why individuals did not use the shortcut may be a) the demands of the task or b) inadequate or conflicting knowledge of the associativity principle. Given the importance of associativity in students' later algebra learning, further research into why the principle is poorly understood is warranted.

DOI: 10.1177/1747021820905739

Competing Interests

The authors have no competing interests.

Ethical Statement

Both studies were approved by the (name removed for blind review) University's ethics (Human participants) sub-committee. All individuals provided written informed consent to take part.

Author's contributions

All authors were involved in the conception and design of both experiments, and interpretation of the data. The first author collected and analysed the data, and drafted and wrote the article. All authors were involved in revising and finalising the article.

Open Practices

The data from the present experiment are publicly available at the Open Science Framework website:

The pre-registration for our analyses for Experiment 1 can be found here (http://aspredicted.org/blind.php?x=jm8vs9)

The pre-registration for our analyses for Experiment 2 can be found here (http://aspredicted.org/blind.php?x=8q6zz7

The data for Experiment 1 and 2 can be found here (https://doi.org/10.17028/rd.lboro.7533755)

The macros used to analyse the output data for each participant can be found here (https://doi.org/10.17028/rd.lboro.7538654)

The PsychoPy scripts and files for experiment 1 can be found here (https://doi.org/10.17028/rd.lboro.7533770).

The PsychoPy scripts and files for experiment 2 can be found here (https://doi.org/10.17028/rd.lboro.7533794)

DOI: 10.1177/1747021820905739

The stimuli and interview questions used in the experiments can be found here (https://doi.org/10.17028/rd.lboro.7533764)Supplementary analyses can be found here (https://doi.org/10.17028/rd.lboro.8343308)

DOI: 10.1177/1747021820905739

References

Alibali, M. (2015). Perceptual processes in strategy discovery. In *How shall I solve this problem? Modeling and stimulating strategy use in mathematics.* Philadelphia, PA: Symposium conducted at the Biennial Meeting of the Society for Research in Child Development.

Alibali, M. W., Crooks, N. M., & McNeil, N. M. (2017). Perceptual suppot promotes strategy generation: Evidence from equation solving. *Developmental Psychology*, *36*, 153–168.

Babai, R., Shalev, E., & Stavy, R. (2015). A warning intervention improves students' ability to overcome intuitive interference. *ZDM - Mathematics Education*, *47*, 735–745. https://doi.org/10.1007/s11858-015-0670-y

Baroody, A. J. (2003). The development of arithmetic concepts and skills: Constructing adaptive expertise. In A. J. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills: Constructive Adaptive Expertise* (pp. 1–490). Mahwah, N.J.: Lawrence Erlbaum Associates, Inc. https://doi.org/10.4324/9781410607218

Bisanz, J., & LeFevre, J. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. Bjorklund (Ed.), *Children's Strategies: Contemporary Views of Cognitive Development* (pp. 213–244). New York: Psychology Press.

Brown, S. A., Menendez, D., & Alibali, M. W. (2019). Strategy adoption depends on characteristics of the instruction, learner, and strategy. *Cognitive Research: Principles and Implications*, *4*, 1–18. https://doi.org/10.1186/s41235-019-0158-3

Bryant, P., Christie, C., & Rendu, A. (1999). Children's understanding of the relation between addition and subtraction: inversion, identity, and decomposition. *Journal of Experimental Child Psychology*, *74*, 194–212. https://doi.org/10.1006/jecp.1999.2517

Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of*

DOI: 10.1177/1747021820905739

Experimental Child Psychology, *92*, 220–246. https://doi.org/10.1016/j.jecp.2005.06.001

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology*, *34*, 882–891. https://doi.org/10.1037/0012-1649.34.5.882

Carr, M., & Jessup, D. L. (1995). Cognitive and metacognitive predictors of mathematics strategy use. *Learning and Individual Differences*, *7*, 235–247. https://doi.org/10.1016/1041-6080(95)90012-8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Lawrence Erlbaum Associates Publishers. https://doi.org/10.1234/12345678

Crooks, N. M., & Alibali, M. W. (2014). Defining and measuring conceptual knowledge in mathematics. *Developmental Review*, *34*, 344–377. https://doi.org/10.1016/j.dr.2014.10.001

Dewolf, T., Van Dooren, W., Ev Cimen, E., & Verschaffel, L. (2014). The impact of illustrations and warnings on solving mathematical word problems realistically. *The Journal of Experimental Education*, *82*, 103–120. https://doi.org/10.1080/00220973.2012.745468

Dubé, A. K. (2014). Adolescents' understanding of inversion and associativity. *Learning and Individual Differences*, *36*, 49–59. https://doi.org/10.1016/j.lindif.2014.09.002

Dubé, A. K., & Robinson, K. M. (2010). The relationship between adults' conceptual understanding of inversion and associativity. *Canadian Journal of Experimental Psychology*, *64*, 60–66. https://doi.org/10.1037/a0017756

Eaves, J. M., Attridge, N., & Gilmore, C. (2019). Increasing the use of conceptually-derived strategies in arithmetic: from inversion to associativity. *Learning and Instruction*, 1–25. https://doi.org/10.1016/j.learninstruc.2019.01.004

Eckstein, M. P., Abbey, C. K., Pham, B. T., & Shimozaki, S. S. (2004). Perceptual learning through

DOI: 10.1177/1747021820905739

optimization of attentional weighting: Human versus optimal Bayesian learner. *Journal of Vision*, *4*, 3. https://doi.org/10.1167/4.12.3

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science*, *17*, 973–980. https://doi.org/10.1111/j.1467-9280.2006.01815.x

Edwards, W. (2013). *Underlying components and conceptual knowledge in arithmetic*. University of Regina, Saskatchewan. Retrieved from https://ourspace.uregina.ca/bitstream/handle/10294/5434/Edwards_William_200275843_MA_EAP_Spring2014.pdf?sequence=1

Faulkenberry, T. J. (2013). The conceptual/procedural distinction belongs to strategies, not tasks: A comment on Gabriel et al. (2013). *Frontiers in Psychology*, *4*, 1–2. https://doi.org/10.3389/fpsyg.2013.00820

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, *34*, 906–911. https://doi.org/10.1037/0003-066x.34.10.906

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., … Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology*, *46*, 1731–1746. https://doi.org/10.1037/a0020662

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology*, *49*, 363–383. https://doi.org/10.1016/0022-0965(90)90065-G

Ghatala, E. S., Levin, J. R., Pressley, M., & Goodwin, D. (1986). A componential analysis of the effects of derived and supplied strategy-utility information on children's strategy selections. *Journal of*

DOI: 10.1177/1747021820905739

Experimental Child Psychology, *41*, 76–92. https://doi.org/10.1016/0022-0965(86)90052-4

Gilmore, C. K., & Bryant, P. (2006). Individual differences in children's understanding of inversion and arithmetical skill. *The British Journal of Educational Psychology*, *76*, 309–331. https://doi.org/10.1348/000709905X39125

Gilmore, C., & Papadatou-Pastou, M. (2009). Patterns of individual differences in conceptual understanding and arithmetical skill: A meta-analysis. *Mathematical Thinking and Learning*, *11*, 25–40. https://doi.org/10.1080/10986060802583923

Glidden, P. L. (2008). Prospective elementary teachers' understanding of order of operations. *School Science and Mathematics*, *108*, 130–136. Retrieved from https://doi.org/10.1111/j.1949-8594.2008.tb17819.x

Haider, H., Gaschler, R., Vaterrodt, B., & Frensch, P. A. (2014). How we use what we learn in Math: An integrative account of the development of commutativity. *Frontline Learning Research*, *2*, 1–21. Retrieved from https://files.eric.ed.gov/fulltext/EJ1090841.pdf

Haider, H., & Rose, M. (2007). How to investigate insight: A proposal. *Methods*, *42*, 49–57. https://doi.org/10.1016/j.ymeth.2006.12.004

Harms, C., & Lakens, D. (2018). Making "null effects" informative: statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research*, *3*, 382–393. https://doi.org/10.18053/jctres.03.2017s2.007

Hiebert, J. C., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: an introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of Mathematics* (pp. 1–27).

Jeffreys, H. (1961). *Theory of probability*. *Oxford Classic Texts in the Physical Sciences* (3rd ed.). Oxford: Oxford University Press, Clarendon press.

DOI: 10.1177/1747021820905739

Jiang, M. J., Cooper, J. L., & Alibali, M. W. (2014). Spatial factors influence arithmetic performance: The case of the minus sign. *Quarterly Journal of Experimental Psychology*, *67*, 1626–1642. https://doi.org/10.1080/17470218.2014.898669

Johnston, W. A., & Dark, V. J. (1986). Selective Attention. *Annual Review of Psychology*, *37*, 43–75. https://doi.org/10.1146/annurev.ps.37.020186.000355

Jollie, A., Ivanoff, J., Webb, N. E., & Jamieson, A. S. (2016). Expect the unexpected: a paradoxical effect of cue validity on the orienting of attention. *Attention, Perception, & Psychophysics*, *78*, 2124–2134. https://doi.org/10.3758/s13414-016-1164-x

Kilpatrick, J., Swafford, J., & Findell, B. (2002). *Adding it Up: Helping children learn mathematics*. Washington, D.C: National Academic Press. https://doi.org/10.17226/9822

Kim, M. S., & Cave, K. R. (1995). Spatial attention in visual search for features and feature conjunctions. *Psychological Science*, *6*, 376–380. https://doi.org/10.1111/j.1467-9280.1995.tb00529.x

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: the concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology*, *54*, 105–116. https://doi.org/10.1037/h0087333

Ladson-Billings, G. (1997). It doesn't add up: African American students' mathematics achievement. *Journal for Research in Mathematics Education*, *28*, 697–708. https://doi.org/10.2307/749638

Lakens, D. (2017). Equivalence tests: A practical primer for T-tests, correlations, and meta-analyses. *Social Psychological and Personality Science*, *8*, 355–362. https://doi.org/10.1177/1948550617697177

Landy, D., & Goldstone, R. L. (2007a). How abstract is symbolic thought? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(4), 720–733. https://doi.org/10.1037/0278-

DOI: 10.1177/1747021820905739

7393.33.4.720

Landy, D., & Goldstone, R. L. (2007b). The alignment of ordering and space in arithmetic computation. *Proceedings of the Twenty-Ninth Annual Meeting of the Cognitive Science Society*.

Landy, D., & Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quarterly Journal of Experimental Psychology*, *63*, 1953–1968. https://doi.org/10.1080/17470211003787619

Landy, D. H., Jones, M. N., & Goldstone, R. L. (2008). How the appearance of an operator affects its formal precedence. *Proceedings of the Thirtieth Annual Conference of the Cognitive Science Society*, 2109–2114.

Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition*, *38*, 322–32. https://doi.org/10.3758/MC.38.3.322

Lemaire, P., & Lecacheur, M. (2011). Age-related changes in children's executive functions and strategy selection: A study in computational estimation. *Cognitive Development*, *26*, 282–294. https://doi.org/10.1016/j.cogdev.2011.01.002

Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: combined influences on operator selection. *Cognitive Psychology*, *31*, 168–217. https://doi.org/10.1006/cogp.1996.0016

Mevarech, Z. R., & Fridkin, S. (2006). The effects of IMPROVE on mathematical knowledge, mathematical reasoning and meta-cognition. *Metacognition and Learning*, *1*, 85–97. https://doi.org/10.1007/s11409-006-6584-x

Mevarech, Z. R., & Kramarski, B. (1997). Improve: A multidimensional method for teaching mathematics in heterogeneous classrooms. *American Educational Research Journal*, *34*, 365–394. https://doi.org/10.3102/00028312034002365

Murphy, G. L. (1982). Cue validity and levels of categorization. *Psychological Bulletin*, *91*, 174–177.

DOI: 10.1177/1747021820905739

https://doi.org/10.1037/0033-2909.91.1.174

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel* (Vol. 37). Washington, D.C. https://doi.org/10.3102/0013189X08329195

Newton, K. J., Star, J. R., & Lynch, K. (2010). Understanding the development of flexibility in struggling algebra students. *Mathematical Thinking and Learning*, *12*, 282–305. https://doi.org/10.1080/10986065.2010.482150

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press. https://doi.org/10.1057/jors.1994.133

Peirce, J. W., Gray, J., Halchenko, Y., Britton, D., Rokem, A., & Strangman, G. (2019). *PsychoPy – A Psychology Software in Python*. Retrieved from http://www.psychopy.org/PsychoPyManual.pdf

Petersen, S. E., & Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annual Review of Neuroscience*, *35*, 73–89. https://doi.org/10.1146/annurev-neuro-062111-150525

Posner, G. J., & Gertzog, W. A. (1982). The clinical interview and the measurement of conceptual change. *Science Education*, *66*(2), 195–209. https://doi.org/10.1002/sce.3730660206

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience*, *13*, 25–42. https://doi.org/10.1146/annurev.ne.13.030190.000325

Posner, M. I., Snyder, C. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, *109*, 160–174. https://doi.org/10.1037/0096-3445.109.2.160

Prather, R. W., & Alibali, M. W. (2009). The development of arithmetic principle knowledge: How do we know what learners know? *Developmental Review*, *29*, 221–248.

DOI: 10.1177/1747021820905739

https://doi.org/10.1016/j.dr.2009.09.001

Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology (General)*, *135*, 207–236. https://doi.org/10.1093/acprof:oso/9780199744282.003.0011

Rittle-Johnson, B., & Schneider, M. (2015). Developing conceptual and procedural knowledge in mathematics. In *Oxford handbook of numerical cognition* (pp. 1102–1108).

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In *The Development of Mathematical Skills* (pp. 75–110).

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*, 346–362. https://doi.org/10.1037/0022-0663.93.2.346

Robertson, I. H., Ward, T., Ridgeway, V., & Nimmo-Smith, I. (1996). The structure of normal human attention: The Test of Everyday Attention. *Journal of the International Neuropsychological Society : JINS*, *2*, 525–534. https://doi.org/10.1017/S1355617700001697

Robinson, K., & Dubé, A. (2009). A microgenetic study of the multiplication and division inversion concept. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Experimentale*, *63*, 193–200. https://doi.org/10.1037/a0013908

Robinson, K. M., & Dubé, A. (2017). Children's understanding of additive concepts. *Journal of Experimental Psychology*, *156*, 16–28. https://doi.org/10.1016/j.jecp.2016.11.009

Robinson, K. M., & Dubé, A. K. (2012). Children's use of arithmetic shortcuts: The role of attitudes in strategy choice. *Child Development Research*, *2012*, 1–10. https://doi.org/10.1155/2012/459385

DOI: 10.1177/1747021820905739

Robinson, K. M., & Dubé, A. K. (2013). Children's additive concepts: Promoting understanding and the role of inhibition. *Learning and Individual Differences*, *23*, 101–107. https://doi.org/10.1016/j.lindif.2012.07.016

Robinson, K. M., & LeFevre, J. A. (2012). The inverse relation between multiplication and division: Concepts, procedures, and a cognitive framework. *Educational Studies in Mathematics*, *79*, 409–428. https://doi.org/10.1007/s10649-011-9330-5

Robinson, K. M., & Ninowski, J. E. (2003). Adults' understanding of inversion concepts: how does performance on addition and subtraction inversion problems compare to performance on multiplication and division inversion problems? *Canadian Journal of Experimental Psychology*, *57*, 321–330.

Schillemans, V., Luwel, K., Ceulemans, E., Onghena, P., & Verschaffel, L. (2012). The effect of single versus repeated previous strategy use on individuals' subsequent strategy choice. *Psychologica Belgica*, *52*, 307–325. https://doi.org/10.5334/pb-52-4-307

Schillemans, V., Luwel, K., Onghena, P., & Verschaffel, L. (2011). The influence of the previous strategy on individuals' strategy choices. *Studia Psychologica*, *53*(4), 339–350. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-84856547469&partnerID=40&md5=04b811bd84901d2d75bb1761847f60a0

Schneider, M., & Stern, E. (2010). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology*, *46*, 178–192. https://doi.org/10.1037/a0016701

Schoenfeld, A. (1985). *Mathematical problem solving: issues in research*. New York: Academic Press.

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*, 405–410. https://doi.org/10.1111/1467-9280.00076

DOI: 10.1177/1747021820905739

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264. https://doi.org/10.1037/0096-3445.116.3.250

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conservation. *Cognitive Psychology*, *28*, 225–273. https://doi.org/10.1006/cogp.1995.1006

Siegler, R. S. (2006). Microgenetic Analyses of Learning. In *Handbook of Child Psychology* (Vol. 2, pp. 464–510). https://doi.org/10.1002/9780470147658.chpsy0211

Siegler, R. S. (2007). Microgenetic Analyses of Learning. In *Handbook of Child Psychology* (Vol. 2, pp. 464–510). https://doi.org/10.1002/9780470147658.chpsy0211

Siegler, R. S., & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. *Advances in Child Development and Behavior*, *33*, 1–42. https://doi.org/10.1016/S0065-2407(05)80003-5

Siegler, R. S., & Lemaire, P. (1997). Older and younger adults' strategy choices in multiplication: Testing predictions of ASCM using the choice/no-choice method. *Journal of Experimental Psychology: General*, *126*, 71–92. https://doi.org/10.1037/0096-3445.126.1.71

Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries:A microgenetic analysis. *Journal of Experimental Psychology: General*, *127*, 377–397. https://doi.org/10.1037/0096-3445.127.4.377

Siegler, R. S., & Svetina, M. (2006). What leads children to adopt new strategies? A microgenetic/cross-sectional study of class inclusion. *Child Development*, *77*, 997–1015. https://doi.org/10.1111/j.1467-8624.2006.00915.x

Starkey, P., & Gelman, R. (1989). The development of addition and subtraction abilities prior to formal schooling in arithmetic. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition*

DOI: 10.1177/1747021820905739

and Subtraction: A cognitive perspective (pp. 99–116). Hillsdale, NJ: Laurence Erlbaum Associates.

Thevenot, C., & Oakhill, J. (2005). The strategic use of alternative representations in arithmetic word problem solving. *The Quarterly Journal of Experimental Psychology. A, Human Experimental Psychology*, *58*, 1311–1323. https://doi.org/10.1080/02724980443000593

Thevenot, C., & Oakhill, J. (2006). Representations and strategies for solving dynamic and static arithmetic word problems: The role of working memory capacities. *European Journal of Cognitive Psychology*, *18*, 756–775. https://doi.org/10.1080/09541440500412270

Verschaffel, L., Luwel, K., Torbeyns, J., & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, *24*, 335–359. https://doi.org/10.1007/BF03174765

Watchorn, R. P. D., Bisanz, J., Fast, L., LeFevre, J.-A., Skwarchuk, S.-L., & Smith-Chant, B. L. (2014). Development of mathematical knowledge in young children: attentional skill and the use of inversion. *Journal of Cognition and Development*, *15*, 161–180. https://doi.org/10.1080/15248372.2012.742899

Wertheimer, M. (1923). Laws of Organization in Perceptual Forms. In W. . Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). London, Routledge.

Zakis, R., & Rouleau, A. (2017). Order of operations: On convention and met-before acronyms. *Educational Studies in Mathematics*, *97*, 143–162. https://doi.org/10.1007/s10649-017-9789-9

Zentall, T. R. (2005). Selective and divided attention in animals. *Behavioural Processes*, *69*, 1–15. https://doi.org/10.1016/j.beproc.2005.01.004

DOI: 10.1177/1747021820905739

Figure captions

Figure 1

The procedure of a trial in the left-prime, right-prime and control conditions in Experiment 1

Figure 2

Trial-by-trial RT data for one participant. The identification-point refers the time-point at which the individual changed from using a less efficient (e.g. left-to-right) strategy, to a more efficient associativity-shortcut strategy.

DOI: 10.1177/1747021820905739

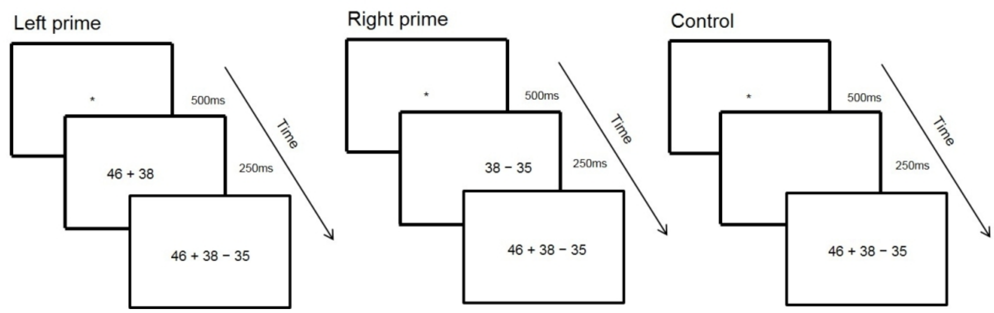

Figure 1: The procedure of a trial in the left-prime, right-prime and control conditions in Experiment 1

249x77mm (150 x 150 DPI)

DOI: 10.1177/1747021820905739

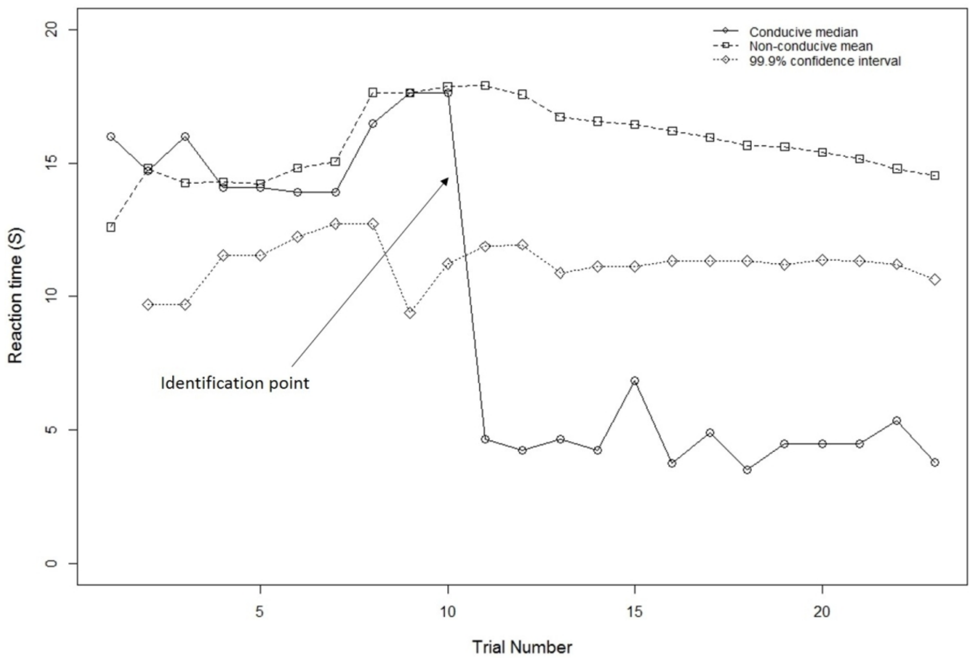

Figure 2: Trial-by-trial RT data for one participant. The identification-point refers the time-point at which the individual changed from using a less efficient (e.g. left-to-right) strategy, to a more efficient associativity-shortcut strategy.

257x190mm (150 x 150 DPI)

DOI: 10.1177/1747021820905739

Table 1

Number of associativity shortcut identifiers and non-identifiers in each condition of Experiment 1.

	Condition			
	Left-prime	Control	Right-prime	Total
Identifier	19	17	25	61
Non-identifier	17	19	11	47
Total	36	36	36	108

DOI: 10.1177/1747021820905739

Table 2

Mean trial number (SD) of identification for the associativity shortcut identifiers in Experiment 1.

	Condition		
	Left-prime	Control	Right-prime
Conducive trial number	8.00 (3.73)	6.71 (3.60)	6.44 (5.36)
Overall trial number	15.58 (7.41)	13.12 (7.11)	12.52 (10.49)

DOI: 10.1177/1747021820905739

Table 3

Mean accuracy and median correct RT across all trials for participants who identified the associativity shortcut (Experiment 1)

	Conducive problems		Non-conducive problems	
Condition	Accuracy (%)	RT (s)	Accuracy (%)	RT (s)
Left-prime	91.99 (8.85)	4.92 (2.77)	78.95 (24.78)	11.03 (4.33)
Control	90.28 (15.26)	6.14 (2.28)	79.54 (20.26)	13.65 (5.40)
Right-prime	94.43 (6.09)	6.61 (6.01)	82.09 (12.70)	12.42 (5.95)

DOI: 10.1177/1747021820905739

Table 4

Number of identifiers and non-identifiers of the associativity shortcut in each condition of Experiment 2.

	Condition			
	Left-prime	Control	Right-prime	Total
Identifier	18	17	20	55
Non-identifier	18	19	16	53
Total	36	36	36	108

DOI: 10.1177/1747021820905739