Frontline Learning Research
An official journal of EARLI

Frontline **L**earning **R**esearch 1 (2014) 1-21
ISSN 2295-3159

How we use what we learn in Math: An integrative account of the development of commutativity

Hilde Haider[a], Alexandra Eichler [a], Sonja Hansen [a], Bianca Vaterrodt [b], Robert Gaschler [c], Peter A. Frensch [b]

[a]University of Cologne, Germany

[b]Humboldt-University Berlin, Germany

[c]University Koblenz-Landau, Germany

Article received 28 May 2013 / revised 6 January 2014/ accepted 16 January 2014 / available online 27 January 2014

Abstract

One crucial issue in mathematics development is how children come to spontaneously apply arithmetical principles (e.g. commutativity). According to expertise research, well-integrated conceptual and procedural knowledge is required. Here, we report a method composed of two independent tasks that assessed in an unobtrusive manner the spontaneous use of procedural and conceptual knowledge about commutativity. This allowed us to ask (1) in which grade students spontaneously apply this principle in different task formats and (2) in which grade they start to possess an integrated concept of the commutativity. Procedural and conceptual knowledge of 8 to 9 year olds (163 second and 180 third graders) as well as 46 adult students was assessed independently and without any hint concerning commutativity. Results indicated procedural as well as conceptual knowledge about commutativity for second graders. However, their procedural and conceptual knowledge was unrelated. An integrated relation between the two measures first emerged with some of the third graders and was further strengthened for adult students.

Keywords: Conceptual Knowledge; Procedural Knowledge; Commutativity; Integrated Concept.

Corresponding author: Hilde Haider, Department of Psychology, University of Cologne, Richard-Strauss-Str. 2
50931 Cologne, Germany, Phone: +49-221-4704719, Email: hilde.haider@uni-koeln.de

http://dx.doi.org/10.14786/flr.v2i1.37

1. Introduction

One major skill in mathematics is the acquisition of adaptive expertise. That is, students should be able to deliberately recognize those constraints that allow to apply a certain mathematical principle (e.g., Torbeyns, De Smedt, Ghesquière & Verschaffel, 2009; Verschaffel, Luwel, Torbeyns & Van Dooren, 2009). For instance, in the PISA mathematical literacy test students have to spontaneously apply mathematical principles in order to solve problems in real-world contexts. Given the important role of self-guided learning and performance in the development of mathematical abilities and concepts, some recent studies have started to focus on spontaneous recognition of mathematical aspects in natural surroundings (e.g., Hannula, & Lehtinen, 2005; Hannula, Lepola, & Lehtinen, 2010; McMullen, Hannula-Sormunen, & Lehtinen, 2011).

An important question with regard to adaptive expertise is how students come to recognize that they can use a certain principle in order to facilitate calculation. Or to put it in other words, what kind of knowledge underlies the ability to adaptively apply a mathematical principle spontaneously *whenever* it facilitates calculation? In the research on adaptive expertise, it is widely accepted that this ability is not only based on procedural knowledge (knowing how to apply a certain strategy), but also on conceptual knowledge (knowing when and why a certain principle applies). Procedural and conceptual knowledge should be integrated and the resulting knowledge base should be abstract enough to ensure flexibility in knowledge application (e.g., Anderson & Schunn, 2000; Baroody, 2003; Gentner & Toupin, 1986; Haider & Frensch, 1996; Koedinger & Anderson, 1990; Star, 2005; Verschaffel et al., 2009). As one example of linking concepts and procedures, this research has revealed that conceptual knowledge is important to guide attention to task relevant information in order to solve problems (e.g., Baroody & Rosu, 2006). An abstract conceptual understanding might also be of particular importance when knowledge has to be transferred from one domain to another (e.g., Goldstone & Sakamoto, 2003; Kaminsky, Sloutsky & Heckler, 2008). It supports flexible shortcut application when problems to which a principle applies are presented mixed with problems to which the principle does not apply. For instance, Siegler and Stern (1998) have shown that second graders relied less on inversion-based procedures when inversion problems (a + b – b) were randomly interspersed with control problems (a + b – c) compared to blocked presentation. Mixed presentation of inversion and control problems hindered the use of inversion short-cuts. This suggests that younger children do not deliberately recognize the constraints important for applying the inversion principle. Rather, they simply seem to know that the strategy applies for a certain class of problems. Likely they cannot rely on a well integrated, abstract understanding of the inversion principle (i.e., they have not yet developed adaptive expertise). But, how and when do children develop an integrated representation of basic mathematical principles?

The first goal of the current study was to develop a method to unobtrusively measure the *spontaneous* application of procedural and conceptual knowledge taking the commutativity principle as a test case. The second goal was to shed some light on the development of an abstract and well integrated representation of the commutativity principle. With regard to the second goal, we pursued to different questions: (1) In which grade are students able to spontaneously apply commutativity knowledge in different task formats and (2) in which grade starts performance expressed in the different tasks to correlate with one-another? For this purpose, we investigated the deliberate use of the commutativity principle in two different situations. The term "deliberate" means that children did not receive any hint about the commutativity principle at all (e.g., Torbeyns et al., 2009). In the first test children simply solved addition problems that sometimes allowed for a shortcut based on the commutativity principle (procedural knowledge). The second test was aimed at assessing conceptual knowledge. Children were instructed to mark – without solving the problem – those problems that they believed could be solved without calculation. Hence, this task required children to realize that the order of addends does not change cardinality. The correlation between these two independent tasks allowed us to gauge how well integrated children's knowledge was. Focusing on just two unobtrusive measures (one procedural and one conceptual), the current work can potentially lay the ground to develop multi-method approaches in the same spirit, safeguarding that multiple testing does not cue participants towards what the test situation is about.

We focused on the commutativity principle as it is one of the most basic properties in mathematics. It refers to the principle that changing the order of operands in addition and multiplication does not change the end result. It is known as a fundamental property of many binary operations. The commutativity of simple operations, such as the multiplication and addition of numbers are usually acquired throughout elementary school. However, many mathematical proofs also depend on this property.

1.1 Development of procedural and conceptual knowledge about commutativity

Former research in the field of developmental psychology has already shown that children acquire informal knowledge of commutativity as an arithmetic principle long before they enter school (e.g., Baroody & Gannon, 1984; Baroody, Ginsburg & Waxman, 1983; Canobi, Reeve & Pattison, 1998, 2002; Cowan & Renton, 1996; Resnick, 1992; Siegler & Jenkins, 1989; Sophian, Harley & Manos Martin, 1995). One potential reason for this early development is that at least the core property of commutativity, the order-irrelevance principle, applies to many non-numerical situations. For example, children may experience that some tasks require a certain sequence (e.g., putting on one's clothes), whereas others do not (e.g., laying the table). Already toddlers have many opportunities to learn that order does not affect the end result in some situations, but does in others.

Order-irrelevance is also a core principle for counting (e.g., Gelman, 1990; Gelman & Gallistel, 1978). Learning to count requires children to learn, on the one hand, that the sequence of number words is relevant. On the other hand, the sequence in which the objects are counted is irrelevant. Consequently, Briars and Siegler (1984) found that children need time to understand order-irrelevance in counting. Furthermore, counting is the dominant skill through which preschool children learn to map concrete objects to numbers. Also, counting is one of the important precursors of addition. Through counting, pre-school children can learn order-irrelevance in a numerical manner before entering school. They thus do not only have the chance to understand order-irrelevance in a non-numerical manner.

However, even though considerable interest in research on counting and addition principles emerged already in the 1980s and still continues (e.g., Baroody, 1984; Baroody & Gannon, 1984; Baroody et al., 1983; Briars & Siegler, 1984; Canobi, Reeve & Pattison, 1998, 2002, 2003; Fuson, 1988; Gelman & Gallistel, 1978; Gelman & Meck, 1983; Resnick, 1992; Sophian & Adams, 1987; Starkey & Gelman, 1982), the central question has not been solved yet: How and when do children acquire integrated knowledge representations, in the sense of true formal arithmetic principles? For instance, Geary (2006) stated that it is not clear when children "explicitly understand commutativity as a formal arithmetical principle" (p. 791).

On the one hand, the difficulties in answering this question are due to the fact that researchers by no means agree upon the characteristics of procedural or conceptual knowledge that must be given in order to conclude that children possess an abstract mathematical concept (cf., Star, 2004). Concerning procedural knowledge, most researchers agree that it refers to the ability to apply a certain strategy when performing a mathematical task (e.g., Hiebert & LeFevre, 1986). Conceptual knowledge or metastrategic competences (Kuhn, Garcia-Mila, Zohar & Andersen, 1995) often is assumed to refer to children's explicit understanding of a certain principle (i.e., why and when it is allowed to a apply a certain strategy; e.g., Baroody, Feil & Johnson, 2007; Hiebert & LeFevre, 1986; Rittle-Johnson, Siegler & Alibali, 2001).

On the other hand, there is no consensus how best to assess procedural and conceptual knowledge. One frequently used approach to measure procedural knowledge is to ask children to solve addition problems and afterwards have them explain their strategies (e.g., Baroody & Gannon, 1984; Baroody, Ginsburg & Waxman, 1983; Bisanz & LeFevre, 1992; Canobi et al., 1998, 2002, 2003; Cowan & Renton, 1996). Conceptual knowledge in these studies has, for example, been assessed by letting children observe a puppet solving problem pairs (see, e.g. Baroody et al., 1983; Canobi et al, 1998). If a child on enquiry stated that the puppet could know the answer to the second problem from looking at the previous one, he or she was asked for reasons and eventually prompted for more detailed explanations. This form of assessment implies that children are being informed about the underlying arithmetic principle – at least they are made aware that different efficient strategies are applicable. Such procedural and the conceptual knowledge tests might guide children's attention to the task-relevant information. Also, it is conceivable that they look at the problems more attentively when they are asked to verbalize their strategies. Consequently, conclusions concerning the question whether a child possesses abstract conceptual knowledge may vary depending on the tests that were applied.

Based on the above-mentioned forms of assessment, the empirical research on commutativity suggests that conceptual and procedural knowledge in this domain are moderately related (e.g., Baroody et al., 1983; Canobi, 2004; Canobi et al., 1998). However, the findings do not allow to exclude that the acquired conceptual knowledge of first or even second graders is still domain-specific rather than akin to an abstract concept representing the formal arithmetic principle of commutativity (e.g., Bisanz, Watchhorn, Piatt & Sherman, 2009; Geary, Hoard, Byrd-Craven & DeSoto, 2004; LeFevre et al., 2006). Therefore, investigating the spontaneous application of commutativity knowledge would complement and broaden this research.

In summary, the goal of our study was twofold: First, we aimed to develop a method to unobtrusively test for *spontaneous* application of procedural and conceptual commutativity knowledge. Our second goal was to investigate the degree of integration of this spontaneously expressed procedural and conceptual knowledge of second and third graders. Additionally, for means of comparison, we also tested adult students. As described above, knowledge about a mathematical principle like, for instance, the commutativity principle, can be said to represent an integrated or abstract concept in the sense of a true formal mathematic principle when learners are able to apply their knowledge *whenever* task constraints permit. That is, learners should be able to deliberately recognize task properties allowing them to apply the mathematical principle irrespectively of task context.

2. Method

2.1. General Method

We investigated the commutativity principle with three-element addition problems (i.e., 5+3+7 = ?)[1]. These three-element problems are unfamiliar at least for younger students. Since we wanted to investigate whether or not students would recognize the applicability of the commutativity principle without any further information about this principle, we needed less familiar problems. Therefore, we accepted that three-element problems implicitly presuppose knowledge about associativity (e.g., Canobi et al., 1998).

The three-element problems were always presented in blocks, one problem beneath the other. Unbeknownst to the participating students, some problems consisted of identical addends in a different order as the preceding problem, and thus could be solved without calculation (commutative problems, hereafter). Students received two different and completely independent task formats.

The first task, the arithmetic task, consisted of two blocks. One block contained interspersed commutative problems, the other one did not. Participants did not receive any information about the existence of these commutative problems. Rather, they were simply asked to solve the two blocks of addition problems as fast and accurately as possible. If students are faster when working on the block that includes three-element commutative problems as compared to the block which does not contain such shortcut options, they can be said to possess procedural knowledge about commutativity.

[1] Some researchers use the term *associativity* instead of *commutativity* when an addition or multiplication problem has more than two addends or factors (Geary et al., 2008). Other researchers (Canobi, et al., 1998) refer to c*ommutativity* as the property that problems containing the same terms in a different order have the same answer independent of the number of terms, whereas *associativity* is the property that problems in which terms are decomposed and recombined in different ways have the same answer $[(a + b) + c = a + (b + c)]$.

In the second task, the so-called judgment task, students were instructed to identify those problems which they believed need no calculation. They explicitly were told to refrain from calculating any problems. If students understand that the order of identical addends does not change the cardinality, they should be able to correctly mark the commutative problems. By virtue of this second task type, we were able to assess conceptual knowledge about commutativity without cueing the concept.

Importantly and in contrast to former experiments (e.g., Baroody et al., 1983; Canobi, 2005), participants in our experiment did not receive any hint about the existence of commutative problems in either task. That is, they were not instructed to further explain their strategies. The rationale behind this procedure was that any instruction to think about the strategies used to solve the problems might trigger active search for regularities, thereby making it impossible to assess the *spontaneously* activated concept of commutativity. If students possess an abstract understanding of the commutativity principle, they should be able to recognize and rely on the relevant task characteristics in any task context and without any hint (e.g., Bisanz et al., 2009; Prather & Alibali, 2009).

To the extent that children have acquired an abstract concept of commutativity, performance should correlate between both of our two tasks reflecting knowledge about this principle. Likely procedural and conceptual knowledge about commutativity becomes iteratively more integrated in the first years of primary school. We should thus find that the relation between procedural and conceptual knowledge is stronger in third graders as compared to second graders (e.g., LeFevre et al. 2006).

2.2 Participants

Overall, 163 second graders (79 girls) with a mean age of 8 years 1 month (SD = 7 months), 180 third graders (91 girls) with a mean age of 9 years 1 month (SD = 8 months) participated in the study. As a control condition, we also collected data of 46 students of the University of Cologne (37 women) with a mean age of 23.6 years (SD = 5.2). Children were recruited from six different elementary schools located in middle socioeconomic status suburbs of Cologne. All children had their parents' or guardians' permission to participate in the study.

2.3 Procedure and Materials

The study consisted of two parts. In the first part, participants received the arithmetic task: one block with interspersed repetitions of addends in changed order in consecutive problems (commutativity block) and one block without such repetitions (control block). In the second part, participants were administered the judgment task. Both tasks were designed as paper-pencil tests and children and adult students were tested in groups of up to 25 participants in a classroom-like setting.

We generated three sets of 30 arithmetic problems with three addends between 2 and 9 (e.g., 3 + 6 + 8 = ?; maximum result was 24; 1 as an addend was not included). The problems in all three sets yielded at least approximately the same totals and within a problem each numeral could only occur once.
The 30 problems of each of the two blocks were distributed over five pages with six problems on each page. In the commutativity block, each page contained two pairs of commutative problems (i.e., one problem and its repetition with a different order of addends). In the control block no such commutative pairs occurred. Instead participants received pairs of control problems which yielded the same results but were composed of different addends. In both blocks, participants were instructed to calculate the problems page by page from top to bottom.

The judgment task consisted of overall 30 problems with 10 problems per page. On each page, three pairs were commutative pairs and the remaining four problems were filler problems. The first page was for practice only. Participants were instructed to first solve all 10 problems from top to bottom on the page. Afterwards they were asked to mark those problems that needed no calculation on that page. In particular, they were told that some of the problems need no calculation and that they should figure out for which of these problems they could have written down the result without calculation. After this practice page, participants were instructed to *only judge* on the next pages whether or not they needed to calculate the result for a problem without actually attempting to solve it. Therefore, all problems on pages 2 and 3 were presented without equal sign. Instead, there was a circle to the right of each problem and participants were told to mark this circle when they believed they did not need to calculate the result. Again, students were instructed to work on the problems from the top to the bottom of each page. Table 1 depicts examples of the problems in each of the two arithmetic blocks and the judgment task.

Table 1

Examples of the problems presented on one page in the two arithmetic blocks (commutativity and control block) and the judgment task

Arithmetic task		Judgment task	
Commutativity Block	Control Block		
3 + 5 + 4 =	5 + 3 + 4 =	2 + 7 + 9	O
4 + 9 + 8 =	8 + 9 + 4 =	9 + 5 + 4	O
4 + 8 + 9 =	6 + 7 + 8 =	**2 + 6 + 5**	O
6 + 2 + 5 =	5 + 2 + 6 =	**6 + 5 + 2**	O
9 + 7 + 2 =	2 + 7 + 9 =	8 + 7 + 5	O
2 + 7 + 9 =	9 + 4 + 5 =	**3 + 5 + 6**	O
		6 + 5 + 3	O
		2 + 9 + 5	O
		6 + 7 + 9	O
		9 + 6 + 7	O

Problems in bold indicate the commutative pairs of the respective task

Each of the two arithmetic blocks was administered as a separate booklet, as was the judgment task. Students only worked with a pencil and were not allowed to use an eraser. Rather, to increase the reliability of the timing measure, they were told to cross out any errors and to write the correct answer right beside the problem. An experimenter instructed all participants in the classroom.

The experiment started with six arithmetic practice problems with three addends. The only goal of this phase was to familiarize the children with the task requirements. Students were given 2 minutes to solve these six warm-up problems. Then, the first of the two arithmetic blocks was presented. Approximately half of the children (second graders and third graders) and all adults in the control condition received the commutativity block first, followed by the control block. The remaining participants started with the control block and subsequently received the commutativity block. The time limit was set to 3 minutes per block (1 minute for adult students) with a 1-minute break between blocks.

One minute after having finished the second arithmetic block, the judgment task was presented. Participants were allowed 2 minutes (adult students again 1 minute) to calculate the problems on the practice page. Afterwards they had the same amount of time for marking those problems they believed they could have solved without calculation. After the practice phase, the same time limit was applied for the two subsequent pages, so that time did not suffice to calculate the problems and to concurrently mark those problems requiring no calculation. In addition, up to four additional experimenters observing small groups of children (up to six) ensured that they were not calculating the problems. After this last block, all children received some sweets. Adult students in the control condition were debriefed about the study.

2.4. Design

Independent variables were grade (second versus third graders) and block type (commutativity versus control block in the arithmetic task). Dependent variables in the arithmetic problem blocks were calculation time per problem in each of the two arithmetic blocks, as well as the number of correct results. Calculation time was computed separately for each participant and each of the two arithmetic problem blocks by dividing the individual number of completed problems by the total time given for the block in the respective age group (three minutes for second and third graders; 1 minute for adults). For the judgment task, the dependent variables were relative number of hits (correctly identified commutative problems) and false alarms (problems incorrectly identified as commutative problems), as well as the sensitivity index d’ from signal detection theory (i.e., the difference between z-transformed hit rate and false alarms rate).

2.5 Split-half Reliability

In order to check if our measures are reliable, we computed split-half reliabilities for each task type. That is, for each of the two age groups and the adults, we calculated correlations between the two arithmetic blocks (control and commutativity block) and between the second and third pages of the judgment task (the practice page of the judgment task was excluded). Table 2 shows the Spearman-Brown corrected correlation coefficients separately for each age group and each task format (arithmetic task and judgment task).

Table 2

Spearman-Brown corrected correlation coefficients for the arithmetic task and the judgment task for all participants and separately for the three age groups (Arithmetic task: correlation between the amount of computed problems in the commutativity block and the control block; Judgment task: correlation between correct responses on the first and on the second pages of the test)

Arithmetic task	Amount of computed commutative problems			
	All participants	Grade 2	Grade 3	Adults
Amount of computed control problems	.90	.88	.88	.95
Judgment task	Amount of correct judgments on the first page			
	All participants	Grade 2	Grade 3	Adults
Amount of correct judgments on the second page	.82	.83	.78	.86

As can be seen from Table 2, the correlation coefficients in each age group ranged between r = .78 and r = .95. Thus, the two tasks used to assess participants’ procedural and conceptual knowledge seem to be reliable measures.

3. Results

Second or third graders were excluded from further analyses if they completed less than 16 problems across the two arithmetic problem blocks (i.e., 2 standard deviations below the group means; 15 second graders, 12 third graders, and 1 adult). They were also excluded from further analyses if they solved all 30 problems in the control and the commutativity block, as calculation times could not be calculated for these participants (2 second graders, 23 third graders, and 8 adults). This led to 146 second graders, 145 third graders, and 37 adult students in the control condition. The following result section is divided into three parts. We first describe the results for the arithmetic problem blocks. Second, we report the performance in the judgment task. Lastly, we analyze the relation between these two tasks.

3.1 Arithmetic Task

As a preliminary analysis did not reveal substantial effects of the order of presentation (commutative problem first followed by control problem or vice versa), we collapsed the data for all participants within the groups of second and third graders. Table 3 depicts the calculation times for problems in the commutativity and the control blocks per age group. Mean calculation times suggest that second and third graders benefitted from the commutative problems whereas adult students did not.

Table 3

Calculation times per task in the commutativity and the control block for each age group. The table holds the means and standard deviations for the different age group in seconds as well as lower and upper limit of the 95-% confidence interval (CI; Loftus & Masson, 1994)

Age group	Commutativity block		Control block		N
	M (SD)	M±95%CI	M (SD)	M±95%CI	
Grade 2	12.33 (3.74)	12.09-12.66	13.28 (4.39)	12.95-13.60	146
Grade 3	9.55 (2.51)	9.34-9.76	9.93 (3.12)	9.72-10.14	145
Adults	3.79 (1.04)	3.66-3.92	3.63 (1.04)	3.50-3.76	37

A 2 (Age group) X 2 (Block type: commutativity block vs. control block) analysis of variance (ANOVA) with calculation time as dependent variable revealed significant main effects of Age group ($F[1, 289] = 66.23$, $MSe = 20.64$, $p < .01$, $\eta^2 = .23$), and of Block type ($F[1, 289] = 15.94$, $MSe = 4.04$, $p < .01$, $\eta^2 = .06$). The interaction between Age group and Block type was close to significance ($F[1, 289] = 2.98$, $MSe = 4.04$, $p = .088$). Planned contrasts revealed that only second graders significantly profited from the commutative problems (second graders: $F[1, 289] = 16.32$, $MSe = 4.04$, $p < .01$, $\eta^2 = .06$; third graders: $F[1,289] = 2.59$, $MSe = 4.04$, $p = .108$). A separate t-test with Block type as within-participants variable revealed that the adult control group did not show a significant benefit from commutative problems ($t < 1$).

In addition, we also analyzed the percentage of correct responses in the commutativity and the control blocks. Table 4 presents the percentage of correct responses in the two age groups for these two types of problems. As can be seen from Table 4, percentage of correct responses was higher for third- as compared to second graders. Accordingly, the 2 (Age group) X 2 (Block type) ANOVA yielded a significant main effect of Age group ($F[1, 289] = 3.8$, $MSe = 118.27$, $p < .05$, $\eta^2 = .01$). No other effect was significant.

Table 4

Mean percent correct responses in the three age groups for the commutativity and control blocks. Also depicted are standard deviations (in parentheses) and the lower and upper limit of the 95-% confidence interval (CI; Loftus & Masson, 1994)

Age group	Commutativity block		Control block		N
	M (SD)	M±95%CI	M (SD)	M±95%CI	
Grade 2	92.66 (9.85)	91.64-93.68	93.11 (9.13)	92.08-94.13	146
Grade 3	94.33 (8.62)	91.64-93.68	94.82 (8.31)	94.04-95.60	145
Adults	96.66 (9.08)	94.81-98.51	95.94 (12.55)	94.09-97.78	37

Overall, the results up to this point show that third graders were faster and less error prone as compared to second graders. Furthermore and more importantly, second graders showed a substantial benefit of commutative problems. Third graders in tendency also profited from these problems, but for them the effect was not significant. Adults, by comparison, did not show such a benefit, probably due to a floor effect based on the simplicity of the problems (for similar results, see Robinson & Dubé, 2009; Robinson & Ninowski, 2003).

3.2 Judgment Task

For each student, we individually computed the hit rate, false alarms rate, and the sensitivity index (d') from signal detection theory[2]. Table 5 depicts the means for these dependent measures separately for each of the two age groups and the adult students. As expected, hit rate was higher than false alarms rate in all age groups. Accordingly, the sensitivity index d' differed significantly from chance (all ts > 2.5, ps < .01). This suggests that students were able to correctly identify at least some of the commutative problems. In addition, d' was substantially higher in third- as compared to second graders ($t[289] = 2.18$, $p < .05$, $\eta^2 = .02$).

[2] Separately for each student within the respective age groups, we computed z-score of his or her hit and false alarms rate. Then, we individually computed the sensitivity index d' from signal detection theory by subtracting the z-transformed false alarms rate from the z-transformed hit rate.

Table 5

Rate of hits, false alarms and d’ in each of the three age groups in the judgment task. Standard deviants are given in parentheses.

Age group	Judgment task		
	Hits	False alarms	d’
Grade 2	.70 (.28)	.36 (.33)	1.81 (2.35)
Grade 3	.81 (.22)	.35 (.32)	2.41 (2.32)
Adults	.82 (.25)	.11 (.18)	3.92 (2.33)

Low sensitivity could result from two different sources: the difficulty to identify commutative problems (hit rate) or a tendency to mark other than commutative problems (false alarm rate). Therefore, we additionally analyzed the hit and false alarm rates in the two age groups. These analyses revealed that higher sensitivity in grade 3 as compared to grade 2 was mainly due to a higher hit rate. Second graders were less able to identify the commutative problems than third graders ($t[289] = 3.44$, $p < .01$, $\eta^2 = .04$). The false alarm rate did not differ significantly between these two age groups ($t < 1$). By contrast, as can be seen from Table 5, the higher sensitivity in adults as compared to third graders resulted from a lower false alarm rate in adults, whereas hit rate was almost identical in these two age groups.

Thus, older participants were better able to discriminate between commutative and control problems than younger participants. This finding from our cross-sectional age-comparison suggests a progress in conceptual knowledge with increasing age (as cohort differences are unlikely).

3.3 Relation between Procedural and Conceptual Knowledge

The results reported up to this point are somewhat counterintuitive. Even though third graders were better able to identify commutative problems in the judgment task, they seemed to rely less on a commutativity-based shortcut during calculation than second graders. In addition, for adults we found no benefit of commutative problems in the arithmetic task. Thus, it seems that either the willingness to use more efficient arithmetic strategies or procedural knowledge of commutativity itself decreases (while conceptual knowledge increases with age).

The last analyses of the relationship between procedural and conceptual knowledge might help to reconcile this picture. These analyses will answer the research question whether or not participants possess an integrated concept of commutativity. If so, we should find significant positive correlations between the use of the commutative-based shortcut in the arithmetic task (procedural knowledge) and the ability to correctly identify the commutative problems in the judgment task (conceptual knowledge).

For procedural knowledge, we used for each participant the average calculation time per problem in the control and the commutativity block of the arithmetic task as well as the difference between these two measures (i.e., savings; with positive values indicating shorter calculation times in the commutativity block). For conceptual knowledge, we used hit rate, false alarms rate, and the sensitivity measure d’. In a first analysis we calculated correlations across second and third graders. Second, we calculated correlations within the two age groups and for the adults. Table 6 depicts the correlation between procedural and conceptual knowledge.

Table 6

Correlation coefficients between procedural and conceptual knowledge depicted separately for all second and third graders as well as for the three age groups. Procedural knowledge is indicated by calculation times in seconds for commutative problems, control problems, and in addition for savings. Hit rate, false alarms, and d' indicate conceptual knowledge

		Arithmetic tasks			
	Judgment task	Commutative problems	Control problems	Savings	N
Second and third graders	Hits	-0.23**	-0.17**	0.001	
	False Alarms	-0.09	-0.09	0.02	291
	d'	-0.05	-0.02	0.01	
Grade 2	Hits	-0.23**	-0.18**	-0.04	
	False Alarms	-0.13	-0.15	-0.03	146
	d'	-0.09	-0.04	-0.008	
Grade 3	Hits	-0.02	0.06	0.11	
	False Alarms	0.06	-0.04	0.01	145
	d'	0.05	0.09	0.06	
Adults	Hits	-0.23	0.03	0.39*	
	False Alarms	-0.15	-0.24	-0.14	37
	d'	-0.11	0.13	0.36*	

** $p < .01$; * $p < .05$

As can be seen from Table 6, hit rate for the entire group correlated negatively with calculation time for commutative *and* control problems. That is, the faster second and third graders solved the arithmetic problems, the better they were able to identify commutative problems in the judgment task. Savings in solution time due to commutative problems were not related to the ability to identify commutative problems, suggesting that their knowledge about commutativity was not very well integrated.

A closer look at the different age groups revealed, however, that adults showed the expected positive correlation between savings and sensitivity. Adults who applied the commutativity-based shortcut in the arithmetic blocks were also those who were better able to identify the commutative problems in the judgment task. This correlation suggests that the tested adults do possess an integrated knowledge representation of the commutativity principle.

In contrast, second and third graders' procedural and conceptual knowledge were only weakly related at best. As Table 6 additionally reveals, second graders' hit rate correlated negatively with calculation time. Again, this correlation suggests that the faster second graders solved the arithmetic problems the better they were able to identify the commutative problems in the judgment task. Thus, second graders' ability to discriminate between commutative and control problems was linked to more general calculation competencies rather than to their procedural knowledge about using the commutativity-based shortcut. Third graders, by contrast, did not show any significant correlation between calculation performance and discrimination.

Overall, these findings suggest that only adults' spontaneous application of commutativity knowledge is based on an integrated concept of the commutativity principle. In contrast, procedural and conceptual knowledge seem to be only weakly related in second and third graders. Note that alternatively, one also could argue that our assessments of procedural and conceptual knowledge are not sufficiently reliable (but, see Table 2). In order to further rule out this latter argument and to better understand the missing correlations between savings in the arithmetic task (procedural) and the sensitivity index in the judgment task (conceptual knowledge), we conducted a final fine grained analysis for second and third graders.

In the judgment task, false alarms rate of second and third graders was rather high (approximately 40%; Table 5) and differed largely between participants in both age groups. Presumably, children with a high false alarms rate might have correctly recognized the commutative problems in the judgment task, but at the same time might have believed that also easy to calculate problems (i.e., those with comparatively small addends) needed no calculation. This might have inflated false alarms rate and thus might have reduced the correlations between procedural and conceptual knowledge within second and third graders.

Following up on these assumptions, we divided the second and third graders into three different groups according to their false alarm rate: Children with no false alarms, with up to 50% false alarms, and children with a false alarm rate higher than 50%. Table 7 presents the number of participants within these three groups as well as the hit rates separately per grade.

Table 7

Mean hit rates for second and third graders with no (FA = 0), medium (FA ≤ 50%), or high (FA > 50%) false alarms rate in the judgment task

	False Alarm rate					
	No false alarms		Medium FA-rate		High FA-rate	
	Hit rate	N	Hit rate	N	Hit rate	N
Grade 2	.81	41	.52	61	.85	44
Grade 3	.88	44	.73	60	.84	41

As can be seen from Table 7, for second and third graders hit rate was high when either the false alarm rate was low or when the false alarm rate was high (i.e., some children marked only the commutative problems while others marked the commutative and many other problems). This might have caused the overall low correlations between procedural and conceptual knowledge within these two age groups. Therefore, we re-analyzed the correlation between hit rates and d' and arithmetic abilities separately for these three groups within second and third graders. Table 8 presents these correlations. In both age groups, only those participants who produced high hit rates without incorrectly marking the filler problems also showed substantial correlations. However, second and third graders differed qualitatively with regard to these correlations.

Table 8

Correlations between procedural and conceptual knowledge for second and third graders with no, medium, or high false alarms rate. Procedural knowledge is indicated by calculation times in seconds for commutative and control problems as well as savings. Hit and false alarms rate (FA) indicate conceptual knowledge

	Grade 2					
	No false alarms (N = 41)		Medium FA-rate (N = 61)		High FA-rate (N = 44)	
	Hits	FA	Hits	FA	Hits	FA
Commutative	-.50**	--	-.11	.07	-.01	-.01
Control	-.39**	--	-.11	-.10	.02	-.04
Savings	.03	--	-.11	-.15	.04	.04
	Grade 3					
	No false alarms (N = 44)		Medium FA-rate (N = 60)		High FA-rate (N = 41)	
	Hits	FA	Hits	FA	Hits	FA
Commutative	-.25	--	.04	.12	.04	-.22
Control	-.01	--	.07	.08	-.04	-.18
Savings	.31*	--	.05	-.03	-.02	-.04

** $p < .01$; * $p < .05$

Once again, the results suggest that the second graders' ability to discriminate between commutative and control problems in the judgment task is mainly related to their general calculation skills rather than to their ability to rely on efficient calculation strategies (i.e., the commutativity-based shortcut strategy). By contrast, third graders with high discrimination abilities seem to already possess integrated procedural and conceptual knowledge, starting to form an abstract understanding of commutativity. They use this knowledge, on the one hand, to identify commutative problems in the context of control problems and, on the other hand, to increase efficiency in solving arithmetic problems.

4. Discussion

With the current study we aimed at presenting an approach to unobtrusively measure the spontaneous usage of procedural and conceptual knowledge of the commutativity principle. Apart from providing a basis to develop the method further (see below), the second goal of our study was to investigate the relation between procedural and conceptual knowledge about the commutativity principle in second and third graders. For this we asked (1) at which grade the different forms of commutativity knowledge can be detected and (2) at which grade they start to correlate with one-another.

Overall, our study yielded three main results: First, as expected, third graders showed higher general calculation proficiency (procedural knowledge) and more conceptual knowledge about commutativity than second graders. Second, a solution time benefit based on the procedural use of the commutativity principle was only found for second graders. They calculated commutative problems faster than control problems. Neither calculation times of third graders nor of adult students reflected significant profit from interspersed commutative problems. Third, the correlation between (a) the benefit resulting from a commutativity-based shortcut and (b) conceptual knowledge of commutativity was rather weak in second and third graders. The relation seems to arise in some of the third graders. The link was also present in the control group (adults). The second and third findings merit some further discussion before we come to the theoretical and practical implications.

The second finding (i.e., that only second graders' calculation performance reflected the exploitation of commutativity whereas that of third graders and adults did not) was somewhat surprising. Interestingly however, Gaschler, Vaterrodt, Frensch, Eichler, and Haider (2013) found similar patterns of results with the identical arithmetic task. Therefore, we assume that this finding is not due to a sample artefact. Nevertheless, it does not fit the general claim that with experience, children become faster and more accurate at solving addition problems and also tend to use more sophisticated strategies, such as order-irrelevant, decomposition, and retrieval strategies (Baroody et al., 1983; Canobi et al., 1998, 2002; 2003; Geary, Brown & Samaranayake, 1991; Goldman, Mertz & Pellegrino, 1989; Resnick, 1992; Rittle-Johnson & Siegler, 1998; Siegler, 1987; but see, McNeil, 2007; Robinson & Dubé, 2009; Robinson & Ninowski, 2003; Torbeyns et al., 2009). It also seems to contradict the results of Baroody et al. (1983), which show that approximately 80% of their third graders applied the commutativity-based shortcut to solve arithmetic problems (see also, Canobi et al., 2003).

One obvious reason for these divergent findings might be that we used three-element addition problems which probably were hard for second graders but (due to the rather small addends) easy for third graders and adults. This may have caused second graders to rely on the more efficient commutativity-based shortcut strategy, whereas third graders and adults were fast in solving the problems anyway, so that they did not consider any gain through using the shortcut strategy. For instance, Siegler and Araya (2005) mentioned that participants are more likely to adopt solution strategies if they contribute to significant performance advantages. This argument is further supported by the results of Gaschler et al. (2013) who found larger benefits when presenting problems with large rather than with small addends.

A second reason might be that in former studies (e.g., Baroody et al., 1983; Canobi et al., 1998; Farrington-Flint, Canobi, Wood & Faulkner, 2010) students were instructed to explain their strategy immediately after responding. By contrast, our participants received no hint about the existence of commutative problems. While in our study the use of any shortcut strategy was spontaneous, it is possible that the explanation required in the Baroody et al.'s study might have triggered students to apply the commutativity-based strategy. For instance, Torbeyns et al. (2009) found less strategy application when students could spontaneously apply different strategies during calculation than when they were instructed to do so. In a similar vein, a yet unpublished study from our labs revealed that second and third graders as well as adults substantially benefitted from instruction (compared to a non-instructed group). Participants reminded of the commutativity principle and alerted to the fact that commutative problems might occur, showed a larger solution time advantage on commutative problems as compared to control problems. As students of all three age groups relied on the commutativity principle after being instructed accordingly, it seems justified to conclude that (with the exception of second graders) students in our study indeed did not profit much from spontaneously applying the commutativity-based shortcut strategy.

Concerning our second research question, we found that the second graders' understanding of commutativity was unrelated to their use of commutativity-based shortcut strategies. First signs of an integrated concept (assessed by the correlation of procedural and conceptual knowledge measures) occurred in a small group of third graders and were substantial only for adult students. Thus, the integration of procedural and conceptual knowledge seems to increase with age. However, it also suggests that second graders may have used the shortcut strategy without entirely understanding the commutativity principle. This finding seems at odds with the early onset assumption of, for example, Baroody and Gannon (1984). Furthermore, Canobi et al. (1998; 2002) had found that second graders' conceptual (assessed by an explanation task) and procedural knowledge (assessed by solving addition problems) correlated moderately. However, as already discussed concerning Baroody et al.'s (1983) findings, Canobi (2009; Canobi et al., 1998, 2002) assessed conceptual knowledge by asking participants to explain their strategies after they had solved an addition problem. Thus, even though Canobi et al. used different tasks for assessing procedural and conceptual knowledge, the knowledge assessed by their addition task might have resulted from a mixture of procedural and conceptual competencies (see also, Robinson & Dubé, 2009). This might have led to a higher correlation between both tests compared to a variant were spontaneous application of procedural and conceptual knowledge is independently assessed.

Alternatively, one might suspect that our measures of procedural and conceptual knowledge were unreliable. However, this is not likely as we did find satisfying split-half reliabilities for all age groups and both task formats (see, Table 2). In addition, our results showed significant correlations (a) for all second graders between the calculation time and percentage of hits and (b) for at least some third graders and all adult students between savings due to the use of commutativity-based shortcuts and hits in the judgment task. Therefore, it seems worthwhile to ask for further theoretical causes concerning our third finding.

4.1 Theoretical Implications

At first glance, the results seem to fit with a procedural-first development of commutativity (e.g., Baroody et al., 2007; Briars & Siegler, 1984; Siegler & Stern, 1998). That is, second graders use the commutativity-based shortcut before having acquired an abstract understanding of the principle. It therefore appears that the development of conceptual knowledge progresses more slowly than that of procedural knowledge – at least as measured in this study and for commutativity (cf. Canobi, 2004; Canobi et al., 1998). However, in their review about relations between children's understanding of mathematical concepts and their ability to execute arithmetic procedures, Rittle-Johnson and Siegler (1998) provided ample evidence that with regard to commutativity, children first acquire conceptual knowledge before then applying corresponding strategies.

In order to reconcile this conflict, we refer to the iterative model of the development of conceptual and procedural knowledge (e.g., Resnick, 1992; Rittle-Johnson et al., 2001). Our findings suggest that second graders possess at least rudimentary conceptual knowledge of the commutativity principle, but their conceptual representation of the commutativity principle is less well integrated (with procedural knowledge) than that of third graders or adult students. This is in line with many findings in the field of mathematic development showing that already second graders possess conceptual knowledge about commutativity (for a review, see Rittle-Johnson & Siegler, 1998). Also, our sensitivity index d' indicated such knowledge. However, the consistent use of this knowledge may still be reduced; that is, their competency to identify the relevant task properties for applying a certain shortcut strategy has not fully developed yet. Therefore, they may need a certain external trigger in order to activate their knowledge about commutativity and the corresponding strategies. Our instructions for the arithmetic and the judgment task did not provide any such trigger which probably made it rather difficult, particularly for second graders and also for most of the third graders, to realize that they should rely on the commutativity principle in both tasks. Consequently, it may be that some participants applied the commutativity-based shortcut strategy to solve the arithmetic problems, but did not use it in the judgment tasks or vice versa. This does not imply that they first learn procedures before they acquire conceptual knowledge. Rather, we assume that such a finding mainly reflects that children's conceptual knowledge is not sufficiently integrated to spontaneously recognize that they could rely on the commutativity principle. In a similar vein, research on expertise (e.g., Anderson & Schunn, 2000; Gentner & Toupin, 1986; Haider & Frensch, 1996; Koedinger & Anderson, 1990) also shows that well-integrated and thus abstract conceptual knowledge is required to identify task relevant information in order to solve problems and to flexibly transfer knowledge from one task domain to another (see also e.g., Sloutsky & Fisher, 2008; Star & Seifert, 2006).

To summarize, we assume that the divergent findings concerning the development of an abstract understanding of the commutativity principle reflect the fact that after students have acquired some procedural and conceptual knowledge in this domain, this knowledge needs to be integrated. This integration of knowledge, we suspect, is done in an iterative way which means that procedures are applied, which then refine the conceptual knowledge (Rittle-Johnson et al., 2001). The conceptual knowledge is then used to guide children's attention to information which is needed to adaptively apply efficient strategies.

4.2 Further Improvements of the Measurement of Spontaneous Application of a Mathematical Principle

With the current study, we took a first step to measure spontaneous usage of commutativity knowledge in a non-reactive way. Participants worked on the paper-and-pencil tasks in a setting very similar to other tests in the classroom. In the arithmetic blocks, we asked for fast and correct solutions to the arithmetic problems and did not mention that regularities in the task material might be exploited for efficient task processing. We inferred procedural knowledge of the commutativity principle from the performance benefits on material containing identical addends in changed order in consecutive problems (as compared to material that did not contain such pairs of problems). Probing for conceptual knowledge, we asked participants to indicate in which cases calculation was not necessary – again without hinting that it might be the commutativity principle that made calculation superfluous. The rationale behind this procedure was that participants who had well integrated knowledge of the commutativity principle should recognize the respective arithmetic problems and consequently should be able to relate it to the task demand (marking problems where calculation was not necessary). Age-related changes in hits and false alarms in the judgment task suggested that this was indeed the case. Similar indirect approaches to measure knowledge have been developed in order to measure insight (cf. Haider & Rose, 2007). When investigating insight, it is not feasible either to directly ask participants again and again if they already have discovered the regularity in the task material – without providing them with a strong hint that such a regularity exists.

One way to further improve the method would be to include control problems which also feature the same numbers as their predecessor problems in changed order but do not allow to apply the commutativity principle (i.e. subtractions). Such interspersed control problems could help to rule out superficial matching strategies (i.e. "same numbers = same result") that do not capture the essence of commutativity. First explorations in our labs indicate that second graders do not confuse subtraction problems containing the same digits as a preceding addition with genuine commutative problems. Furthermore, it would be interesting to implement our instruments within a multi-method approach, administering multiple measurements per person and construct (cf. Prather & Alibali, 2009). As a first step one would have to estimate to what extend repeated testing of spontaneous usage of a mathematical principle induces participants to recognize and use the principle – and by this spoils the possibility to assess spontaneous usage.

Paper-and-pencil-based testing in the classroom has the advantage that the test situation is similar to other tests the students take. In a parallel line of research, we have started to employ eyetracking to obtain process measures related to commutativity knowledge (e.g., Gaschler et al. 2013; Godau, Wirth, Hansen, Haider, & Gaschler, in press). For instance, it is possible to quantify the extent to which a child searches for repetitions of addends in subsequent addition problems. However, when they are tested individually with an eyetracking system, children are aware that the measurement is about where they look and how they calculate. Class-based testing in computer labs within schools might offer the possibility to obtain process measures while keeping up the character of the assessment as allowing to measure spontaneous application of the principle knowledge.

4.3 Theoretical Conclusions and Practical Implications

Recently, Prather and Alibali (2009; see also, Bisanz et al., 2009; Schneider & Stern, 2010) called for multifaceted knowledge assessment in the context of arithmetic development. Our use of the arithmetic and judgment tasks in order to independently assess procedural and conceptual knowledge can be seen as a first step in this direction. Complementing earlier work on commutativity knowledge (e.g., Baroody & Gannon, 1984; Canobi et al., 2002), our findings suggest that, when second graders and third graders are not alluded to rely on the commutativity principle, second graders and most of the third graders show but weak signs of spontaneous application of commutativity and interrelation of different forms of commutativity knowledge. This suggests that they do not possess well-integrated knowledge about commutativity in the sense of an abstract formal mathematical principle.

Our results suggest that even if children use procedures that suggest integrated conceptual knowledge about commutativity, the learning process has by far not reached an endpoint. Rather, it still progresses, before leading to a well-integrated, abstract representation of the mathematical principle as with our measures found in adults. As long as children do not possess such an abstract representation, they will not be able to flexibly and adaptively use the commutativity principle in different task contexts. Accordingly, we suspect that increasing experience in the field of mathematics is needed in order to better integrate conceptual knowledge about various arithmetic principles. This might explain why transfer of knowledge from one context to another is often found to be rather weak (e.g., Frensch & Haider, 2008; Kaminsky, Sloutsky & Heckler, 2008; Siegler & Stern, 1998; Sloutsky & Fisher, 2008). Therefore, helping students to develop well-integrated knowledge concepts should be one of the most important tasks education has to fulfill (see, e.g., Geary et al., 2008; Prather & Alibali, 2009; Verschaffel et al. 2009).

In more practical terms, if children are taught the commutativity principle in the context of addition, they seem to learn that they can use this principle to avoid unnecessary labor. However, our results suggest that this does not mean that they concurrently acquire an idea of the abstract principle of cardinality. We suspect that many children only acquire a procedure (or a strategy) that they can easily apply for two-element addition problems. In order to help students to understand the abstract principle of commutativity, it might be worthwhile to activate students' prior knowledge of this principle, such as the order-irrelevance principle they already use in counting. When introducing the commutativity principle in addition (or multiplication) it might be helpful to tell students that they already have used this principle in other contexts and explain how and why it works in all these different situations. This then might help them to understand the commutativity principle in a more abstract manner and probably also to understand task properties needed to correctly apply this principle. Further, it may help to support children in recognizing the consequences of using alternative strategies in order to ensure representational redescription (e.g., Baroody & Gannon, 1984).

Keypoints

- Procedural and conceptual commutativity knowledge increase with increasing age.
- Second graders show no signs of an integrated concept of commutativity.
- First signs of an integrated concept of commutativity emerge in grade three.

Acknowledgements

This research was supported by the German Research Foundation (DFG; HH- 1471/12-1). Some of the results were presented at the Kongress der Deutschen Gesellschaft für Psychologie 2010 in Bremen, Germany. We thank Annette Bräutigam, Yvonne Radermacher, Pia Blase and Ester Jung for help with data collection.

References

Anderson, J. R. & Schunn, C. D. (2000). Implications of the act-r learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 1-33). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Baroody, A. J. (1984). More precisely defining and measuring the order-irrelevance principle. *Journal of Experimental Child Psychology, 38*, 33-41. doi:10.1016/0022-0965(84)90017-1

Baroody, A. J. (2003). The Development of Adaptive Expertise and Flexibility: The Integration of Conceptual and Procedural Knowledge. In A. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills. Constructing Adaptive Expertise* (1st ed., pp. 1–34). Mahwah, NJ: Lawrence Erlbaum Associate Publishers.

Baroody, A. J., Feil, Y. & Johnson, A. R. (2007). An alternative reconceptualization of procedural and conceptual knowledge. *Journal for Research in Mathematics Education, 38*, 115-131.

Baroody, A. J. & Gannon, K. E. (1984). The development of the commutativity principle and economical addition strategies. *Cognition & Instruction, 1*, 321-339. doi:10.1207/s1532690xci0103_3

Baroody, A. J., Ginsburg, H. P. & Waxman, B. (1983). Children's use of mathematical structure. *Journal for Research in Mathematics Education, 14*, 156-168. doi:10.2307/748379

Baroody, A.J., & Rosu, L. (2006). *Adaptive expertise with basic addition and subtraction combinations - The number sense view*. Paper presented at theAnnual Meeting of the American Educational Research Association (April) San Francisco, CA.

Bisanz, J., & LeFevre, J. (1992). Understanding elementary mathematics. In J. D. Campbell (Ed.) , *The nature and origins of mathematical skills* (pp. 113-136). Oxford England: North-Holland. doi:10.1016/S0166-4115(08)60885-7

Bisanz, J., Watchorn, R. P. D., Piatt, C., & Sherman, J. (2009). On “Understanding” Children's Developing Use of Inversion. *Mathematical Thinking and Learning, 11*, 10–24. doi:10.1080/10986060802583907

Briars, D. & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology, 20*, 607-618. doi:10.1037/0012-1649.20.4.607

Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development, 19*, 81-93. doi:10.1016/j.cogdev.2003.10.001

Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology, 92*, 220-246. doi:10.1016/j.jecp.2005.06.001

Canobi, K. H. (2009). Concept-procedure interactions in children's addition and subtraction. *Journal Of Experimental Child Psychology*, *102*(2), 131-149. doi:10.1016/j.jecp.2008.07.008

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. doi:10.1080/01443410903473597

Canobi, K. H., Reeve, R. A. & Pattison, P. E. (2002). Young children’s understanding of addition concepts. *Educational Psychology, 22*, 513-532. doi:10.1080/0144341022000023608

Canobi, K. H., Reeve, R. A. & Pattison, P. E. (2003). Patterns of knowledge in children’s addition. *Developmental Psychology, 39*, 521–534. doi:10.1037/0012-1649.39.3.521

Cowan, R. & Renton, M. (1996). Do they know what they are doing? Children's use of economical addition strategies and knowledge of commutativity. *Educational Psychology, 16*, 407-420. doi:10.1080/0144341960160405

Farrington-Flint, L., Canobi, K.H., Wood, C. & Faulkner, D. (2010). Children’s patterns of reasoning about reading and addition concepts. *British Journal of Developmental Psychology, 28*, 427–448. doi:10.1348/026151009X424222

Frensch, P. A. & Haider, H. (2008). Transfer and Expertise: The Search for Identical Elements. In H. L. Roediger, III (Ed.), *Cognitive Psychology of Memory. Vol. [2] of Learning and Memory: A Comprehensive Reference* (pp. 579-596) Oxford: Elsevier. doi:10.1016/B978-012370509-9.00177-7

Fuson, K. C. (1988). *Children's counting and concepts of number*. New York, NY: Springer-Verlag Publishing.

Gaschler, R., Vaterrodt, B., Frensch, P. A., Eichler, A. & Haider, H. (2013). Spontaneous usage of different shortcuts based on the commutativity principle. *PLoS ONE 8*(9): e74972. doi:10.1371/journal.pone.0074972

Geary, D. C. (2006). Development of Mathematical Understanding. In W. Damon, R. M. Lerner, & N. Eisenberg (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (Vol. 3, pp. 777–810). John Wiley and Sons.

Geary, D. C., Boykin, A. W., Embretson, S., Reyna, V., Siegler, R., Berch, D. B., et al. (2008). Report of the task group on learning processes. In *National mathematics advisory panel, reports of the task groups and subcommittees (pp. 4-1–4-211)*.

Geary, D. C., Brown, S. C. & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 27,* 787-797. doi:10.1037/0012-1649.27.5.787

Geary, D. C., Hoard, M. K., Byrd-Craven, J. & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88*, 121–151. doi:10.1016/j.jecp.2004.03.002

Gelman, R. (1990). First principles organize attention to and learning about relevant data: Number and the animate-inanimate distinction as examples. *Cognitive Science, 14,* 79-106.

Gelman, R. & Gallistel, C. R. (1978). The child's understanding of number. In. Cambridge, MA: Harvard University Press.

Gelman, R. & Meck, E. (1983). Preschoolers' counting: Principles before skill. *Cognition, 13*, 343-359. doi:10.1016/0010-0277(83)90014-8

Gentner, D. & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science*, *10*, 277-300. doi:10.1207/s15516709cog1003_2

Godau, C., Wirth, M., Hansen, S., Haider, H., & Gaschler, R. (in press). From marbles to numbers - Estimation influences looking patterns on arithmetic problems. *Psychology.*

Goldman, S. R., Mertz, D. L. & Pellegrino, J. W. (1989). Individual differences in extended practice functions and solution strategies for basic addition facts. *Journal of Educational Psychology, 81*, 481-496. doi:10.1037/0022-0663.81.4.481

Goldstone, R. L., & Sakamoto, Y. (2003). The transfer of abstract principles governing complex adaptive systems. *Cognitive Psychology*, *46,* 414-466. doi:10.1016/S0010-0285(02)00519-4

Haider, H. & Frensch, P. A. (1996). The role of information reduction in skill acquisition. *Cognitive Psychology, 30,* 304-337. doi:10.1006/cogp.1996.0009

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

Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction*, *15*, 237-256. doi:10.1016/j.learninstruc.2005.04.005

Hannula, M. M., Lepola, J., & Lehtinen, E. (2010). Spontaneous focusing on numerosity as a domain-specific predictor of arithmetical skills. *Journal of Experimental Child Psychology*, *107*, 394-406. doi:10.1016/j.jecp.2010.06.004

Hiebert, J., & 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). Hillsdale, NJ: Erlbaum.

Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). Learning Theory: The Advantage of Abstract Examples in Learning Math. *Science*, *320*, 454–455.

Koedinger, K. R. & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, *14*, 511-550. doi:10.1207/s15516709cog1404_2

Kuhn, D., Garcia-Mila, M., Zohar, A., & Andersen, C. (1995). Strategies of knowledge acquisition. *Society for Research in Child Development Monographs*, *60* (4), Serial no. 245.

LeFevre, J.-A., Smith-Chant, B. L., Fast, L., Skwarchuk, S.-L., Sargla, E., Arnup, J. S., et al. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through grade 2. *Journal of Experimental Child Psychology*, *93*, 285-303. doi:10.1016/j.jecp.2005.11.002

McMullen, J.A., Hannula-Sormunen, M.M., & Lehtinen, E. (2011). Young children’s spontaneous focusing on quantitative aspects and their verbalizations of their quantitative reasoning. In Ubuz, B. (Ed.). *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education*, *3*, pp. 217-224. Ankara, Turkey: PME.

McNeil, N. M. (2007). U-shaped development in math: 7-year-olds outperform 9-year-olds on equivalence problems. *Developmental Psychology, 43*, 687–695. doi:10.1037/0012-1649.43.3.687

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:10.1016/j.dr.2009.09.001

Resnick, L. B. (1992). From protoquantities to operators: Building mathematical competence on a foundation of everyday knowledge. In G. Leinhardt, R. T. Putnam & R. A. Hattrup (Eds.), *Analysis of arithmetic for mathematics teaching* (pp. 373-429). Hillsdale, NJ, UK: Lawrence Erlbaum Associates, Inc.

Rittle-Johnson, B. & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 75-110)). East Sussex, UK: Psychology Press.

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. doi:10.1037/0022-0663.93.2.346

Robinson, K. M. & Dubé, A. K. (2009). Children’s understanding of addition and subtraction concepts. *Journal of Experimental Child Psychology, 103*, 532–545. doi:10.1016/j.jecp.2008.12.002

Robinson, K. M., & Ninowski, J. E. (2003). Adults’s 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. doi:10.1037/h0087435

Schneider, M. &. Stern, E. (2010). The Developmental Relations Between Conceptual and Procedural Knowledge: A Multimethod Approach. *Developmental Psychology*, *46*, 178–192. doi:10.1037/a0016701

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. doi:10.1037/0096-3445.116.3.250

Siegler, R. S. & Araya, R. (2005). A computational model of conscious and unconscious strategy discovery. In R. V. Kail (Ed.), *Advances in child development and behavior (vol. 33)* (pp. 1-42). Oxford, UK: Elsevier.

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

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

Sloutsky, V. M. & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development, 79*, 639-651. doi:10.1111/j.1467-8624.2008.01148.x

Sophian, C. & Adams, N. (1987). Infants' understanding of numerical transformations. *British Journal of Developmental Psychology, 5*, 257-264. doi:10.1111/j.2044-835X.1987.tb01061.x

Sophian, C., Harley, H. & Manos Martin, C. S. (1995). Relational and representational aspects of early number development. *Cognition & Instruction, 13*, 253-268. doi:10.1207/s1532690xci1302_4

Star, J.R. (2004, April). *The development of flexible procedural knowledge in equation solving.* Paper presented at the annual meeting of the American Educational Research Association, San Diego.

Star, J. R. (2005). Reconceptualizing procedural knowledge. *Journal for Research in Mathematics Education, 36*, 404-411.

Star, J. R. & Seifert, C. (2006). The development of flexibility in equation solving. *Contemporary Educational Psychology, 31*, 280-300. doi:10.1016/j.cedpsych.2005.08.001

Starkey, P. & Gelman, R. (1982). 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 and subtraction: A cognitive perspective* (pp. 99–116). Hillsdale, NJ: Erlbaum.

Torbeyns, J., De Smedt, B., Ghesquière, P. &Verschaffel, L. (2009). Acquisition and use of shortcut strategies by traditionally schooled children. *Educational Studies in Mathematics, 71*(1), 1-17.

Verschaffel, L., Luwel, K., Torbeyns, J. & Van Dooren, W. (2009). Conceptualizing, investigating, and enhancing adaptive expertise in elementary mathematics education. *European Journal of Psychology of Education*, *24*(3), 335-359. doi:10.1007/BF03174765