The potential of symbolic approximation.

Disentangling the effects of approximation vs. calculation demands in nonsymbolic and symbolic representations

Inauguraldissertation

zur Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionsordnung vom 10.05.2010

vorgelegt von

Sonja Maria Hansen

aus

Bochum

Juni 2016

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im Oktober 2016 angenommen.

Abstract

Mathematical skills that we acquire during formal education mostly entail exact numerical processing. Besides this specifically human faculty, an additional system exists to represent and manipulate quantities in an approximate manner. We share this innate approximate number system (ANS) with other nonhuman animals and are able to use it to process large numerosities long before we can master the formal algorithms taught in school. Dehaene´s (1992) Triple Code Model (TCM) states that also after the onset of formal education, approximate processing is carried out in this analogue magnitude code no matter if the original problem was presented nonsymbolically or symbolically. Despite the wide acceptance of the model, most research only uses nonsymbolic tasks to assess ANS acuity. Due to this silent assumption that genuine approximation can only be tested with nonsymbolic presentations, up to now important implications in research domains of high practical relevance remain unclear, and existing potential is not fully exploited. For instance, it has been found that nonsymbolic approximation can predict math achievement one year later (Gilmore, McCarthy, & Spelke, 2010), that it is robust against the detrimental influence of learners´ socioeconomic status (SES), and that it is suited to foster performance in exact arithmetic in the short-term (Hyde, Khanum, & Spelke, 2014). We provided evidence that symbolic approximation might be equally and in some cases even better suited to generate predictions and foster more formal math skills independently of SES.

In two longitudinal studies, we realized exact and approximate arithmetic tasks in both a nonsymbolic and a symbolic format. With first graders, we demonstrated that performance in symbolic approximation at the beginning of term was the only measure consistently not varying according to children´s SES, and among both approximate tasks it was the better predictor for math achievement at the end of first grade. In part, the strong connection seems to come about from mediation through ordinal skills. In two further experiments, we tested the suitability of both approximation formats to induce an arithmetic principle in elementary school children. We found that symbolic approximation was equally effective in making children exploit the additive law of commutativity in a subsequent formal task as a direct instruction. Nonsymbolic approximation on the other hand had no beneficial effect. The positive influence of the symbolic approxi-

mate induction was strongest in children just starting school and decreased with age. However, even third graders still profited from the induction.

The results show that also symbolic problems can be processed as genuine approximation, but that beyond that they have their own specific value with regard to didactic-educational concerns. Our findings furthermore demonstrate that the two often confounded factors 'format' and 'demanded accuracy' cannot be disentangled easily in first graders numerical understanding, but that children´s SES also influences existing interrelations between the different abilities tested here.

Content

Figures

Tables

Appendices

1 Introduction

Mathematical competencies are an indispensable part of our daily life. This is not only the case in the business world where mathematical abilities are necessary for not falling victim to fraud, or to plan your investments and wins realistically. No, it is even more fundamental than that: From early on, we need to understand what 'sharing' means and at least estimate how we have to divide things according to the number of people we want to share them with. We try to make the smartest choice between alternative offers in terms of quantity; and soon we have to coordinate our time with our obligations or our food stocks with our consumption as well as with the next opportunity to refill them.

These are only a few examples of situations in which quantitative or even mathematical abilities play a role in our everyday life. These examples also indicate that different degrees of accuracy can be applied in mathematical processing. In business contexts we may aim to compute and plan as exactly as possible, whereas we probably administer a rule of thumb estimate when it comes to shopping for the household (at least as long as we do not have to worry too much about our financial situation…).

But what does 'estimation' actually mean? Geary (2006, p. 793) provides the following definition:

"*Estimation typically involves some type of procedure to generate an approximate answer to a problem when calculation of an exact answer is too difficult or is unnecessary.*"

As can already be deduced from Geary´s very general definition, 'estimation' is a term that is widely used in various contexts. You estimate the approximate result of a sum, the distance you have to walk to the closest grocery store, the numbers of peas in a jar, or how long you will need to finish a piece of work. Not enough of the confusion, Dowker (2012) furthermore states that estimation is not a unitary process, but consists of numerous components and ranges from mere quantification without the need of any kind of calculation up to computational (or arithmetical) estimation that includes the use of derived fact strategies as well as estimation techniques like rounding, truncation or compensation.

We would rather call the processes this thesis will be about 'approximation', instead of estimation. We will describe that approximation is not only an umbrella term to all quantitative processing that is carried out without the explicit goal to produce an exact result; something that we probably do intuitively when the time or ability for exact processing is lacking. Instead, approximation as a 'coarse, intuitive sense of magnitude' (Ganor-Stern, 2015), that is the ability to grasp and process quantities in an imprecise manner, can be found not only in humans already in the first hours of their lives (Bijeljac-Babic, Bertoncini, & Mehler, 1993), but also in other nonhuman animals like rats, pigeons, or monkeys (Gallistel, 1990; Meck & Church, 1983). The evolutionary use of the ability to approximate and to compare numbers is obvious when we think of ingroup and outgroup encounters of rivaling herds. In situations like this, quantity might be one important factor in an animal´s 'decision' to fight or flight.

Curiously, despite the importance and the almost automatic turning towards approximation when nobody demands an exact result from us, approximate processing has long been neglected in research (Booth & Siegler, 2006; Dehaene, 1992). While the cognitive background of exact number processing like counting and arithmetic has been studied already in the first half of the 20th century (i.e. Brownell, 1935), actual approximation processes have first been studied considerably later (Rubinstein, 1985) and only reached the focus of attention in the 21st century (Dowker, 2003). Since then, however, a lot has been discovered about these so-called 'primary quantitative competencies' (Geary, 2006) or 'preverbal numerical knowledge' (Dehaene, 1992). In the next chapter, some important differences between this knowledge and the so-called formal calculation processes will be outlined.

2 Qualitative differences of quantitative processing: An overview about numerical abilities

We often take the quantitative skills that we employ in everyday life for granted. When we hear the term 'math' or 'arithmetic', we rather tend to think of the rules, constraints and algorithms we have learned in school. And indeed there are important characteristics and differences between the way adults carry out complex mathematics (formal calculation processes) and the abilities that already infants and even animals use (preverbal numerical knowledge). To avoid the ambiguous meaning of the word 'number', we follow Stanislas Dehaene´s (1992) suggestion and speak of the actual physical quantity with the term 'numerosity', and use the term 'numeron' when we are talking of a mental representative of a given numerosity (which may be an Arabic number symbol, or a number word). In the following, we will give a short overview of formal calculation processes and then contrast them with preverbal numerical knowledge. We will describe empirical findings which flowed into the most influential account to connect formal and informal numerical skills in one model, the 'Triple Code Model' (TCM) by Dehaene (1992; Dehaene, Piazza, Pinel, & Cohen, 2003). This model, which will be described in more detail after an overview about the empirical findings that formed its basis, is about the quality of a quantity´s representation in the learner on the one hand. We will learn from the next sections that numerical magnitudes can be represented in a format independent manner as well as in a symbolic way. The latter can consist of Arabic numerals, of number words, or other numerons an individual might know to represent quantities. On the other hand, in the TCM it is distinguished between the varying degrees of accuracy that is demanded in different kinds of operations. From the findings that are reported below, Dehaene concluded that approximate processing is carried out in the format independent representational code, while exact processing always happens with some kind of assigned symbols or numerons.

Formal calculation processes as we know them depend and are based on *positional number notations* (Dehaene, 1992). Understanding this numerical notation system requires at least two competencies: First, we have to gain lexical knowledge about the – in our case – Arabic number notation system to know which number symbol belongs to which numerosity. Second, we also need syntactic knowledge about how the numerons (here the Arabic numbers) are constructed in our notational system. For example, we have to understand that the position of each digit indicates its power of tens (e.g., in the case of 345 we know that 5 represents the units, the 4 has to be multiplied by ten, and the 3 has to be multiplied by hundred), and that the resulting numerons then have to be added (300 + 40 + 5 = 345). These skills depend on the human language faculty and cannot be acquired before general language acquisition has started. During the first acquisition phase, lexical errors (e.g., counting *one, two, four*) as well as syntactical errors (overgeneralization of rules, like *twentyten, twentyeleven* etc.) can occur in children. Furthermore, dissociations are observed between both skills in brain damaged adults, demonstrating the distinctiveness of the two competencies (Dehaene, 1992).

The formal calculation procedures differ regarding the number space in which they are performed, that is, with the size of the operands. For multi-digit calculation, we are taught efficient algorithms, e.g. to multiply two three-digit numbers like 345 x 271. Using these algorithms can be dissociated from single-digit calculation in that it is mostly carried out mechanically. No or only little semantic awareness of the actual numerosities one is manipulating is needed, or of the rules that (hopefully) make sure that the numerosities enclosed in the number symbols are preserved in our calculation (Dehaene, 1992). Single-digit calculation on the other hand often seems to be performed by retrieving memorized facts, e.g. in form of stored addition and multiplication tables. This kind of retrieval is supported by findings of spreading activation in these tables: An incorrect result of a given problem takes longer to be falsified when it is in the same row or column as the correct results (Koshmider & Ashcraft, 1991). After some years of formal schooling, frequently encountered problems in the low number range are so overlearned that the arithmetic store actually seems to be accessed automatically. This is for example visible in findings that the time to falsify a result increases when the suggested number is the result of a different operation with the same numbers (e.g., 4 + 6 = 24, see Dehaene, 1992). Before children are able to retrieve solutions to frequently occurring arithmetic problems from memory (that is, the time that they have built this arithmetic store), they use various calculation strategies. Usually, counting-based strategies dominate their efforts at first (Dehaene, 1992; Geary, 2006). With concrete objects, 3 year-old children use manipulations (e.g., pointing to each object) to create a one-to-one correspondence of elements and counting words. In kindergarten age, children count both operands to solve addition problems. This can either happen with the CAL procedure ('counting all', e.g. for the problem 3 + 4 they would count *one, two, **three**,* *four, five, six,* ***seven***), with CON ('counting on', that is they would directly start with *three, four, five, six,* ***seven***) or COL ('counting on from larger', the more economical way of starting counting with the cardinality of the larger addend, no matter if it actually was the first operand in the presented problem: *four, five, six,* ***seven***). These counting-based strategies lead to memory representations of the basic number facts and thus later enable children to directly retrieve solutions or at least decompose problems to components they have a stored solution for (for example, 6 + 7 becomes 6 + 6 + 1 because the learner already knows that 6 + 6 is twelve). Despite the general trend to develop from counting-based to memory-based strategies in calculation, there is no fixed sequence that each child goes through, not even a stable development *within* a child. Instead, always a strategy mix is used; a mix that may contain more counting-based procedures in younger than in older children (Siegler and Booth, 2005, call this the 'overlapping waves model'), but that preserves flexibility on every age level. Even adults will sometimes use the seemingly primitive counting-based strategies, when the problem context makes it the most efficient solution process.

Preverbal numerical knowledge, as opposed to formal calculation processes, enables to *quantification* as its most basic feature. Quantification is the process to capture the numerosity of a given set and the ability to tie an according numeron to it. Dehaene (1992) differentiates the three quantification processes counting, subitizing, and approximation[1].

Genuine ***counting*** (in the sense of enumerating, not only reciting the count list) requires understanding of five principles (Geary, 2006; Gelman & Gallistel, 1978):

(1) one-on-one correspondence, meaning that each numeron can only be used for one of the to-be-counted objects;

(2) stable order, so that the numerons are mapped in a reproducible manner to the elements;

(3) cardinality, that is the understanding that the last numeron of the count represents a property of the entire set – its numerosity;

[1] Dehaene (1992) actually speaks of 'estimation', which is a term we already mentioned in Chapter 1 and which is used more frequently especially in earlier studies in this field. However, as he described the same process that is today known by 'approximation', and to distinguish the meaning of 'estimation' in everyday language from the mathematical context, we stick to the term 'approximation' in the following.

(4) abstraction, the understanding that any collection of all kinds of objects (or even only mental entities) can be counted with the same sequence, even when the set consists of heterogeneous objects; and

(5) order irrelevance, meaning that within a set, an object can be counted at an arbitrary position (e.g., one does not have to count strictly from left to right).

These principles leave the definition of counting quite open when it comes to the quality of the numerons. So children, people from cultures without numerical notations, and even animals could be regarded as adequately counting (Dehaene, 1992). Several studies have explored children's understanding of the five principles and related it to their counting skills. For example, Wynn (1992) let children judge counting processes (in her experiment, a doll did the counting). She found that, while children understand order irrelevance quite early, understanding of cardinality or the abstraction principle are not fully existent when children start counting themselves. Meck and Church (1983) were interested in animals' ability to discriminate numerosities. They succeeded in training a rat to press two different levers depending on hearing two beeps vs. hearing eight beeps. In an older study of Rilling and McDiarmid (1965), it was shown that animals' numerical discrimination can extend to up to 50 events, but with decreasing accuracy.

There are also situations, however, in which quantification seems to differ from classical counting. This becomes visible in experiments measuring reaction times (RT) until a participant names the numerosity of a set of objects presented to him. If he/she counts the objects, one would expect a linear increase in reaction time the larger the numerosity gets. This, however, is only the case for numbers of an intermediate number range (Mandler & Shebo, 1982). For small sets between one and three (sometimes four) objects, RTs remained relatively constant and accuracy high. This often replicated result lead to postulating a separate quantification process for small numerosities up to three or four, the so-called subitizing (Butterworth, 2011; Burr, Turi, & Anobile, 2010; Dehaene, 1992; Siegler & Booth, 2005). This process will be described in the following, before turning to another exception to classical counting in quantification processes, that is, approximation.

Subitizing and approximation. Subitizing in the terms of discriminating between sets of one to three simultaneously presented objects based on numerosity has also been observed in infants well before they learned to speak (Starkey & Cooper, 1980). Up to date, there is no consensual explanation of what mechanism underlies subitizing. The economically convincing theory (Mandler & Shebo, 1982) that numerosities that do not exceed 3 or 4 always result in a canonical geometrical configuration (two dots build a line, three dots build a triangle) may work for the visual domain, but not for auditory sequences for which subitizing has been demonstrated in humans and animals, too (Bijeljac-Babic et al., 1993). Also the hypothesis that subitizing is not qualitatively different from approximation (see below) has been questioned. It seems plausible, given that the accuracy of approximation results decreases the larger the numerosity gets, that subitizing is just approximation with very small numerosities – and hence maximal accuracy – that results in exact quantification for numerosities up to three. However, Burr et al. (2010) found that the mechanism of subitizing and approximation must differ, as only the latter required attentional resources. More recent accounts thus postulate an own distinct system for subitizing, e.g. the 'small numerosity system' (Butterworth, 2011). Others associate subitizing with the ability of multiple object tracking (MOT) (Chesney & Haladjian, 2011), which resembles subitizing in the most prominent feature, that is the restriction to small numerosities. MOT is the mechanism of perceiving up to four moving objects as individual elements and follow them through space and time. These skills have been found to activate different neuronal areas than the ANS (posterior parietal and occipital cortex, see Piazza, 2010), vary between individuals and develop during the first year of life.

In RT experiments like that by Mandler and Shebo (1982), usually another exception from the linear increase with larger quantities is found. Within a given time limit, also for larger numerosities (- depending on the time limit, numerosities larger than 7 suffice to find the following effect -) the RTs do not increase much further, but remain largely constant. Contrary to the subitizing number range, however, at the same time the accuracy of the given answers decreases the larger the set size gets. This process is called ***approximation*** (originally ***estimation,*** Dehaene, 1992). The research of approximation processes since then has yielded remarkable results and elucidated a lot about the kind of representations that enable us to this rapid but imprecise quantification of large numerosities.

One important and robust finding in studies that do not even explicitly tackle questions regarding quantification is the so-called SNARC-effect (spatial numerical association of response codes): When adults had to judge parity of numbers, and sometimes had to indicate the possible answers 'even' and 'odd' with their right vs. their left hand (balanced out), it was found that independently of the answer, they were faster in responding to large numbers with their right and to small numbers with their left hand. The effect has also been found for preschoolers (Hoffmann, Hornung, Martin, & Schiltz, 2013; Patro & Haman, 2012) and supports the metaphor of a mental *number line* that represents quantities oriented from left to right. Further specification of the concept of the mental number line comes from findings of an important difference between approximation perception tasks (like the reported experiments from Mandler and Shebo, 1982) and approximation production tasks (in which participants have to produce an array of dots that approximately matches a given Arabic number). In approximation perception tasks, the numerosity tends to be stronger *under*estimated the higher the quantity gets. On the other hand, in approximation production tasks, with set size the produced arrays are increasingly *over*estimating the given number (Crollen, Castronovo, & Seron, 2011; Mandler & Shebo, 1982). These and other findings indicate that our genuine mental number line is compressed (Booth & Siegler, 2006), consisting rather in a logarithmic than a linear scale (see Figure 1). This would explain the described under- and overestimation results in terms of the bidirectional mapping hypothesis (Castronovo & Seron, 2007; Crollen & Seron, 2012): the underestimation in approximation perception tasks fits a mapping procedure from a logarithmic representation of the dot array to the linear representation of Arabic numbers (Figure 1, a), while the overestimation in the approximation production task becomes explainable by the reverse mapping route from the linear representation onto a logarithmic one (Figure 1, b).

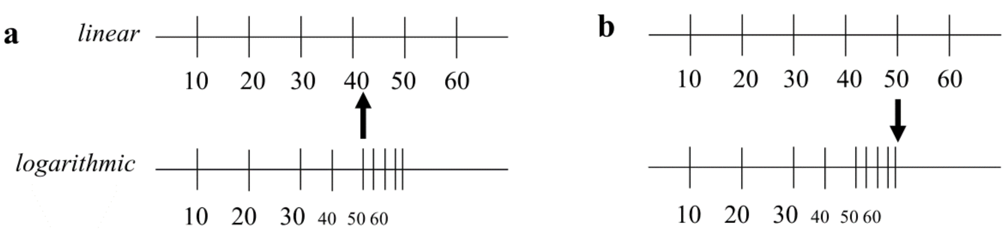

Figure 1. Logarithmic vs. linear number representation. The figure is based on Crollen et al., 2011, p. 41

It was observed that children before and right after starting school at first still map this early logarithmic representation on the newly acquired numerical notations, that is, Arabic symbols, as visible in so-called 'number line estimation tasks'. While they show the logarithmic compression of magnitude representation when asked to place a certain numeron on a number line (Booth & Siegler, 2006; Siegler & Opfer, 2003), it has also been found that during the course of first or second grade, this representation is calibrated into an increasingly linear one (Booth & Siegler, 2006). However, multiple representations depending on context are possible. The logarithmic-to-linear development at first only applies to numbers up to 100. For higher numbers the representation is still logarithmic in second grade (Booth & Siegler, 2006). A similar situation could also be found in adults – they had multiple representations according to numerical context, too – only for higher number ranges (Siegler & Opfer, 2003).

To bring the knowledge about formal calculation procedures as well as preverbal quantitative abilities together in a comprehensive model of numerical understanding, Dehaene (1992; see also Dehaene, Piazza, Pinel, & Cohen, 2003) postulated his Triple Code Model (TCM), which still can be considered the most influential account in the literature (Klein, Moeller, & Willmes, 2013).

2.1 The Triple Code Model (TCM)

Dehaene (1992, Dehaene et al., 2003) merged the above reported results to his account of adult numerical processing. Later on, his model has been found to also account for typical observations made in children who have just started formal education, respectively already know some things about number symbols (Kolkman, Kroesbergen, & Lesemann, 2013). The model is built on two premises: That numerosities can be represented in three different codes, and that each numerical process is specifically bound to one of these three representational codes. Both premises will be outlined in further detail in the following.

(1) Numerosities can be represented in three different codes.

There are two symbolic codes, the *auditory verbal word frame*, which depends on and is manipulated in general-purpose language modules. The second symbolic code is the *visual Arabic number form*. This code (mainly) serves formal calculation processes and is manipulated for example by means of the algorithms.

The third, nonsymbolic code is the *analogue magnitude code* which later was described as the “coarse, intuitive sense of magnitude” (Ganor-Stern, 2015), or, as Dehaene (1992,

p. 30) puts it, “variable distributions of activation over an oriented number line obeying Weber-Fechner´s law”. In later publications, for this kind of processing Dehaene and others also coined the term 'number sense' (Dehaene, 1999; Dowker, 2012; Lyons & Beilock, 2011; Mazzocco, Feigenson, & Halberda, 2011; Park & Brannon, 2014; Wilson, Dehaene, Dubois, & Fayol, 2014) or the 'ANS', the 'approximate number system' (see for example Butterworth, 2011; Bonny & Lourenco, 2013; Cantlon, Platt, & Brannon, 2009; de Smedt, Noël, Gilmore, & Ansari; 2013; Ganor-Stern, 2015; Piazza, 2010; Lyons, Ansari, & Beilock, 2012).

Each of the three codes (see Figure 2) provides the possibility to transform input and output according to its respective needs. For example, a reading procedure can transform written Arabic numbers into the internal Arabic code. On the other hand, a writing procedure can transform our internal Arabic code by means of a motor program into written number symbols.

There are bidirectional translating paths between all three codes, meaning every representation can be transformed from a given code to each of the others.

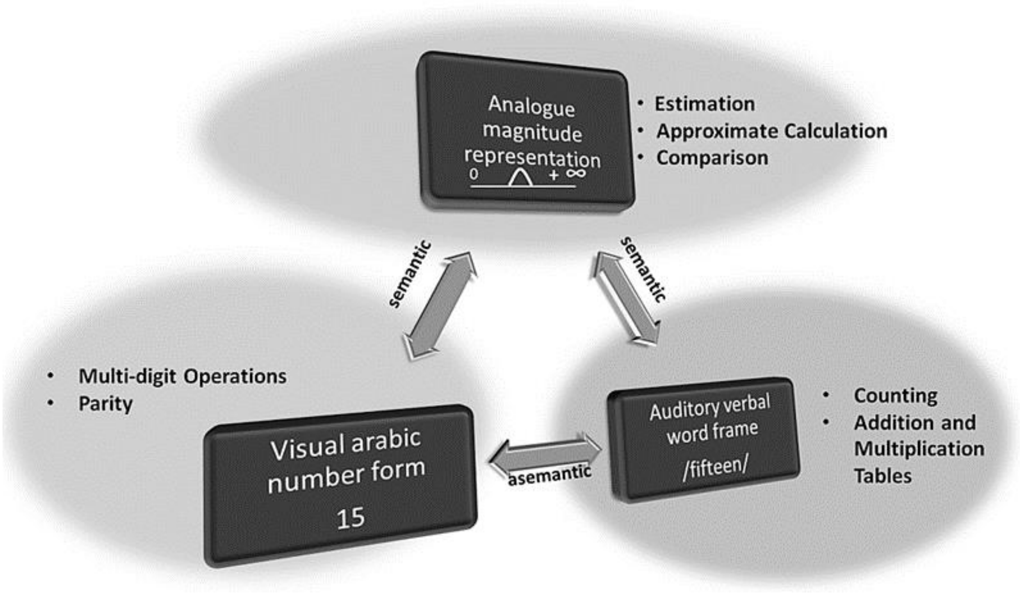

Figure 2. The Triple Code Model as proposed by Dehaene (1992)

However, only the translation into the analogue magnitude code is regarded as containing semantic understanding of numerosities (Crollen et al., 2011). When transferring

Arabic numerals or number words into the analogue code, the symbolic input is first approximated (e.g. in terms of the highest power of ten) to activate the according segment of the number line. In the reverse case, through categorization of the activated number line segment, an approximately adequate numeron (for adults in western cultures probably an Arabic symbol or number word) can be assigned. The path from analogue to verbal code contains no information or understanding of how a proper German or English number word is constructed and can only lead to a coarse estimate, like “two hundred”, but will not lead to an answer like “two hundred and seven” (Dehaene, 1992). Thus, the translating routes from and to the analogue magnitude code can only work in an approximate manner. Translating between the two symbolic codes, on the other hand, involves no access of the position of the numerosity on the mental number line and is therefore considered to be an asemantic route (Crollen et al., 2011). Note that Dehaene left it open if the direct translation paths between the Arabic and analogue code coexist with the translation paths between the verbal and analogue code; or if one of the symbolic codes is preferred for 'communication' with the analogue code.

(2) Each numerical process is specifically bound to one of the three representational codes.

Dehaene positions himself explicitly against hypotheses proposing that it might depend on interindividual cognitive differences or differences in learning history, what representation code is preferred for specific mathematical processes (see for example McCloskey et al., 1991). He states that there are genuine constraints as to what process can be performed in which code.

Judgments about parity or multi-digit operations have to be performed in the Arabic number form. The ability to perform such calculations is fully dependent on a positional notation system of number, the development of which in turn also recruits linguistic abilities (Dehaene, 1999). However, as the Arabic notation system is solely dedicated to numbers and its mastery has been found to be dissociable from linguistic impairments in adults (Anderson, Damasio, & Damasio, 1990; Cohen, Verstichel, & Dehaene, 1997), it is considered as a with age increasingly distinct system.

Stored facts like addition- or multiplication tables, on the other hand, must be retrieved in a verbal representation code. This is also the case for counting. Dehaene states that these mathematical processes are parasites of the general language processing system and the understanding of the verbal numerical notation. By that they are performed on a level that is not exclusively dedicated to numbers. Reciting the counting sequence is not qualitatively different from reciting the alphabet (Dehaene, 1992).

Finally, all approximate processing will recruit the analogue magnitude code. This includes magnitude comparisons of any kind, approximation, and also approximate calculations. As already toddlers before language acquisition as well as animals share these fundamental abilities, this code seems to be a distinct preverbal system of arithmetic reasoning.

To sum it up, according to Dehaene´s Triple Code Model, numerical understanding cannot be seen as a unitary representation or understanding anymore, but instead as a “fractionated set of numerical abilities, among which faculties such as quantification, number transcoding, calculation or approximation may be isolated” (Dehaene, 1992, p. 34, see also Dowker, 2012). In the next section, the analogue magnitude code will be described in further detail as the terminology has changed over the recent years, and several interesting findings could be added to our knowledge about the characteristics of approximate processing.

2.2 The analogue magnitude code or the approximate number system (ANS)

While Dehaene himself in more recent publications uses the term 'number sense' for the analogue magnitude code, among the majority of researchers the term ANS (approximate number system) has been established. Descriptions of the ANS describe nothing different than what was reported for the analogue magnitude code. However, since the postulation of the TCM, knowledge about this system has been complemented and thus it shall be reported in a little more detail in the following.

As reported above, on the representational level of the ANS, nonsymbolic and symbolic quantities are represented in an imprecise manner (Knops, Viarouge, & Dehaene, 2009; Le Corre & Carey, 2007; Noël & Rouselle, 2011; Wagner & Johnson, 2011). The precision of the approximate representation decreases the larger the quantity gets. This leads to a characteristic signature of the ANS, the *ratio effect* (also *distance effect*, see Brannon, 2002). It means that the logarithmic representation of magnitudes obeys Weber-Fechner´s law (see Dehaene, 1992): For an individual to be able to discriminate two quantities, not a specific *absolute* difference but a certain minimal *ratio* of both magnitudes cannot be undercut. So a person might be able to discriminate the quantity 3 from the quantity 5 (ratio 3:5 = 0.6) with the absolute difference of 5 – 3 = 2, but cannot discriminate the magnitude 100 from the magnitude 102 (or may be considerably slower in it) although both numbers differ by the same absolute value (= 2). The same ratio of 0.6 which would be realized in the comparison of 100 to 166.67, on the other hand, would be equally easy for the individual to discriminate as 3 versus 5. With age and experience, the largest discriminable ratio increases and approaches 1. While neonates are able to discriminate only a ratio not exceeding 1:2 (0.5), adults can reliably discriminate at a ratio of 7:8 (0.875) (Bonny & Lourenco, 2013; Piazza, 2010; Piazza, Facoetti, Trussardi, Berteletti, Conte, & Lucangeli, 2010). This effect is thought to occur because the mental representations of numerical magnitudes are coarse so that magnitudes that are close to each other overlap. This makes it harder to discriminate two similar magnitudes than magnitudes with a larger distance to each other (Izard & Dehaene, 2008; Xenidou-Dervou, van Lieshout, & van der Schoot, 2014). Or, put differently, the variance of the estimates of a given numerosity stays constant. So, in higher segments of the number line (where the quantities due to the logarithmic scaling are closer to each other), the same absolute difference results in greater overlap and is thus harder to discriminate.

The ANS enables us not only to discriminate between two quantities (e.g. dot arrays, see e.g. Barth, Baron, Spelke, & Carey, 2009; number of tones, see Bijeljac-Babic et al., 1993; area, see Odic, Libertus, Feigenson, & Halberda, 2013) and to recognize the larger one (van vanMarle, Chu, Li, and Geary, 2014). By recruiting the ANS, kindergartners can also perform simple arithmetic tasks like mentally adding two quantities and comparing the result to a third (Gilmore, McCarthy, & Spelke, 2010), or solve simple multiplication or division problems (Barth, Baron, Spelke, & Carey, 2009). They can do that with nonsymbolic representations of number (Barth, La Mont, Lipton, & Spelke, 2005) as well as with number symbols (Gilmore, McCarthy, & Spelke, 2007). This processing runs by the term 'approximate arithmetic' and seems to differ in several aspects from exact mathematical problem solving and number representation. In accordance with the TCM, several studies found exact arithmetic to be connected to language-specific neuronal circuits like the left inferior frontal circuit which is also activated when making associations between words (Dehaene, 1999; Spelke, 2011; Walsh, 2003). Approximate arithmetic, on the other hand, seems to be language-independent and is connected to a bilateral parietal visuo-spatial network (Dehaene, 1999), especially the mid intraparietal sulcus (Piazza, 2010).

3 Relationships between the ANS and exact skills – Behavioral evidence

There is an ongoing debate if the ANS can be considered as the fundamental building block of the uniquely human (de Smedt et al., 2013) capacity of exact number processing (Piazza, 2010; Spelke, 2011). Although there are a lot of interesting theoretical accounts to the question if the characteristics of the ANS or of specialized subsystems like subitizing can explain and lead to the adult representation of integers (see for example Butterworth, 2011; or Rips, Bloomfield, & Asmuth, 2008), the discussion of these accounts would go beyond the scope of this thesis.

On the other hand, researchers try to approach the subject by collecting evidence in neurophysiological as well as behavioral studies. In both fields, remarkable connections have been found between performances in tasks that tap the ANS and exact arithmetic skills or standardized math tests assessing formal math performance.

In this chapter, we will give an impression of the numerous literature and research exploring correlative and predictive relationships between different ANS measures and formal math performance. We will demonstrate that a consistent picture of these relations has not been drawn yet and that this might in part be due to an imbalance in the tested ANS measures, namely a frequently occurring confounding of demanded accuracy and format of the tasks. More specifically, we will point to the complete omission of symbolic approximate arithmetic tasks in current research. We argue that this measure is promising not only as a predictor of formal math performance, but also in terms of an intervention or training for subsequent exact mathematical processing. In this regard, we will discuss the possibility of fostering not only simple arithmetic performance, but rather the procedural and maybe even conceptual knowledge of more advanced arithmetic principles.

As mentioned above, in the majority of research taxing connections between the ANS and exact math skills, we observed the tendency to confound representational code (nonsymbolic and symbolic format) with the demanded accuracy. Authors make their conclusions about detected or not detected relationships or effects of ANS and formal (meaning exact and symbolic) mathematical skills, but often they actually only tested nonsymbolic ANS measures. This might in many cases be based on a more or less silent assumption that Arabic number symbols cannot recruit the ANS. And in fact, there are doubts on Dehaene´s (1992; Dehaene et al., 2003; see also Gilmore, Attridge, de Smedt, & Inglis, 2014; Hyde, Khanum, & Spelke, 2014; Knops et al., 2009; Piazza, 2010; Walsh, 2003) implication of the TCM, that approximation of any kind is performed in the analogue magnitude code (and hence on the representational level of the ANS). Amongst these supporters of the position that nonsymbolic and symbolic approximation are two distinct abilities (de Smedt et al., 2013; Ebersbach, Luwel, & Verschaffel, 2013, Kolkmann et al., 2013; Moeller, Klein, Nuerk, & Cohen Kadosh, 2012; Rips et al., 2008; Xenidou-Dervou et al., 2013), there are several accounts tackling the genesis of these abilities. Some state that they are distinct at first, but become integrated only in the course of formal education (Kolkmann et al., 2013). Others, on the contrary, argue that at first connected representations of symbols and analogue magnitudes become separated in a formal learning context. So after some years of formal schooling, number symbols have no connection to the corresponding actual quantity representations anymore. This might be reflected in the transition from logarithmic to linear magnitude representation that was observed during the first two grades of elementary school (Lyons et al., 2012).

Still, there is strong cumulative evidence for a tight connection between nonsymbolic and symbolic approximate skills in children as well as in adults. We thus believe in the original TCM proposition that both formats[2] are being processed in the analogue magnitude representation when approximation is demanded (Dehaene, 1992; Dehaene et al., 2003). Piazza (2010) reviewed empirical evidence in favor of this position.

For example, it has been found that traces of ANS-constraints like the ratio effect are also existent in adults´ performance in symbolic comparison tasks (Moyer & Landauer, 1967). Neuroimaging techniques revealed format independent activation when only approximate processing was involved as in comparison tasks (Notebaert, Nelis, & Reynvoet, 2011), and transfer of quantity-related responses of the mid-intraparietal cortex from nonsymbolic to symbolic format (Piazza, Pinel, LeBihan, & Dehaene, 2007). Furthermore, Piazza (2010) argues that the characteristic initial steps of symbolic num-

[2] When we speak of two formats (instead of three codes) we refer to nonsymbolically presented problems on the one hand, and Arabic symbolic notation on the other hand.

ber acquisition become understandable in terms of their dependence from the ANS. So, it has been observed that children between two and four years of age successively learn to match the numbers 'one' to 'four' to the corresponding cardinalities, each new number taking up to six months. This decelerated development can be explained by the above mentioned ratio effect. While in the first years in life, the ANS represents small numbers with high precision, the exact representation of a given set size N is first possible when the child has learnt to discriminate that set from the adjacent sets N-1 and N+1. Thus, an exact representation of the numerosity "three" is possible when children can successfully discriminate at a 3:4 ratio which is the case after the third year of life – the time at which children usually acquire the concept of the numeral "three" (LeCorre & Carey, 2007, but see Turr, Buri, & Anobile, 2010). Additionally, it has been found that it is difficult to teach children new number words as long as the corresponding quantities cannot be discriminated by their ANS (Huang, Spelke, & Snedeker, 2010). Finally, a third line of research reports correlations between impaired ANS processing and developmental dyscalculia as well as decreased mid-parietal activation in dyscalculic children when approximating as well as when calculating with Arabic numerals (Mussolin, Mejias, & Noël, 2010).

3.1 Correlative relationships and predictive power of the ANS for exact mathematical skills

As outlined above, a large proportion of current research advances the position that the ANS lays the ground for later exact mathematical skills (Dehaene, 1992). Given the importance of mathematical proficiency in modern society, there have been attempts to identify early indicators and risk factors of later math performance. Consequently, these attempts also tax competencies associated with the ANS: One line of research focuses on the question if ANS measures correlate with later skills in exact mathematics as they are demanded in school and professional contexts, or even predict them. As will be outlined in the following, these results do not reveal a consistent picture yet. In part, this might be due to the fact that different measures of ANS performance can be (and are!) applied. Most frequent is the use of the already introduced comparison tasks. However, there is also a growing body of research that acknowledges the possibility for very young children to perform *arithmetic* in an approximate manner (Barth et al., 2005; Gilmore et al., 2007; 2010; Gilmore et al., 2014; Pinheiro-Chagas et al., 2014). Thus, in the following, a selection of the most interesting findings is presented sorted by the respective measure that was used to determine ANS acuity; that is, comparison vs. approximate arithmetic tasks. Following the above mentioned question if also symbolic representations can recruit processes on the level of the ANS, special attention will be paid to the formats (nonsymbolic vs. Arabic or verbal numerals) of the administered ANS measures.

Comparison Tasks. In their literature review, de Smedt et al. (2013) found that robust evidence seems to exist for a positive relationship between performances in symbolic comparison tasks and mathematical skills. The results for an analogue relation between nonsymbolic comparison and math performance however were more mixed. This is also mirrored in the selection of recent research we will present in the following.

For example, Libertus, Feigenson, and Halberda (2013) found a nonsymbolic comparison measure of ANS acuity to predict significant unique proportions of variance in math performance half a year later in preschoolers. Similarly, Lyons and Beilock (2011) found a correlative relationship between a nonsymbolic comparison task and exact arithmetic performance in a cross-sectional study with adults. However, they also found that this relationship was completely mediated by participants´ symbolic numerical ordering ability (see also Chapter 5). More recently, Lyons et al. (2014) tested correlations and predictive power of both symbolic and nonsymbolic comparison tasks for math performance in the children of the grades 1-6. They found only zero-order relationships between nonsymbolic comparison and math performance, but at no age level a significant cross-sectional regression. Symbolic comparison, on the other hand, could in fact predict simultaneously assessed math performance in the first 5 years of school. Contrary to the cross-sectional studies reported above, Attout, Noël, and Majerus (2014) longitudinally tested children in kindergarten, in first and in second grade with a symbolic comparison task. Contrary to Lyons et al. (2014), they found no correlations at a given testing point between acuity in the symbolic comparison task and calculation ability. But they extended the results by demonstrating a genuine (longitudinally measured) predictive relationship: Performance in the symbolic comparison task in kindergarten could predict variance in calculation performance in first and second grade.

Thus, the reported studies suggest that performance in a nonsymbolic comparison task can successfully longitudinally predict formal math performance in preschoolers (Libertus et al., 2013) but not anymore in first to sixth graders (Lyons et al., 2014: correlative rather than predictive relationship) and adults (full mediation found by Lyons & Beilock, 2011). However, the latter regressions were cross-sectional ones.

As de Smedt et al. (2013) concluded in their review, symbolic comparison performance on the other hand was a stable longitudinal as well as cross-sectional predictor in preschool as well as during the first five years of school (Attout et al., 2014; Lyons et al., 2014).

Approximate arithmetic tasks. Barth et al. (2005) provided evidence that preschoolers can perform simple arithmetic as long as they only have to do it on the representational level of the ANS. This is the case when no exact result needs to be calculated and the expected decision can be made based on approximation processes. While Barth et al. (2005) still stated that preschoolers only could perform simple and approximate arithmetic operations when these were presented in a *nonsymbolic* format, later research revealed that this is not true. Rather, the results of Barth and colleagues were based on symbolic and nonsymbolic approximate arithmetic tasks that differed in working memory demands. They had presented symbolic problems only verbally to the children, while the nonsymbolic problems were presented as images. When in a later study of Gilmore et al. (2007) this artefact was eliminated, preschoolers showed better-than-chance performance also in *symbolic* approximate arithmetic tasks (see also Barth et al., 2009; Gilmore & Spelke, 2008). Gilmore and Spelke (2008) assume that children might understand logical principles of exact arithmetic at first in form of approximate relationships and operations before they map this understanding onto exact representations. This view is supported in their later finding that preschool performance in a nonsymbolic approximate arithmetic task could predict formal math performance at the end of first grade (Gilmore et al., 2010).

Especially the finding that children can perform symbolic arithmetic in an approximate manner already before starting school has drawn considerable attention in the scientific community. However, these findings have led to surprisingly little effort to further determine the relationship between approximate arithmetic and formal math skills (but see the work of Pinheiro-Chagas et al., 2014, reported below). Also, it has to be noted that there seem to be almost no studies that test approximate *arithmetic* performance and its relationship to exact mathematical skills, despite the finding of a domain specific predictive relationship between children's nonsymbolic approximate arithmetic performance and their math performance at the end of first grade (Gilmore et al., 2010; see also Mazzocco et al., 2011).

A shortcoming in Gilmore et al.´s study (2010) is that they did not include a symbolic approximate arithmetic task; even though the same research group had shortly before found out that preschoolers actually are able to perform also symbolic arithmetic approximately (Gilmore, McCarthy & Spelke, 2007).

The omission of the symbolic approximate arithmetic task is not trivial, given the discussion about the question if the ANS is format-independent and the inconsistent findings regarding nonsymbolic comparison tasks. We believe that approximate arithmetic both in a nonsymbolic as well as symbolic representational code relies on the same (analogue) magnitude representations and arithmetic procedures (Dehaene, 1992; Dehaene et al., 2003; Knops et al., 2009). Thus, also aside from Gilmore et al.´s (2010) promising findings, we expect that also symbolic approximate arithmetic should predict later math skills: In approximate arithmetic tasks, not only pure ANS acuity (which already for itself seems to be predictive) accounts for variance in children´s results (Gilmore et al., 2014), but also their understanding of the arithmetic operations (typically addition) is assessed. The latter is the ability needed and fostered in formal math education and should thus further contribute to a positive correlation between ANS measures and math achievement.

First results in this direction stem from Pinheiro-Chagas et al. (2014). They systematically compared interrelations between three (all nonsymbolic) ANS measures: nonsymbolic comparison, nonsymbolic approximation (here meaning: assigning a number word to a dot array), and nonsymbolic approximate addition. Furthermore, they computed different regression models and mediation analyses to explore which of the measures would predict exact calculation performance in children from grade 1 to grade 6, and if the found relationships were direct or rather mediated from one ANS measure by another. They found that all three ANS measures correlated with each of the others and also with the exact calculation performance. But the regression analyses disclosed that only the nonsymbolic approximation and approximate addition performance explained unique proportions of variance in exact calculation, while comparison did not. Mediation analyses specified that the relationship between nonsymbolic comparison and exact mathematical skills (which is the most frequently tested connection in current research) was partly mediated by the approximation performance, but, more crucially, *fully* mediated by the approximate addition task. The authors hence propose that their results have to be attributed to differential cognitive processes that are demanded in the three ANS-tasks. While a comparison task consists of simple discrimination (but note that also an ordinal aspect could be included, as discussed in Chapter 5), a nonsymbolic approximation task contains the mapping of an internal represented analogue magnitude to number words, and approximate arithmetic recruits rather complex manipulations of numerosities. Pinheiro-Chagas et al. (2014) conclude from their results that the latter is the critical mechanism driving the association between approximate, analogue magnitude processing and exact mathematics.

If they are right, paradoxically, the more direct measure for the connection between ANS and formal math is also the one administered least often in longitudinal studies; and even when it is, it is only tested in a nonsymbolic format. If the ability to manipulate internal quantities is the actual interface between approximate and exact numerical abilities, a direct comparison of these abilities in both a symbolic as well as nonsymbolic code is needed. However, given the finding of Lyons and Beilock (2011) that numerical ordering ability in turn mediated the positive relationship between nonsymbolic comparison and exact arithmetic; it might also be that not approximate arithmetic per se, but the more basic ordinal knowledge accounts for the connection between nonsymbolic arithmetic and exact math skills that Pinheiro-Chagas et al. (2014) found in their sample. We do not think that is likely, however, because in their study, nonsymbolic comparison performance which in itself contained an ordinal judgment (see Chapter 4) did not explain a significant proportion of variance. To our understanding, this advances the hypothesis that the more specific aspect of understanding arithmetic and mentally manipulating quantities is responsible for the relationship rather than ordinal knowledge.

Besides the academic void, there is another, more practical reason to start focusing on approximate arithmetic tasks in both representational formats. That is, a potentially fostering effect of ANS-related training on later exact number representation or even exact arithmetic abilities. The findings that several ANS measures seem unaffected by socio-economic factors (see Chapter 4) and the connections between ANS and formal math (see discussion above) make it a promising outlook to use this connection for intervention. Some of the according research will be topic in the next section 3.2.

3.2 Beneficial effects of ANS-related interventions on exact mathematical skills

The repeated – although not very consistent – findings of predictive relationships between the ANS and formal mathematics have so far led to only few attempts to not only predict, but actively foster children´s math skills by means of ANS-related training tasks.

Interventions taxing general arithmetic skills. Wilson et al. (2007) did find positive effects of their computer game "Number Race" in low-SES kindergartners. The game included nonsymbolic, symbolic and verbal comparison and approximate addition tasks. Playing the game benefitted symbolic approximate measures, but not nonsymbolic approximate or (symbolic or nonsymbolic) exact ones measured after a 14 week training period. On the other hand, Obersteiner, Reiss, and Ufer (2013) compared effects of the "Number Race" with their own computer-training in a mixed-SES sample. For that, they built an analogue game, but this time incorporating the tasks demanding exact rather than approximate processing. They found that the approximate original benefited *all* subsequently assessed approximate measures, no matter if symbolic or nonsymbolic, if comparison or arithmetic; but not the exact ones. The exact version of the game on the contrary only did enhance exact skills like subitizing. To sum it up, both studies that tested the original "Number Race" did not find a positive effect of the various approximation tasks included in the game on later exact mathematical performance. Note, however, that in the "Number Race", comparison and arithmetic tasks were mixed and their specific effects could not be disentangled. Contrary, there are some studies that test the suitability of ANS-based interventions in a more controlled and local context.

For example, Park and Brannon (2013) trained adults with nonsymbolic approximate addition and subtraction tasks in ten sessions. Subsequently, they found an improvement in participants´ exact symbolic addition and subtraction performance. In a second experiment, they compared the fostering impact of solving the nonsymbolic approximate arithmetic tasks to that from solving a training task focusing on numerical ordering, which Lyons and Beilock (2011) had found to be an even stronger predictor of symbolic arithmetic in adults. However, Park and Brannon found the approximate nonsymbolic arithmetic tasks to be more effective in boosting the exact symbolic arithmetic performance.

Hyde et al. (2014) on the other hand compared the impact of two kinds of comparisons and two kinds of approximate arithmetic tasks to test the specificity of the relationship between ANS and symbolic addition. They compared possible positive effects of different training tasks directly before assessing exact symbolic addition skills in first graders. According to their specific condition, as a training children worked through (1) a nonsymbolic approximate addition task, (2) a line length addition task, (3) a nonsymbolic comparison task, or (4) a brightness comparison task. Hyde and colleagues found that only the training tasks with numerical content ((1) and (3)) had a positive influence on subsequent exact arithmetic. By employing the other training tasks, the authors could rule out possible alternative hypotheses that would explain the fostering effects differently than by a direct connection between ANS and symbolic arithmetic. The non-numerical tasks (2) and (4) did not lead to comparable benefits as the numerical ones. This contradicts the assumption of a domain-general magnitude system (Walsh, 2003) for lengths, brightness, magnitude, numerosity etc. That the results are due to common cognitive operations beyond the numerical domain can be ruled out by the fact that exact arithmetic profited only from the approximate addition task and the nonsymbolic comparison task. The tasks (2) and (4) demanded analogue abilities in a non-numerical context, but had no positive impact on exact arithmetic. Furthermore, the possibility of "easier arithmetic" warming up the more difficult one was refuted by the finding that nonsymbolic comparison worked equally well as approximate addition.

Interventions taxing principle use and understanding. Last but not least, another training study should be mentioned here. This study, conducted by Sherman and Bisanz (2009), does strictly speaking not explicitly test the effect of approximation. Instead, the researchers investigated the effect of solving equivalence problems with concrete (thus nonsymbolic) material on the subsequent understanding of equivalence problems in exact symbolic format. They first instructed second graders to solve the concrete equivalence problems and afterwards administered a symbolic equivalence task. In a control condition, participants received the reverse order. They found that solving the concrete problems at first facilitated the performance in symbolic problems, whereas the reverse was not the case.

This study is reported to call attention to another shortcoming of current research: So far, ANS-based interventions have only been tested regarding their impact on simple arithmetic, but it has not been actively tried to extend their beneficial influence to less basic arithmetic *principles*. Sherman and Bisanz (2009) demonstrated that by addressing informal (though exact) mathematical skills with concrete operands, a principle like the law of equivalence could be induced and enabled second graders to transfer this knowledge to a formal, symbolic and exact context.

Up to date, there is no study that explored the suitability of approximation to induce understanding beyond number representation or merely procedural competencies like performing simple arithmetic[3]. The principles effective in (approximate) addition problems are still rather general. For example, a basic principle like "addition makes more" can already be understood by 5-months old infants. Wynn (1992) found that children of this age already look longer at 'impossible outcomes' (two dolls disappear behind an occluder, and when the occluder is removed there is only one doll left) than at possible ones. However, there are also more abstract principles that children can also understand on the representational level of the ANS already before starting school. So, it has been demonstrated that children are able to understand the inversion principle ($a + b - b = a$, Gilmore & Spelke, 2008) on an approximate level, or even the additive law of commutativity (Cowan, 2003). Children can experience its core principle, which is order-irrelevance, in many non-numerical as well as numerical everyday situations long before entering school. For instance, a child may learn that the order is irrelevant when laying the table or when putting on one's socks. By contrast, putting on underpants and trousers clearly does require a strict order. Consequently, already toddlers might know that order is irrelevant in some situations and relevant in others. If order is irrelevant, they also learn that combining two different sets of objects leads to the same result regardless of the order (e.g. Canobi, Reeve, & Pattinson, 2002; Cowan & Renton, 1996; Gallistel & Gelman, 2000, Resnick, 1992; Sophian, Harley, & Manos Martin, 1995;). Another indicator of understanding additive commutativity is the use of the COL strategy (see Chapter 2, Baroody & Gannon, 1984; Baroody & Ginsburg, 1983; Canobi et al.,

[3] This topic is explored in our Experiments 3 and 4. These experiments, including parts of the theoretical background given for these experiments, have been prepublished with the permission of the Dean of Research of the University Cologne and can also in greater detail be found here:
Hansen, S. M., Haider, H., Eichler, A., Godau, C., Frensch, P. A., & Gaschler, R. (2015). Fostering formal commutativity knowledge with approximate arithmetic. *PloS ONE,* DOI:10.1371/journal.pone. 014255.

2002; Gallistel & Gelman, 1992; Siegler & Jenkins, 1989). Hence, there is good evidence that even preschoolers already have precursory knowledge about commutativity. This is in line with data reported by Dowker (2009; 2014) who compared the derived fact strategy use in 6–7 year-olds. Derived fact strategies refer to the ability to extract new arithmetic facts from known facts on the basis of arithmetic principles like commutativity, associativity or the inversion principle (Dowker, 2012). In Dowker's study, children had to solve addition and subtraction problems slightly too difficult for them to compute, on the basis of a previously given result of a related problem. The relationship between the two problems consisted of a specific arithmetic principle. Among several principles, Dowker found commutativity to be the one used the second-most. Only the basic identity principle (understanding that the exact repetition of an arithmetic problem will result in the same total) was even easier for the participants.

The exploration of less basic principles (like for instance commutativity) holds an important advantage: It allows the differentiation between procedural and conceptual knowledge. Procedural knowledge is defined as 'knowing how', the ability to apply a particular strategy in a specific problem context (Hiebert & LeFevre, 1986). Conceptual knowledge on the other hand refers to 'knowing why', the understanding of the abstract principle that underlies the constraints and boundary conditions for using a specific procedure. These definitions indicate that the correct application of a procedure or strategy in a mathematical problem only indicates procedural knowledge, but does not necessarily imply that the individual also has conceptual understanding. It is not certain that the problem solver actually knows why and under which conditions the certain procedure works (Hatano, 1988; Hiebert & Wearne, 1996; LeFevre et al., 2006; Rittle-Johnson, Siegler, & Alibali, 2001; Schneider, Rittle-Johnson, & Star, 2011). As Baroody (2003, p.27) puts it, "computational efficiency can be achieved without understanding". Abstract principles like commutativity allow the use of time-saving procedures in problem solving: If you understand that 2 + 5 results in the same total as 5 + 2, you do not have to calculate anew when you encounter a problem that is commuted to one you have solved shortly before, thus saving time. These time-savings can be measured and thereby uncover strategy use. Conceptual knowledge on the other hand has to be measured separately to avoid inducing certain answers with one´s questioning that would not be shown spontaneously by the learners.

To sum up, there is reason to believe that nonsymbolic approximate tasks, both comparison and arithmetic, can benefit exact arithmetic performance in adults as well as in first graders, on the short as well as on the long term. But again, as it also is the case in the studies targeting to predict formal math performance by measures of ANS acuity, also among the intervention research so far there is to our knowledge no study explicitly testing symbolic approximate arithmetic regarding its potential of boosting exact mathematical skills. Positive effects have been stated for both nonsymbolic comparison but also nonsymbolic approximate arithmetic (Park & Brannon, 2013; Hyde et al., 2014), as well as for nonsymbolic exact arithmetic (Sherman & Bisanz, 2009).

This lack of using symbolic approximate arithmetic as an induction is surprising. The idea suggests itself that transfer from the representational level of the ANS to formal mathematics, and hence an exact and symbolic, level might be especially facilitated by addressing the ANS also in a symbolic format. Speaking in terms of the Triple Code Model (Dehaene, 1992; Dehaene et al., 2003), the required mapping process from the Arabic code to an internal analogue magnitude representation might pave the way for abstracting intuitively understandable regularities from ANS back to Arabic code and hence to a formal learning context. In line with our knowledge about children´s precursory concepts of rather abstract mathematical principles and first attempts to strengthen their use and understanding of such domain-inherent regularities (e.g., Sherman & Bisanz, 2009), it might even be possible to refer to this early stages of principle understanding (Resnick, 1992) by means of approximation.

With our Experiments 3 and 4 we thus aim to not only test this efficacy of approximate arithmetic tasks as a training for exact numerical skills, but also try to disentangle if a possible beneficial effect is restrained to procedural knowledge, conceptual knowledge, or might even foster both kinds of knowledge or their integration to an increasingly abstract concept.

4 Influence of the SES on mathematical competencies

From grade 1 to adult age, individuals from low-income families or families of linguistic or ethnic minorities have been found to perform worse in several numerical tasks than their peers of higher socioeconomic status (SES) (Starkey & Klein, 2008). SES has already been found to correlate and to predict many childhood performance measures, like language ability and even executive functioning (Calvo & Bialystok, 2014; Huttenlocher et al., 2010; Noble, Norman, & Farah, 2005). However, in the light of recent large-scale assessments of learning processes like TIMMS (Bos, Wendt, Köller, & Selter, 2012), little has received as much attention like the tight connection between children´s math performance and SES.

In the following we will describe the most frequent account to explain the connection between SES and math performance, that is, impaired symbolic number knowledge in low-SES children due to differences in math related content in their home environments. We will also report empirical findings regarding the question which numerical competencies are affected by socioeconomic factors and which are not. It will be shown that these findings are not consistent and leave room for an alternative route on which home environment might gain impact on children´s math performance (alternative to symbolic number knowledge). This alternative route might consist in a mediating role of language abilities which are known to be related both to SES as well as to exact numerical skills. It will be outlined what kind of study is missing to gain further insight if the alternative route might account for the connection of SES and math performance; and also the importance of an inclusion of symbolic approximate, but even more of non-symbolic exact problems is demonstrated.

4.1 Explanations for the relationship

About what channels might a child´s socioeconomic situation influence his/her math skills? The most logical explanation is that in low-SES children's home environment, there is less encouragement to engage in mathematical activities. This differing level of support children receive at home can consist in the degree of proacademic activities they encounter (Blevins-Knabe & Musun-Miller, 1996). That is, reading with their parents, talking about numbers and quantities in everyday activities and games as well as playing board games. Starkey and Klein (2008) found middle-class mothers to engage in more (and also more complex) mathematical activities with their children than working class mothers. For example, the relative frequency of including mathematical activities in daily routines as well as making up math-games was higher in the middle-class group (Starkey & Klein, 2008, p. 263). While findings that also math-related books, board-games, and math software are more seldom in lower SES-households could be attributed to differing financial resources of the families, the absence of math related activities as described above cannot be solely explained with income (or the lack thereof). Starkey and Klein's (2008) assumption is that according to their own education and classes, the parents from middle- to high-SES backgrounds put a bigger value on numerical activities than their low SES counterparts.

Jordan et al. (2008, see also Jordan & Levine, 2009) specified the kind of impairments that low-SES children have in the numerical domain. For example, they showed that low income children started using their fingers in counting about one year later (i.e., in first grade) and for a prolonged period of time compared to high income children. It has to be noted that in kindergarten, the use of finger counting is positively correlated to accuracy in number combinations, while in second grade this correlation already has become a negative one. They concluded that the development of low-SES children´s number competence was decelerated (instead of qualitatively different) and that this is most likely due to the fact that children from low-SES families have less knowledge of number words and symbols. This, in turn, should be based on home-environmental factors described above, which they then receive when starting school.

If impaired symbolic number knowledge is the main factor in differences of numerical processing between children of different socioeconomic backgrounds, one should expect to mainly find differences in tasks that are presented symbolically. Crucially, this should also apply to approximate operations because the impeded processing of the number symbols would in turn at least slow down the respective approximation task.

A somewhat different explanation for the origin of the relationship between numerical abilities and SES consists in a more indirect path. The assumption is based on findings that not all neurocognitive systems seem to be influenced to an equal degree by socio-economical factors. For instance, Noble, Norman, and Farah (2005) compared low and middle SES children in their performance of visual cognition, spatial cognition, memory, language abilities and executive functions. They found the strongest differences between the SES groups in tasks taxing language and executive functions. They also provided hints that the latter only were correlated with SES because language mediated between SES and executive functions. As outlined above, linguistic competencies also play a big role in the culturally acquired numerical competencies. Given the strong relationship between linguistic factors and symbolic and exact mathematical processing (see Dehaene, 1992; Dehaene et al., 2003), language skills might very well be a mediating factor between SES and math performance, too. Fernald, Marchman, and Weisleder (2013) found that already at the age of 18 months, children showed significant differences in vocabulary and language processing efficiency according to their SES. In 2 year-olds, low-SES children lacked 6 months behind their high-SES peers in language processing efficiency. Thus, there is the possibility that also the differences in math performance according to SES might actually go back to linguistic disadvantages of children with lower SES. This should be reflected in a stronger influence of SES on every kind of *exact* processing. Also when the problem is presented nonsymbolically, the operation would nevertheless have to be carried out in one of the two symbolic codes. These have been found to rely heavily on the human language faculty (Dehaene, 1992; Dehaene et al., 2003; Starkey & Klein, 2008; Spelke, 2011; Walsh, 2003).

4.2 Empirical Evidence

For the whole course of formal education and even for adult life, a strong relationship between one´s socioeconomic background and individual performance in various measures of numerical abilities has been established. Even in kindergarten age, children from high SES families outperform children from low SES backgrounds in formal as well as informal mathematical knowledge (Jordan, Kaplan, Locuniak, & Ramineni, 2007; Starkey & Klein, 2008). Contrary to the questions discussed above (prediction, intervention), the use of arithmetic tasks in this research domain is not so seldom. However, also in this field, a systematic realization of the possible combinations of format and demanded accuracy has not yet been conducted. In the following, it will be outlined what is missing and why this matters.

Exact and Approximate Arithmetic Tasks. One very popular and recent study in the field is the already mentioned work of Gilmore et al. (2010). They found first graders from high SES background to perform significantly better in formal (exact symbolic) mathematic at the end of first grade than children with middle SES background. On the other hand, the same children did not differ in their nonsymbolic approximate addition skills. Gilmore et al. (2010) assumed due to their results that children might understand the logical structure of arithmetic at first on an approximate, nonsymbolic level, and that this early understanding on the representational level of the ANS is robust to environmental factors. Performance in symbolic approximate arithmetic was not compared between their two SES groups.

A reason to take into account explanations that differ from or respectively specify the conclusions of Gilmore et al. (2007; 2008; 2010) are previous results that had been obtained by Jordan, Huttenlocher, and Cohen Levine (1992). On the one hand, they found significant differences between 5-6 year old children from middle and low-income families in the performances of different kind of *verbal* problems that tested exact addition and subtraction. The tested formats were (1) story problems; (2) word problems that did not provide a story context but named object referents in the problem; and (3) number fact problems (“How much is 3 + 5?”). On the other hand, nonverbal (and in this regard also nonsymbolic) addition and subtraction problems did not elicit any differences between the SES groups. To make these tasks nonverbal, the experimenter slid two arrays of small disks (the 'addends') under an occluder right in front of the child. The child could at no time see the two arrays simultaneously and should then, on his/her own desk, reproduce the exact number of disks that the experimenter had under the occluder. Thus, here it seems that arithmetic problems do not have to be approximate to be robust against socioeconomic factors, but rather only nonsymbolic (see the previous chapters for a review of the differences and according implications).

It has to be noted that Jordan et al. (1992) appear to be the first who in their study included nonsymbolic arithmetic problems that demanded an *exact* solution. Their results can thus be a motivation to pursue the newer studies reported above. In these it was found that children´s performance in nonsymbolic approximate arithmetic problems were not affected by different SES, contrary to symbolic and exact problems (Gilmore et al., 2010). However, note that only symbolic exact problems were tested. The results summarized above show that it needs to be specified if there is *one* aspect that determines if environmental factors like SES can disturb the development of early numerical abilities, and if there is – which aspect is it? The mere *format of problem presentation*, or rather the *demanded accuracy* of an operation (in terms of the TCM: the representational code *in which the operation is performed*, which is not necessarily equal)? Or is it not possible to clearly distinguish between the two factors? To answer these questions, it is necessary to test the numerical ability in question in a symbolic as well as in a non-symbolic representation, both exact and approximate, resulting in a 2x2 combination design.

For exact calculation processes, Jordan et al. (1992) started to resolve the already mentioned frequent confounding of problem format (nonsymbolic vs. symbolic) and demanded accuracy (exact vs. approximate). As described above, their results already suggest that the (sometimes detrimental) influence of children´s SES might not apply to all exact processing. While Gilmore et al. (2010) found that nonsymbolic approximate arithmetic did not vary with SES while exact symbolic processing did; Jordan et al. (1992) have already specified that children of the same age group did not differ according to their SES in nonsymbolic exact arithmetic either.

If thus the assumption is true that mainly an impaired knowledge of numerical symbols (and not the representational *code of the operation*, which would be based on vs. not be based on language) accounts for the poor performances of low-SES participants, symbolic approximate arithmetic should be affected by differing SES-levels, too: A decelerated understanding of Arabic number symbols would in turn result in a decelerated transcoding between the symbolic code and the analogue magnitude representation, and hence also impair approximation processes if the problems are presented symbolically.

On the other hand, if impaired linguistic skills in low-SES children mediate the influence SES has on numerical processing, it should be expected that performance in both exact tasks (symbolic and nonsymbolic) should be worse in low-SES participants because these tasks have to be carried out in one of the two symbolic codes whose development is based on human language (Dehaene, 1992). Symbolic approximate arithmetic should probably not be influenced in this case, because the according process can be carried out in the language-independent analogue magnitude code. However, the latter is a speculative aspect of this hypothesis and should be treated with caution, because it may also be the case that the according representation of magnitudes is already impaired before being transcoded from the language based Arabic to the analogue code. If this is the case, only the nonsymbolic exact performance should be robust against SES influence.

In line with the account of little symbolic number knowledge as the cause of SES-effects, Wilson et al. (2007) found in their intervention study that the approximation tasks contained in the “Number Race” benefitted low-SES preschoolers’ symbolic comparison performance significantly, but not their performance in nonsymbolic comparison. From that they conclude that it is actually the mapping between symbolic and analogue magnitude code that is impaired because of little symbolic number knowledge in low-SES groups, not one of the representational codes per se. However, as discussed above, a mediating role of linguistic skills *might* produce the same result (influence on symbolic, but not nonsymbolic approximation). One would actually have to test nonsymbolic exact processing to disentangle both explanative accounts (mediating linguistic skills would lead to an impairment of nonsymbolic exact arithmetic in low-SES children, poor number knowledge would not).

A further insight to the question promises the study of McNeil, Fuhs, and Alibali (2011). They tested low- and high-SES preschoolers (4-6 year-old) and obtained puzzling results. Contrary to Gilmore et al. (2010), they found that *both* the performances in a nonsymbolic (!) as well as a symbolic approximate arithmetic task were influenced by children’s SES. They furthermore report that only high-SES children (despite their better general performance in both tasks) showed an effect of non-canonical vs. canonical approximation trials. Canonical trials are those which present a (nonsymbolic or symbolic) addition problem on the left side of the screen, and the to-be-compared reference quantity on the right side. Non-canonical trials accordingly presented the operation on the right and the reference quantity on the left side. McNeil et al. (2011) found that only the high-SES children showed better performance in the canonical approximate arithmetic trials than in the non-canonical ones. Low-SES children on the other hand performed similarly in both kinds of trials. The authors attribute this finding to the fact that the canonical trials better match the unidirectional view of addition ('combining the operands *leads to* the result' instead of an understanding of the *equivalence* between the two sides of the equation sign) (Baroody & Ginsburg, 1983). This view is reinforced by exposure to exact arithmetic problems. Because their participants had not started school yet, McNeil et al. (2011) assume that this exposure happens to a stronger degree in high-SES households which would explain the observed interaction of SES and kind of trial (canonical vs. non-canonical). They thus conclude that already before starting school, exact arithmetic knowledge shapes children´s ability to use their ANS for arithmetic reasoning. They hypothesize that in children already attending school, this interaction of canonical vs. non-canonical trials and SES could either be even more pronounced (because these children have been exposed to formal arithmetic to a stronger degree) or it might alternatively be annihilated. The latter might be the case because in the course of learning formal arithmetic, children not only encounter the canonical problem-format more often, but also the requirement to produce an exact result when encountering a canonical arithmetic situation. This might compromise the solution of approximation problems in canonical representation and thus level out its advance compared to non-canonical problems.

To sum it up, results so far show an inconsistent picture of the relationship between SES and different numerical abilities, especially nonsymbolic approximate arithmetic, whereas symbolic approximate arithmetic as well as nonsymbolic exact processing has rarely been tested.

However, some additional results in related domains have been obtained. Mejias and Schiltz (2013) tested the exact and approximate quantity representations of kindergartners coming from different SES in a symbolic as well as in a nonsymbolic format. They found that in second kindergarten grade (age 4-5 years), only exact quantity representation (non-symbolic as well as symbolic) was worse in children with low SES than in middle SES children. However, they further found that in the *third* year of kindergarten (5-6 years), the exact *as well as* the approximate number representation seemed to be reduced in low SES children compared to children from middle SES families. Additionally, with the exception of the nonsymbolic exact and the symbolic approximate measure, all tasks were significantly intercorrelated now, whereas it had only been the nonsymbolic tasks one year before.

On the one hand, these results suggest a different interpretation than those of Jordan et al. (1992). It seems that, at least in number representation as opposed to arithmetic, in second year kindergartners, the influence of SES differs less with the format (symbolic vs. nonsymbolic) but more with regard to the demanded accuracy (approximate vs. exact) of the presented task and hence the representational code in which it has to be performed. This indicates the possibility of a mediating role of language skills. It has to be noted that children of second kindergarten grade in which this result was obtained were one year younger than Jordan´s participants. In the sample of similar age, the detrimental influence of low socioeconomic status had spread from exact to approximate processing also in Mejias and Schiltz´s study (2013), in this point matching the results of McNeil et al. (2011).

All in all, one important difference in the results of Mejias and Schiltz (2013) and Jordan et al. (1992) is that Mejias and Schiltz found no differential influence of SES on symbolic vs. nonsymbolic exact problems. This at first seems to complicate answering the above raised question, if it is the format of of the numerosities rather than the representational code of the operation which determines a performance´s robustness against socioeconomic influences. However, Mejias and Schiltz did not study children´s arithmetic proficiency, but their general number representation – this might be an important difference because performing arithmetic includes several steps going beyond just representing a quantity (see also Pinheiro-Chagas et al., 2014). Furthermore, a problematic aspect in Mejias and Schiltz´s study is their symbolic approximation measure. The respective task was to determine the approximate numerosity of an *array of dots* with a number word. This is a mix-up of nonsymbolic and symbolic representation and hence cannot be considered purely symbolic. The problem is illustrated in the fact that later on, Pinheiro-Chagas et al. (2014) used the same task as their *nonsymbolic* approximation measure.

Thus, a study that compares children of different socioeconomic backgrounds according to their performances in exact and approximate arithmetic presented symbolically as well as nonsymbolically is still missing. This could provide further evidence for a possible mediating role of linguistic abilities in the relationship of SES and numerical skills. Following the suggestion of McNeil et al. (2011), this study should be carried out after the participants´ have started school, to test for a possible effect of non-canonical vs. canonical approximation trials.

Of special interest in such a study is the exact nonsymbolic task. It promises distinct results regarding an influence of SES, depending on the adequacy of one of the two explanations for the connection between SES and math performance outlined above. According to the – unfortunately rather inconsistent – results of the reported studies, two competing hypotheses are formulated.

If the knowledge of symbolic numbers, and hence the format the problems are presented in, takes more responsibility for the relationship between numerical abilities and SES, both nonsymbolic tasks (approximate as well as exact arithmetic) should not be influenced by low SES. The processing of symbolic problems, on the other hand, should be influenced in both approximate as well as exact operating.

If the influence of SES is mediated by linguistic skills, however, one should find that low-SES children are impaired in the Arabic and the verbal code, as both require the use of language modules. While we cannot reliably make a prediction if symbolic approximation might already suffer in low-SES participants because of its original representation in the language-based code, the distinct finding here should be lower performance in the low-SES group in the nonsymbolic exact task, as exact calculations cannot be performed on the level of the analogue magnitude system.

5 The role of numerical ordering ability for the development of exact processing

In the following chapter an additional aspect of the connection between ANS and formal math performance is described. In current research, it has been frequently proposed that ordinal knowledge might be of special importance in the relationship. We will present empirical findings and accounts of why understanding of ordinality might be a logic interface between the coarse representations of the ANS and an increasingly precise grasp of exact quantities. Furthermore, we will show, that the mediating relationships between ANS, numerical ordering, and formal math performance that have been demonstrated so far might in part go back to the selection of tasks and measures that were used in this research. Finally, we will outline that the connection between formal math skills and approximate arithmetic tasks should not be fully mediated by ordinal knowledge, and that these tasks should even outnumber numerical ordering ability in its predictive capacity for formal mathematical knowledge.

Besides (and beyond) the aspect of quantity, numerical representations also include an aspect of order (Vogel, Remark, & Ansari, 2015). The latter consists in the knowledge that the quantity of 'six' comes after 'five' and before 'seven', for analogue as well as for symbolic representations of magnitude (Brannon, 2002; Lyons & Beilock, 2009). This ability should not be confused with an understanding of cardinality. While understanding cardinality might just mean to know that three bananas are *different* from five bananas, ordinal knowledge goes beyond that in actually understanding that five bananas are *more* than three bananas. Or, as Brannon (2002, p. 224) puts it, the difference between the two concepts tackles the question if to a child twoness is to threeness what a blender is to a chair, or if twoness and threeness are conceived as different positions on the numerical continuum (ANS or analogue magnitude system). She found that already 11 months-old infants were able to discriminate sequences of dot patterns according to their numerically ascending or descending order, clearly demonstrating ordinal understanding in this domain (Brannon, 2002).

This and other findings (Anderson & Cordes, 2013; Berteletti, Lucangeli, & Zorzi, 2012; Brannon & Van de Walle, 2001) that locate numerical ordering ability already on the representational level of the ANS fit a proposal from Lyons and Beilock (2011). They assume that the understanding of the ordinal relations between quantities might be an important step in the development from the approximate representations of the ANS to later representation and arithmetic processing of exact numerosities (Lyons & Beilock, 2011). They had found that numerical ordering abilities were helpful in acquiring novel numerical symbols. Adults who had to map approximate quantities (dot arrays) to the novel symbols performed better in using the symbols in a numerical context when they had explicitly focused on ordinal relationships between the symbols. Same participants who had relied most on ordinal information in the novel numerical symbol acquisition task were also the ones who in a second experiment showed greater ordinal knowledge in a formal context with Arabic numerals.

But why should the understanding of ordinal relations be an important transfer mechanism between approximate and exact representations of number? Lyons and Beilock (2011) reason that a system that represents quantity mainly in relation to other quantities might be suited best to transform the imprecise magnitude representations of the ANS into increasingly exact ones. They hypothesize that the ability to determine the relative ordinal position of a magnitude (numerical ordering ability, NO) offers this special characteristic. So, one should expect to find positive correlations between ANS acuity and NO as well as between NO and exact representations or processing. And in fact, in an adult sample they not only found these correlations, but beyond that they showed that the relationship between the ANS measure (nonsymbolic comparison) and exact arithmetic was completely mediated by numerical ordering ability (Lyons & Beilock, 2011).

It has to be noted, though, that comparison tasks themselves might include an ordinal component. If participants not only have to judge equality vs. inequality of two quantities, but instead have to answer which of two is 'bigger' or 'more' than the other (which is the case in almost all of the reported studies so far), one can actually assume that this is an ordinal judgment (but see Vogel et al., 2015). Then it would be less surprising that an ability tested with Arabic numerals mediated the relationship between the same ability measured nonsymbolically and performance in exact arithmetic. However, the fact that it was a *complete* mediation shows the importance of the format change – to transfer the understanding of ordinal relations onto Arabic symbols might actually have been the important step here.

However, approximate *arithmetic* tasks that tax ANS acuity as well as understanding the logic of addition or subtraction imply the understanding of even more basic principles than ordinality (like for example 'addition makes more' or part-whole relationships, see Cowan, 2003). Approximate arithmetic could thus be an even better predictor for exact arithmetic than a pure numerical ordering measure that only demands sorting of quantities.

First hints that this might be true stem from the before mentioned intervention study of Park and Brannon (2013). They demonstrated that ten sessions with a nonsymbolic approximate arithmetic task improved adults´ performance in exact symbolic arithmetic to a significantly greater degree than a pure numerical ordering intervention (sorting strings of three numerals in ascending or descending order). In accordance with what Lyons and Beilock (2011) had found for adults, the latter training also improved math performance. However, it did so only to the same extent as a general knowledge intervention that was given to the control condition. The latter seems surprising when thinking of the primal role Lyons and Beilock (2011) hypothesized for numerical ordering and its relation to formal math performance. However, the – compared to the approximate arithmetic task – relatively small effect of the NO intervention in Park and Brannon´s work (2013) fits the explanation that the complete mediation Lyons and Beilock (2011) had found was mainly based on the process that was required in both their ANS measure (comparison) and their numerical ordering task. If in turn the ANS measure and the task assessing formal math performance share more requirements than ordinal knowledge, the mediation should disappear. Park and Brannon´s (2013) findings can tentatively be interpreted in that way, just like their conclusion that approximate arithmetic and symbolic arithmetic share a stronger cognitive foundation than numerical ordering does with either one of them.

For younger participants, the picture is a different one anyway. vanMarle et al. (2014) found in a sample of preschoolers a positive association of numerical ordering with nonsymbolic ANS measures as well as with exact math performance. However, NO did not mediate the relationship between ANS and math performance. This is especially noteworthy given the fact that also in this study, the ANS measure consisted in nonsymbolic comparison which constitutes a judgment of ordinality itself. But also other studies failed to find mediating relationships between children´s performances in comparison tasks, numerical ordering and exact math skills.

For example, in line with vanMarle et al.´s (2014) results, in the already mentioned more recent study of Lyons et al. (2014), the authors compared different nonsymbolic and symbolic comparison tasks with numerical ordering ability regarding their respective relationship with formal math performance in children from first to sixth grade. For NO, they found a steady increase in predictive capacity from grade 1 (poor predictor) to grade 6 (strongest of all predictors), with significance first being reached at grade 3. Consistent with their previous account (Lyons & Beilock, 2011), they conclude that only in the course of formal education, the ordinal information in numbers becomes more important than the relative magnitude. This is in accordance with the results of Attout and Noël (2014) who found in their sample of kindergartners, first- and second graders that numerical ordering was only simultaneously (that is, correlative) related to calculation ability, but not longitudinally (no significant regression).

As in all research fields reported above, the absence of research administering symbolic approximation measures is striking. To our knowledge, there is no study that tested the according relationships between approximate symbolic arithmetic, numerical ordering ability, and exact arithmetic. The only study using a nonsymbolic approximate arithmetic task at all tested adults (Park & Brannon, 2013). So, also in this field there is the need to test children with the two formats of approximate arithmetic and the possible role of numerical ordering ability as a mediating mechanism. According to the assumptions outlined above, we expect no (or at least no complete) mediation of the relationship between both arithmetic tasks. Instead, it might even be the case that performance in symbolic approximate arithmetic mediates a connection between nonsymbolic approximate arithmetic and subsequent formal math performance.

6 Research questions and hypotheses

The previous chapters have focused on the early development of human numerical abilities. It has been outlined that humans and nonhuman animals share a fundamental and evolutionary old system to represent large numerosities. This system today is called the ANS (approximate number system), but is also known under the terms 'analogue magnitude system' or 'number sense'. A lot of research has been dedicated to this system because of the still ongoing debate if the ANS might be the crucial building block of the uniquely human later exact mathematical skills (formal math performance). Amongst others, important lines of research have focused on possible concurrent or even predictive relationships between the ANS and formal math performance.

Problematic, though, is an observed imbalance in this research. By far the most studies only administer nonsymbolic measures to assess ANS acuity. This points at an important proposition of the TCM (Triple Code Model), the so far most influential account and overview of numerical processing (Dehaene, 1992; Dehaene et al., 2003).

Dehaene had suggested in his model that all approximate processing (quantification, comparison, arithmetic) is carried out in the analogue magnitude code and hence on the representational level of the ANS. This would be the case irrespective of the format the respective problem is presented to us originally, may it be symbolic (in Arabic numerals or verbally), or nonsymbolic.

However, this special proposition of the TCM seems to be an explicit or implicit topic of disagreement. While in the explicit discussion there is some research presenting evidence that symbolic and nonsymbolic approximation might in fact be qualitatively different (Ebersbach et al., 2013, Kolkmann et al., 2013; Xenidou-Dervou et al., 2013;), and on the contrary also evidence demonstrating its deep relationship (Gilmore et al., 2007, Moyer & Landauer, 1967; Mussolin et al., 2010; Notebaert, 2011, Piazza et al., 2007, Pinheiro-Chagas et al., 2014); in other studies this aspect simply seems to be forgotten and is not even mentioned when confining the applied ANS measures to the nonsymbolic format.

In this chapter, we will present our hypotheses and research questions which are basically centered around the omission of approximate arithmetic in general, and more specifically, symbolic approximate arithmetic in current research. We report which knowledge gaps result from the selective use of tasks regarding the questions which ANS measures are suited best to predict later formal math performance, what might mediate this relationship, how socioeconomic factors influence numerical competencies and how approximation tasks can possibly foster exact symbolic skills as they are required in formal math.

6.1 Prediction: Which ANS measures can predict later formal math performance?

The upper mentioned problem is especially noteworthy given the fact that the studies that actually did apply a symbolic ANS measure more often found a strong connection between ANS and formal math performance than was obtained with nonsymbolic measures. However, to narrow the academic void, one has to note that amongst several possible ANS measures (e.g. approximation production, approximation perception, comparison, approximate arithmetic) the vast majority of this research also restrained itself on the comparison measure. Nevertheless, Gilmore et al. (2010) had found that also preschoolers´ performance in nonsymbolic approximate *arithmetic* could predict their formal math performance one year later. Pinheiro-Chagas et al. (2014) found with children in grades 1 to 6 that predictive capacity of nonsymbolic comparison for formal math was completely mediated by their performance in a nonsymbolic approximate *arithmetic* task. Also in an adult training study, an advantage of nonsymbolic approximate arithmetic compared to a numerical ordering intervention was found (Park & Brannon, 2013). We take this as a hint that the ability to manipulate quantities and to understand the inherent logic of arithmetic might be the actual interface between approximate and exact numerical abilities. This should apply to symbolic approximate arithmetic, too. Consequently, we expect that both a symbolic as well as a nonsymbolic approximate arithmetic task should predict later exact math skills.

6.2 Numerical Ordering: To what extend does NO mediate between approximate arithmetic and formal math performance?

To shift the focus from comparison tasks to approximate arithmetic tasks should also implicate specific relationships with another numerical ability, that is the understanding of ordinal relations, or numerical ordering ability (NO). NO, measured by sequencing strings of Arabic numerals, has been found to completely mediate the relationship between ANS (measured by a nonsymbolic comparison task) and exact arithmetic in adults (Lyons & Beilock, 2011). However, a comparison task demanding the participant to indicate which magnitude is the larger one (and not only to judge equality or inequality) in itself consists an ordinal judgment. A tight connection between these two tasks is thus not surprising. Nevertheless, a mediation through the *symbolic* measure of judging ordinality indicates that the aspect of format (symbolic sequences) plays an important role in the connection. It thus might be that the *mapping* of understanding (in this case of ordinality) from the analogue level onto the Arabic representational code is the more fundamental basis of the relationship.

A similar study with preschoolers (vanMarle et al., 2014) also points in this direction. They could not replicate the complete mediation of the relationship between nonsymbolic comparison and formal math by numerical ordering in this age group. It thus might be that in this age group, other numerical aspects are more important than NO. In line with this thought, vanMarle et al. (2014), who tested several additional symbolic skills in their study, found other symbolic control measures like numerical recognition or counting to be amongst the mediators for the relationship between ANS and formal math. This confirmed findings that numerical ordering becomes a more important skill for predicting success in formal mathematics not before grade 3 (Lyons et al., 2014), that is after the shift from logarithmic to linear representation. Based on these results, it seems likely that not numerical ordering per se is the mediating source of the connection between ANS and exact math performance. It might rather be that this relationship at a given stage of development is specifically determined by the ability to map the understanding of numerical processes and operations one has achieved on the level of the ANS to onto the formal, symbolic representation. While this might be ordinal knowledge from grade 3 onwards, for preschoolers it is understanding of cardinality (Lyons et al., 2014). As children already are able to understand the basic logic of arithmetic when entering school (Gilmore et al., 2007), one should assume that in a study testing the ANS with arithmetic tasks, not numerical ordering but rather the mapping of arithmetic understanding onto symbolic representations might be a mediating factor.

Coming from these results, in a study employing approximate arithmetic tasks, numerical ordering should not fully mediate a possible relationship between a nonsymbolic (or symbolic) approximate arithmetic task, because there is more to understanding the logic of arithmetic than ordinal knowledge. If furthermore the ability to transcode the according understanding onto symbolic representations mediates the relationship between ANS measure and formal math performance, it should be expected that symbolic approximate arithmetic might be a stronger predictor for math success at the end of first grade than nonsymbolic approximate arithmetic.

6.3 SES and its connection to approximate and exact arithmetic in symbolic and nonsymbolic format

The robust finding of a relationship between children's socioeconomic status and their math performance in school has caused furore not only in Germany in the context of large-scale assessments of learning processes. Consequently, similar to the search for early indicators of the development of formal math skills, there is also a line of research that tries to disentangle numerical skills that are influenced by environmental factors like the SES, from numerical abilities that are not. The most prominent explanation of how a child´s socioeconomic status might affect his/her numerical proficiency is the learning environment at home. For example, Starkey and Klein (2008) found that in middle-class households there were more math-related activities like board games, learning software, but also more inclusion of mathematical content in daily routines than in working-class homes. Poorer symbolic number knowledge in the first group of children is supposed to be the result of these differences and the reason for worse performance in tasks assessing different numerical skills in low-SES participants. Thus, the format in which a mathematical problem is presented should determine the impact socioeconomic factors have. This explanation is debatable, as a review of research also allows for alternative accounts.

During the last 25 years, increasing interest has been put into the question of the fine tuning of this relationship. For instance, Gilmore et al. (2010) recently found that preschoolers´ nonsymbolic approximate arithmetic skills did not vary with their respective socioeconomic status, contrary to their later formal math performance which was significantly worse in low-SES children. Earlier, Jordan et al. (1992) have reported that preschoolers of different SES showed an equal performance in exact addition and subtraction problems that were presented to them nonsymbolically, whereas the same problems were solved considerably worse by low-SES children than by their higher-SES peers when presented in a symbolic format. This supports the upper mentioned view that understanding of Arabic numerals or number words might be the important impediment in low-SES children. Although exact calculation presented both in a symbolic as well as in a nonsymbolic format should be carried out in one of the two symbolic codes[4], Jordan et al. (1992) found an influence of SES only for the exact *symbolic* processing that demanded understanding of the culturally acquired Arabic or verbal representatives. Mejias and Schiltz (2013), however, did find an effect of SES also on preschoolers´ nonsymbolic exact number representation in slightly younger children.

So, these and other (see Chapter 4.2) results are not conclusive, in that they leave room for different explanations of the influence that socioeconomic factors have on mathematical performance. Like in other research fields, in many studies approximate processing has not been tested symbolically; and the early study of Jordan et al. (1992) as well as the one by Mejias and Schiltz (2013) represents a big exception in that they administered a nonsymbolic *exact* task. Taken together, a look at these results again is a reminder that 'nonsymbolic' cannot automatically be considered as a recruitment of the representational level of the ANS, and that the manipulation of symbolically presented quantities not necessarily has to happen in one of the two symbolic codes (meaning: exactly). The inconsistent results also allow for alternative accounts to specify the connection between SES and math performance. For example, it is possible that it is the representational code in which an operation has to be *performed* (which depends on the demanded accuracy of calculations: exact vs. approximate operations, see Dehaene, 1992) instead of an impairment of the mere symbolic number knowledge which is crucial for the relationship. This would advance an alternative explanation for the connection between SES and math performance. It has been found that low SES also affects children's linguistic abilities. Given the proposal of the TCM that both symbolic codes have developed in dependence of human language skills, the underlying mechanism of the mentioned alternative account (demanded accuracy of the operation) might consist in linguistic abilities mediating between SES and numerical processes. To test both possibilities, a full realization of the 2x2 design has to be carried out. Two alternative hypotheses are formulated:

If impaired symbolic number knowledge is the main factor in differences of numerical processing between children of different socioeconomic backgrounds, one should ex-

[4] According to the TCM, nonsymbolic exact processing would only be possible with small numerosities to which one would have to generate whatever natured internal numerons

pect to mainly find differences between children´s performances in tasks that are presented symbolically. Crucially, this should also apply to approximate operations because the impeded processing of the number symbols would in turn at least slow down the respective approximation task. Performances in both nonsymbolic tasks, on the other hand, should be resistant to SES influence.

If on the other hand impaired linguistic abilities in low-SES children drive the relationship between SES and formal math, this should be reflected in a stronger influence of SES on the exact tasks, both symbolic and nonsymbolic. These tasks have to be carried out in one of the two symbolic codes which have been found to be affected by language skills (Starkey & Klein, 2008; Spelke, 2011; Walsh, 2003). The approximate tasks, on the other hand, can be carried out in the language-independent analogue magnitude code and hence the performances should not vary according to SES. For symbolic approximation, however, we assign this hypothesis with a question mark, because impaired language skills might already hinder the representation of the according quantities before processing them approximately, and in consequence also decelerate symbolic approximate processing in low-SES children. It thus might be that both approximate task remain unaffected by SES, or that only the nonsymbolic approximate task is robust against its influence. However, the nonsymbolic exact performance should vary with SES.

6.4 Training: In what ways can approximation benefit exact symbolic math skills?

Recent findings suggest that activating children’s precursory knowledge by presenting nonsymbolic problems or approximate arithmetic problems (both symbolic as well as nonsymbolic) can positively influence their performance of exact symbolic arithmetic. However, up to now only few studies have investigated whether approximate arithmetic also enhances the understanding of less basic arithmetic principles. Approximate arithmetic might activate existing precursory conceptual or procedural knowledge that is useful in later exact calculation. In our second series containing two intervention experiments, our goal was to test if activating mathematical knowledge of an arithmetic principle in approximate arithmetic tasks boosts the exploitation of the same knowledge in exact arithmetic problems.

As there is a predictive relationship between approximate arithmetic and performance in exact (formal) arithmetic, there is also the possibility to use this relationship in terms of fostering children´s formal mathematic abilities by means of approximation. Some research has already been dedicated to this aim and positive results have been obtained. For instance, Hyde et al. (2014) found that brief nonsymbolic approximation served as a beneficial influence on subsequent exact symbolic arithmetic performance in first graders. However, some further questions remain unresolved. On the one hand, to our knowledge, no one has tested a possible beneficial effect of symbolic approximate arithmetic yet, although it seems plausible that transfer from the representational level of the ANS onto a formal math context might be facilitated when the original problem is already presented in a symbolic format. Furthermore, the few studies that tested interventions based on approximation, only tested the influence on general arithmetic performance. Given the fact that already very young children can understand less basic principles like inversion or commutativity, it seems promising to further foster this understanding by referring to children´s approximate precursory concepts of these principles. This also allows the differentiation if approximation only benefits *procedural* knowledge, or a deeper understanding in terms of *conceptual* knowledge. We assume that both nonsymbolic as well as symbolic approximate arithmetic should be suited to induce an arithmetic principle in children at the beginning of formal education. However, if only procedural knowledge can profit from this reference to precursory understanding of principles, or even conceptual understanding, is our open research question.

7 Rationale of Studies 1 and 2

It was outlined that the relationship between exact and approximate mathematical processing in both nonsymbolic as well as symbolic representation is not yet clear when it comes to predictive power of specific kinds of approximation for formal math performance. At the same time, also the effects of socioeconomic status on specific aspects of exact and approximate processing need further elucidation.

We aim to explore the longitudinal interrelations between exact and approximate arithmetic competencies in both a nonsymbolic as well as a symbolic representation format. Furthermore, we try to differentiate the influence of socioeconomic status on each of these early mathematical competencies. Thus, Study 1 (exploration) and Study 2 (replication) were based mostly on the experiments of Gilmore, McCarthy, and Spelke (2010) as well as the study of Mejias and Schiltz (2013). So far it has been found that nonsymbolic approximate competencies are not/less prone to the detrimental influence of socioeconomic factors (Gilmore et al., 2010). We disentangle the demanded accuracy (exact vs. approximate processing to reach the answer) and the format of the presented problems (nonsymbolic vs. symbolic). To that aim we presented first graders at the beginning of formal schooling with four kinds of tasks testing their mathematical competencies: A nonsymbolic exact arithmetic task, a nonsymbolic approximate arithmetic task, a symbolic exact arithmetic task and a symbolic approximate arithmetic task. At the end of the school year, we tested children´s math achievement with a standardized measure. This enabled us to assess longitudinal predictive relationships between our two measures of the ANS and formal math performance as it is required in school.

As control variables we included an adapted working memory test (Study 1 and 2) and a measure of numerical ordering ability as well as a standardized measure of reading competence (both only Study 2).

We chose to study the period from the beginning to the end of first grade. The studies reported in chapters 3 and 0 explored similar research questions with younger children and obtained diverging results. To assess approximate arithmetic also symbolically, children need to have some knowledge of symbolic numbers, which should for all children be the case at the beginning of first grade. Furthermore, there are practical implications of this choice. First grade is an educational period in which teachers get to know their pupils and need to adapt their teaching contents and methods to their level. Hence teachers and educators would profit most from knowing as much as possible about the quality and state of their students’ knowledge when they enter the school system.

8 Rationale of Experiments 3 and 4

In the Experiments 3 and 4 we turn to the question if approximate arithmetic based on a specific arithmetic principle can foster the understanding and application of this principle in a more formal exact symbolic context. As far as we know, all studies using approximation tasks as an intervention or training so far have only tested the respective influence on exact arithmetic as such (that is, general addition or subtraction problems). However, the inclusion of more advanced principles like inversion or commutativity not only promises results about the possibility to foster more complex arithmetical reasoning in children just starting formal education. In addition, principles like the law of commutativity allow differentiating between the mere strategy use and a deeper meta-knowledge, because of a specific feature: That their exploitation is time-saving and thus enables to faster problem solving. Hence one can see if a principle has been used (= procedurally) by comparing the solution times of problems that allow principle-based shortcuts (derived-fact strategy use, see Dowker, 2012) with the solution times of problems that are matched in difficulty but do not allow this principle use. Conceptual knowledge on the other hand is usually tested in a combined procedure: After each problem, participants are directly asked how they worked out the answer; and then these answers are categorized according to the demonstrated meta-knowledge. However, this procedure is debatable in that some children might be better in verbalizing their thoughts than others. Furthermore, the wording of the question as well as the repeated asking might trigger children to use the specific shortcut, as well as to give one specific answer again and again. So, this typical assessment is not suitable at least if one is interested in a 'baseline', which is the spontaneous demonstration of procedural and/or conceptual knowledge.

So, we aimed to assess procedural and conceptual principle knowledge separately and subsequent to an induction that presents the principle in an approximate way. We decided to test the additive law of commutativity, which states that in binary operations of addition, the order of the operands does not affect their sum (cf., $a + b = b + a$; see Cowan, 2003). Its core property, the order-irrelevance principle, is ubiquitous in everyday situations and is understood by children younger than one year (Wynn, 1992).

From this precursory understanding follows a good chance that commutativity can be induced in children without explicitly teaching them. In Experiment 3, we constructed two alternative inductions displaying commutativity in approximation problems to second graders. One was a symbolic approximate arithmetic task containing interspersed commutative trials, the other was an analogue nonsymbolic approximate arithmetic task. In a third condition, instead of an induction we instructed participants explicitly about the law of commutativity to assess the ecological value of any possible fostering effect. If a direct instruction works significantly better in making children use specific strategies and understand the underlying logic, in practical terms an also working induction without the need of an explanation would not represent real progress. After receiving their induction or explanation, children of all three conditions worked through two tasks taken from Haider et al. (2014). In the 'computation task', procedural knowledge of commutativity was assessed. The task consisted of two booklets of addition problems matched in difficulty. In one booklet, however, commutative problem pairs were interspersed between the filler problems. The other booklet did not contain the possibility for such a shortcut. By comparing the number of solved problems within the given time limit, we were able to determine if a child exploited the shortcut and hence solved more problems in the commutative booklet. The second task was the 'judgment task' (Haider et al., 2014), which was constructed to measure conceptual commutativity knowledge separately from the strategy use. The booklet of this task also contained addition problems interspersed with commuted problem pairs. But in this task, children were instructed not to actually solve the problems. Instead, they should only look at the problems and mark those that – if one was to solve them – would not need calculation to get the results. The only problems that actually provided this characteristic were the commutative ones. By computing the hits, false alarms and sensitivity we were able to determine children´s metacognitive, conceptual knowledge without triggering specific answers with our questioning. Thus, the design of the two experiments should allow us to assess potential beneficial effects of commuted pairs in nonsymbolic and symbolic approximate arithmetic problems on the procedural and the conceptual knowledge of commutativity. Calculating correlations between the two kinds of knowledge promises insights in when an integration of them to an increasingly abstract concept of commutativity might occur. According to Haider et al. (2014), this should not be the case before grade 3. However, also a possible beneficial effect of our inductions on the integration of conceptual and procedural knowledge can be explored.

9 Study 1

This pilot study was conducted to test our self-constructed main tasks as well as to explore the field of our research interest. We administered the four main tasks including addition and subtraction problems in a nonsymbolic exact, nonsymbolic approximate, symbolic exact and symbolic approximate representation at the beginning of first grade. As we found that the symbolic exact problems produced results close to a ceiling effect in our participants, which could be explainable by methodological issues (see Chapter 9.2.1), we retested this format in the middle of grade 1 again, with some adjustments that were further to be improved in our replication experiment (Study 2).

At the end of the school year, the same children completed a standardized test of math achievement (DEMAT 1+) for us to assess their mathematical abilities as demanded in the German school system, as well as an adapted working memory measure.

9.1 Method

Participants. Participants were 109 first graders (57 girls) from five classes in two primary schools located in two high SES suburbs of Cologne. In the first two classes we administered our first version of the four main tasks which at that time contained 14 problems per task. It turned out that this were too many problems for the first graders, at least according to our time scheme. For the remaining three classes we used the second version of our tasks and reduced the number of problems per task to ten (approximation tasks) respectively eight (exact tasks) and only did our final testing session at the end of grade 1 in these three classes. Thus, only the 64 participants (32 girls) from these three classes, with a mean age of 6 years and 9 months (*SD* = 3.7 months), remained in the analyses reported below.

Material. In the following the tasks are described. We start by providing the common structure of the exact tasks in summary, before specifying the respective features of their symbolic vs. nonsymbolic realization. After that, the same is given for the approximate tasks.

Exact tasks

For the symbolic and nonsymbolic exact task, problems were identical. In both exact tasks, we presented the first graders with eight two-operand arithmetic problems (four addition and four subtraction problems). For the addition problems, the smallest addend was 3, the largest 6, with results ranging from 9 to 11. For the subtraction problems, minuends ranged from 7 to 10, subtrahends from 3 to 7. Each subtraction problem resulted in a total of 3 or 4.

We aimed not to include operands within the subitizing range because in a nonsymbolic representation, quantities within this range (up to 3 or 4) can be easily and exactly assessed at a glance even by very young children. Furthermore, there are hints that processing these kinds of quantities is different to the processing of larger numbers (Lonnemann, Linkersdorfer, Hasselhorn, & Lindberg, 2011, recommend not to mix quantities within and beyond the subitizing range in nonsymbolic problems; see also Burr et al., 2010; Siegler & Booth, 2005). Furthermore, we ensured that there were no ties in the problems because adding two identical numbers leads to a characteristic and systematic underestimation of the result (Charras, Brod, & Lupiáñez, 2012). The position of the larger addend was counterbalanced across the problems because depending on the computational strategy the individual child used, problems 'starting' with the larger addend might be easier to solve. For instance, this might be the case when still using the counting-on strategy instead of the more advanced counting-on-larger strategy (see Baroody & Gannon, 1984). The problems are depicted in Appendix A.

Symbolic exact task

The exact task consisted of the addition and subtraction problems as described above. In this task, problems were presented in Arabic numerals via a Microsoft PowerPoint presentation (see Figure 3), each problem on its own slide. From the onset of each slide, the problem was visible for 10 seconds and then disappeared. After that the next slide appeared with a symbol that indicated that children were now to give the answer in their booklet.

Nonsymbolic exact task

We adapted an idea of Levine et al. (1992) who were the first who included measurement of *exact* nonsymbolic processing. In their so-called nonverbal task, an experimenter slid two arrays of disks behind an occluder while being observed by the participating child. During the experiment, these two collections of disks (the 'addends') were at no time seen simultaneously. The child then had to reproduce the exact number of disks that was now behind the occluder on his/her own desk. Thus neither verbal nor symbolic labeling of the quantities was necessary. We adapted this task to a group setting and presented the same addition problems as in the exact symbolic task, again via a Microsoft PowerPoint presentation (see Figure 4). This time children saw an array of large dots above the image of a container. After 2 seconds, the dots started falling into the container one after another, each taking 0.5 seconds for this path. For the addition problems, two seconds later another array of dots appeared above the container, and after 2 further seconds the dots started falling into the container again. Two seconds after the last dot had reached the container, the next slide appeared with a symbol indicating that the children should now give the answer in their booklet. After 8 seconds, a small symbol appeared in the corner of the slide. At this point, the experimenter told children to come to an end with their answer. The answering sheet was a collection of fifteen loosely scattered circles. Participants were instructed to cross out as many of the circles as there were now dots in the container. We constructed the erratically distributed answering field because a regular matrix might suggest a counting and computing strategy (Siegler & Booth, 2005; Starkey & McCandliss, 2014). For the subtraction problems, two seconds after the last dot of the first array (minuend) had disappeared in the container, a subset of the occluded dots (subtrahend) 'jumped' out of the container and together remained visible for 2 seconds. After that, the slide with the answering symbol appeared.

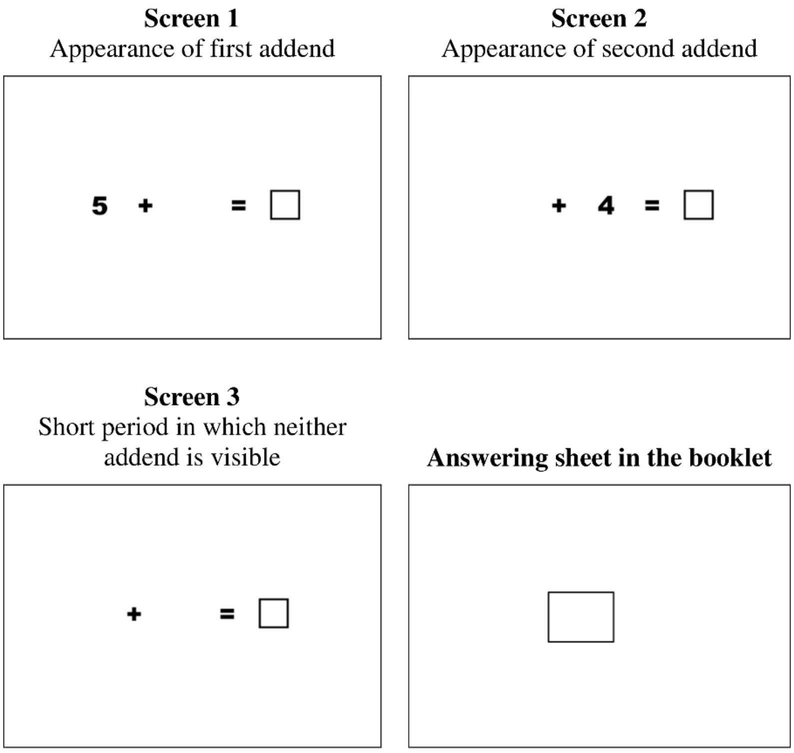

Figure 3. Study 1: Example of a trial in the retested symbolic exact task. The problems were administered as a Microsoft PowerPoint presentation, but the solutions had to be given in an individual answering booklet by each child. In the original symbolic exact task, addends did not disappear but both addends were simultaneously visible for ten seconds

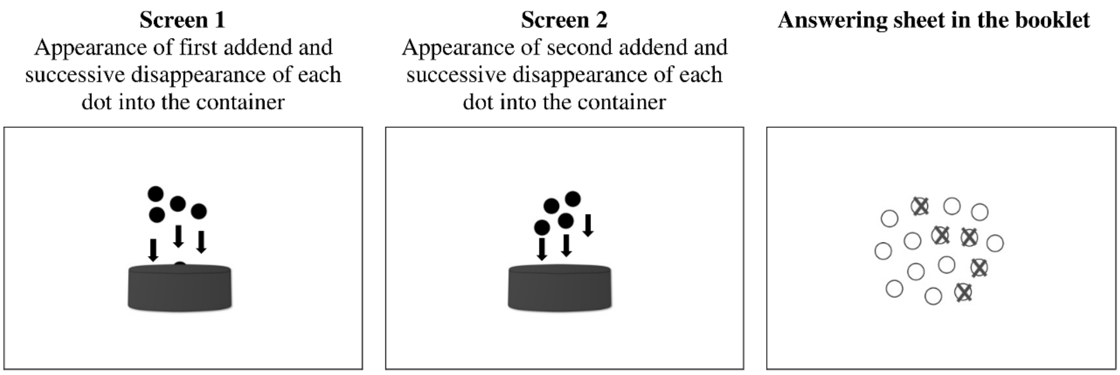

Figure 4. Study 1: Example of a trial in the nonsymbolic exact task. The problems were administered as a Microsoft PowerPoint presentation, but the solutions had to be given in an individual answering booklet by each child. The black arrows in this figure are added to indicate the motion in the original presentation

Approximate tasks

We administered a nonsymbolic and a symbolic approximate addition task, both widely accepted measures of the ANS (Gilmore et al., 2014). Both tasks consisted of the same ten trials. In each trial, children had to compare the result of a two-operand arithmetic problem (five addition and five subtraction problems) to a reference quantity to indicate if the result or the reference quantity was the larger one. The quantities were presented as either numerals within images of candies (symbolic task) or arrays of marbles (non-symbolic task), and belonged to two children, Tim and Lisa. On the next screen, one of the children received some more candies/marbles in addition to the ones he/she already owned (addition), respectively had to give away some of his/her candies (subtraction). This process formed the arithmetic problem of the trial. The other child did neither lose nor get additional candies/marbles in the respective trial, his/her candies thus forming the reference quantity. At the end of each trial, participants had to indicate which of the two children had more candies/marbles. For addition as well as subtraction problems, in three of the five respective trials, 'Tim' was the correct answer, in the remaining two it was 'Lisa'. In one subtraction and two addition trials, the larger number was associated with the arithmetic operation, in the remaining trials with the reference quantity.

Children could not calculate but had to estimate the total of each problem (which made the comparison itself an approximate one, too), because the numbers were too large for a first grader to compute them; respectively the arrays were too large to be exactly counted (and thus computed). In the five addition problems, the smallest addend was 13, the largest 20, with results ranging from 30 to 39. The five according reference quantities ranged from 21 to 65. We realized three different ratios of the resulting quantities that had to be compared. Four comparisons had to be made in the easiest ratio 3:5, three comparisons in the ratio 2:3 and three in the ratio 3:4. Same ratios were realized in the subtraction problems. The minuends ranged from 27 to 42 with subtrahends between 13 and 27. Results of the subtraction problems were between 12 and 20, reference quantities between 12 and 30.

Both tasks were again presented via a Microsoft PowerPoint Presentation. The problems are depicted in Appendix B.

Symbolic approximate task

In this task, each operand (addends, minuends and subtrahends) and the reference quantity were presented as Arabic numerals within the image of a large candy (see Figure 5). With the onset of the slide, the first addend was directly visible and was positioned beneath one of the two children. After 2 seconds the candy with the addend fell down into a container located below the original position of the candy. In the addition problems, two seconds later the second addend appeared in a candy image on this original position and followed the first addend after 2 more seconds into the container of the respective child. In the subtraction problems, a candy with a different (smaller) number, the subtrahend, 'jumped' out of the container and remained visible for 2 seconds. After that, for both kinds of problems, an Arabic numeral within a candy image appeared beneath the other child as the reference quantity. It remained visible until the next slide appeared after 4 seconds with a symbol to indicate that the children were now to give the answer. The answering sheet contained the images of the two children with one box below the image of each child. The participants had to cross out the box beneath the child they felt had the most candies at the end of the respective trial.

Nonsymbolic approximate task

The nonsymbolic approximate task was identical to the symbolic approximate task described above with the exception that the addends below the children were not presented as Arabic numerals in candy images anymore. Instead, this time the quantities were presented as arrays of 'marbles' (dot patterns) that disappeared into the containers (see Figure 6). Note that contrary to the exact tasks, due to the higher operands in the approximate tasks the arrays that represent these operands disappeared *together* in the container, not one after the other.

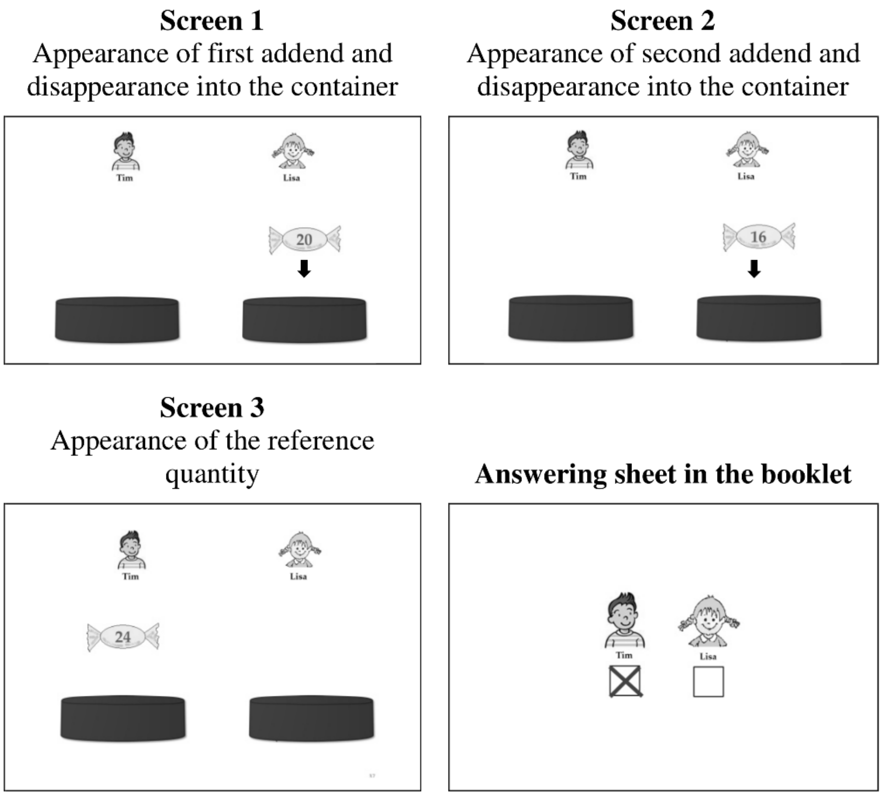

Figure 5. Study 1: Example of a trial in the symbolic approximate task. The problems were administered as a Microsoft PowerPoint presentation, but the solutions had to be given in an individual answering booklet by each child. The black arrows in the figure are added to indicate the motion in the original presentation

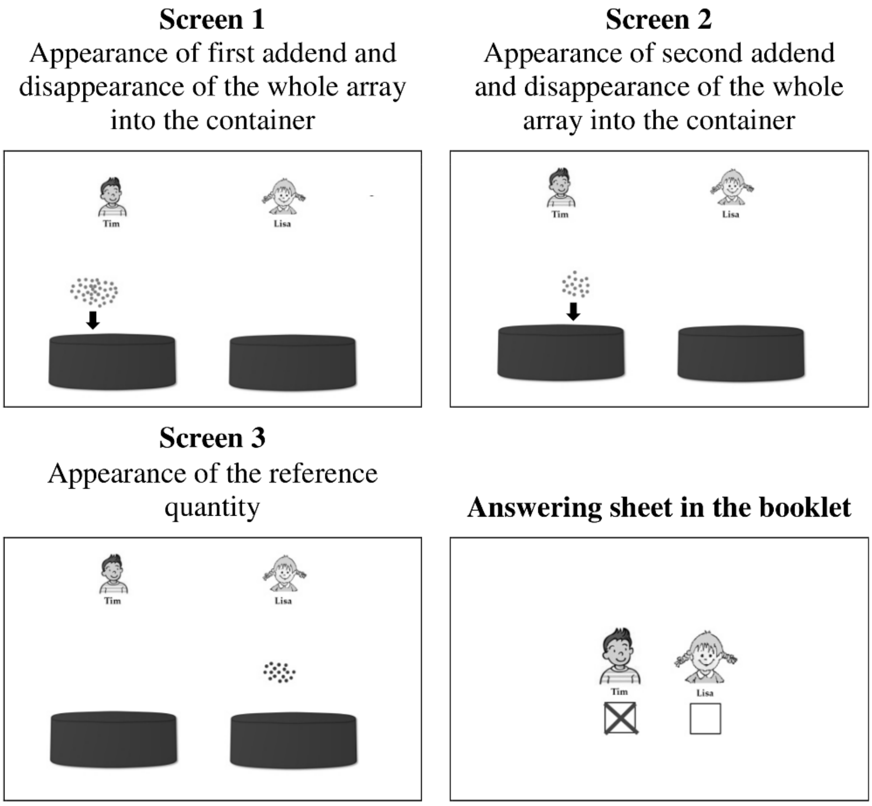

Figure 6. Study 1: Example for a trial in the nonsymbolic approximate task. The problems were administered as a Microsoft PowerPoint presentation, but the solutions had to be given in an individual answering booklet by each child. The black arrows in the figure are added to indicate the motion in the original presentation

Working Memory (WM) Task

Our task to assess children's working memory capacity was adapted from the ZN (*Zahlen Nachsprechen*, repeating numbers) subtest of the HAWIK IV (Petermann & Petermann, 2007). To administer the test in a group setting, we constructed an answering booklet in which children should write down the string of single-digit numbers we read out loud to them. There were always two trials with the same string length. The first two of the twelve trials started with a string length of two numbers, the last two trials consequently contained seven numbers. The strings were read out loud only once to the group, and children were instructed and supervised to only start writing down the numbers when the experimenter was finished reading the sequence. Note that we did not aim to compare our participants to other samples of their age regarding their working memory, as WM was only added as a control variable. This should ensure that possible difficulties with nonsymbolic quantities, especially the larger addends in the exact nonsymbolic task, were not only due to larger working memory load, as Berends and van

Lieshout (2009) found that illustrations might interfere with children´s math performance in that they overload participants´ working memory.

Math Performance: DEMAT 1+

The DEMAT 1+ (Deutscher Mathematiktest für erste Klassen, Krajewski, Küspert, & Schneider, 2001) is a German standardized math achievement measure for children at the end of first or beginning of second grade. It contains 36 problems in the nine subtests 1) quantities-numbers, 2) number space, 3) addition, 4) subtraction, 5) decomposition and finding the missing number, 6) part-whole knowledge, 7) chain problems, 8) inequations, and 9) story problems. The test is curricular valid in that it was constructed according to the mathematics curricula of all German constituent states and the manual reports a retest reliability of $r = .65$. In the following, short descriptions of each subtest are provided.

1) Quantities – Numbers. In this subtest, understanding of quantities with regard to non-symbolically presented arrays is tested in three problems.

2) Number space. To test children’s orientation in the number space, they have to name a marked position on the number line as an Arabic numeral, or, in reverse, map a given Arabic numeral to a place on the number line (four items).

3) Addition and 4) Subtraction. In this mixed subtest the flexibility in the problem solving algorithms is tested in addition and subtraction problems in the number space up to 20. Thus, also carrying procedures have to be undertaken (eight items + one practice problem).

5) Decomposition and finding the missing number. Children solve place-holder problems. Numbers either have to be decomposed into two components or already are composed of two components (four items + one practice problem).

6) Part-whole knowledge. In these problems, children have to understand that a number as a “whole” can be decomposed in different “parts” (e.g. 8 can be decomposed in 5 + 3 *and* in 6 + 2) (four items + one practice problem).

7) Chain problems. This subtest contains addition and subtraction problems in the number space up to 20 consisting of four operators (four items + one practice problem).

8) Ineqations. Here the understanding of the relational “greater-than-/less-than-sign” as well as the equal sign is tested (four items + one practice problem).

9) Story problems. In this subtest, children saw four images. For each image (and at the beginning for one practice problem), the experimenter read aloud a story problem to the children that tested the arithmetic understanding of exchange, comparison and equalize processes (see Riley, Greeno, & Heller, 1983)

Socioeconomic Status (SES)

To assess children´s SES, teachers gave us their classification of the children´s parents’ profession according to the EGP classes (Erikson, Goldthorpe, & Portocarrero, 1979). In this widely used scheme (Müller et al., 2007), families are allocated to social classes according to their employment status. The goal is that a given EGP category comprises individuals who are comparable in their levels of income, economic security, and degree of autonomy at their work place. In the EGP system, '1' stands for the highest, '7' for the lowest status. The classifications were made anonymous in that we only received the information in relation to participants’ subject number.

Procedure. The general timeline of our testing and an overview of the applied measures can be seen in Table 1. In the first quarter of the school year we administered our four main tasks as a group testing in each class. Explaining and processing the tasks took between 90 and 120 minutes including one short break. Each task had its own answering booklet covered with a different color. Before each task, the experimenter told the children which booklet they should place in front of them. For the approximate non-symbolic task, she explained the cover story of Tim and Lisa to the children and solved one practice trial together with the children. When there were no questions and all children paid attention, the presentation of the test trials started. After five problems a slide with a give way sign containing an exclamation mark appeared. The experimenter told the children that now something would be different in the trials to announce the change to subtraction problems. She explained that in the following trials, one of the two presented children would lose some of his/her marbles instead of getting additional ones. Also for subtraction, a practice problem was solved together with the children to ensure that they had understood the task. Then the five test trials were presented.

The approximate symbolic task followed after a short break. Children were told that now they would work with something new – large numbers. Again there would be Tim and Lisa, only that this time they had candies instead of marbles. And even if children would see only one candy for Tim respectively Lisa, the numbers in the candies would tell them how many candies the child has. The following procedure was analogue to the approximate nonsymbolic task, including the announcement of the change to subtraction. In both tasks, children were explicitly told that the marbles were too numerous / the numbers in the candies were too large to be counted respectively computed, and that they should only estimate the answer.

After the approximate tasks there was a break of about 10 minutes. Then the exact tasks were administered. The exact nonsymbolic test was the first for the majority of participants. The experimenter told the children that now, in contrast to the tests before, they should find the exact answer to the problems. A practice trial was shown and children were asked to mark as many dots on the according page of their answering booklet as there would be dots in the container. The experimenters asked and checked if the children marked the correct number of dots. When there were no further questions, the test trial presentation started. After the four addition problems, again the change to subtraction problems was indicated by the give way sign and the experimenter.

The exact symbolic test followed. Because arithmetic problems in Arabic notation are familiar to school children, there was not much explanation needed. Nevertheless, the experimenter solved a practice problem together with the children; and between the addition and subtraction problems she checked that they understand the concept of subtraction and the minus sign.

Our second (planned) testing session took place at the end of the school year. Again we administered our tests in a group setting in the class rooms. First, we tested the mathematical performance with the standardized measure DEMAT 1+. In its group version, it takes about 35 minutes. After that, we administered our adapted WM measure. Numbers were those of the Zahlen Nachsprechen (ZN) subtest of the HAWIK IV (Petermann & Petermann, 2007), instructions were adjusted to the group testing which required to not say the actual numbers but to write them down. This test took 10 minutes.

Table 1. Study 1: Timeline of testing

	T1 beginning of first grade	T2 Retesting	T3 end of term
administered measures	approximate symbolic task		DEMAT 1+
	approximate nonsymbolic task		working Memory (WM)
	exact symbolic task	exact symbolic task (adjusted)	
	exact nonsymbolic task		

9.2 Results

Results will be reported starting with a general overview of children´s performances in the approximate symbolic, approximate nonsymbolic, exact symbolic and exact nonsymbolic task, followed by presenting the signature of approximate processing (that is, the ratio effect) and a description of the retesting of the exact symbolic problems. After that, results will be presented ordered along our hypotheses and research questions as brought up in Chapter 6. In the longitudinal studies, our interest was in children´s performances in the 2x2 tasks and especially the question which tasks would intercorrelate; the influence the SES might have on each of the tasks; and the question what informal numerical ability predicts math performance best at the end of first grade. Note that our hypothesis concerning the possible mediating impact of numerical ordering ability is not tested before Study 2 and thus does not appear in this chapter.

9.2.1 Overview of T1 performance and retesting

In Table 2 performances in the four T1 tasks, which are exact symbolic, exact nonsymbolic, approximate symbolic, and approximate nonsymbolic arithmetic, are reported. To rule out the interpretation that potential differences in the accuracy scores are due to motivational differences in symbolic vs. nonsymbolic tasks, we first compared the number of solved problems within each of the two levels of accuracy[5] (nonsymbolic approximate vs. symbolic approximate; nonsymbolic exact vs. symbolic exact).

The number of solved problems did not differ between both approximate tasks (see Table 2). There were only two children who left out one trial in the nonsymbolic approxi-

[5] We did not compare the accuracy in approximate with that in exact tasks because of the differing answering format – it is much easier to guess the correct one out of *two* solutions ('Tim' or 'Lisa') than to independently produce an exact result to an arithmetic operation.

mate task, and one child that left out one trial of the symbolic approximate task. In the exact tasks, on the contrary, not all children actually solved each presented problem. However, the number of unsolved problems was very small, and a paired sample t-test indicated that there was no significant difference between the numbers of solved problems in the exact symbolic (*M* = 7.84, *SD* = .44) vs. the exact nonsymbolic task (*M* = 7.95, *SD* = .21), *t*(63) = -1.84, *p* = .07, see Table 2). This result indicates that there seem to be no motivational differences in processing problems of two different formats. For approximate as well as exact tasks, we thus chose proportions of correctly solved problems relative to all presented (addition/subtraction) problems of the respective task as our measure of accuracy.

Table 2. Study 1: Performances in the T1 tasks (SD in brackets)

Task		**Number of solved problems**	**Proportion correct**
approximate	nonsymbolic	9.969 *(.175)*	.772 *(.151)*
	symbolic	9.984 *(.125)*	.705 *(.186)*
exact	nonsymbolic	7.844 *(.444)*	.504 *(.227)*
	symbolic	7.953 *(.213)*	.768 *(.218)*

Accuracy. Children performed above chance in all four main tasks (see Table 2, note that even the 50% correct solutions in the exact nonsymbolic task are above chance as these problems were open). We were especially interested in the differential outcome of symbolic vs. nonsymbolic presented arithmetic within each level of accuracy. In the approximate tasks, children performed significantly better in the nonsymbolic (*M* =.77, *SD* = .15) than the symbolic task (*M* = .71, *SD* = .19, *t*(63) = 2.88, *p* = .005, *d* = .41); whereas in the exact tasks the reverse was the case (exact nonsymbolic *M* = .50, *SD* = .23 vs. symbolic *M* = .77, *SD* = .22, *t*(63) = -7.08, *p* < .001, *d* = -.724).

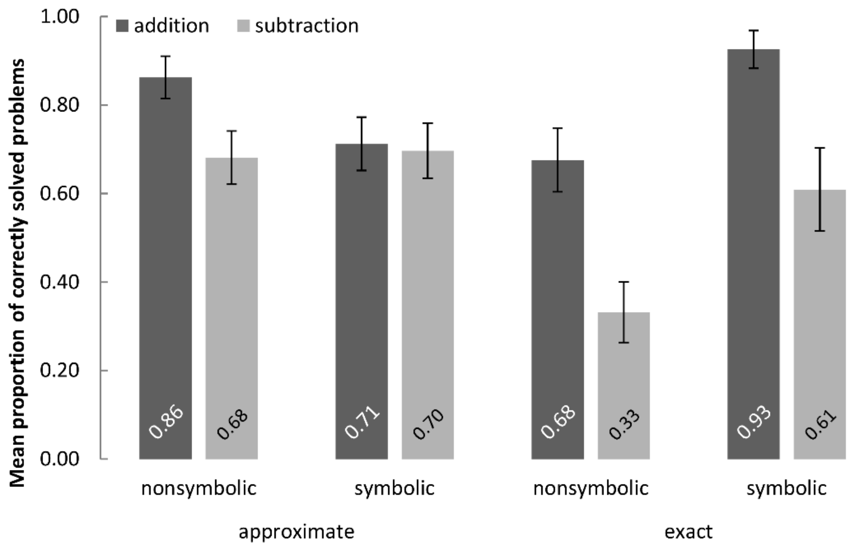

Figure 7. Study 1: Accuracy in the main tasks: addition vs. subtraction. Proportion correct solutions/presented problems for addition vs. subtraction problems in each task

We also compared the accuracy in each task separately for addition and subtraction problems. We found that in each task except for the approximate symbolic problems, addition problems were solved with significantly greater accuracy than subtraction problems (approximate nonsymbolic: $M = .86$, $SD = .19$ vs. $M = .68$, $SD = .24$, $t(62) = 4.62$, $p < .001$, $d = .058$; exact nonsymbolic: $M = .68$, $SD = .29$ vs. $M = .33$, $SD = .27$, $t(63) = 8.29$, $p < .001$, $d = .77$; exact symbolic: $M = .93$, $SD = .17$ vs. $M = .61$, $S = .38$, $t(63) = 6.55$, $p < .001$, $d = .70$, see Figure 7). Furthermore, there was no significant correlation between performance in the addition and subtraction problems within each task (see Table 3), except for the exact nonsymbolic one. We thus decided not to analyze subtraction performance separately in the following analyses, but kept subtraction problems in the general performance measures due to the small number of problems in each task.

Table 3. Study 1: Correlation of subtraction and addition in the main tasks. Correlation coefficients between performances for addition vs. subtraction problems for each task

Correlations of addition and subtraction accuracy within each main task

		r	*p*	N
Approximate	Nonsymbolic	-.043	.736	64
	Symbolic	.143	.259	64
Exact	Nonsymbolic	**.304***	.015	64
	Symbolic	.160	.207	64
	symbolic retested	.182	.157	62

Signature of the ANS in the approximation tasks. The ratio-effect, an increase in solution accuracy with increasing distance between the to-be-compared magnitudes, is a well-known sign of approximate processing. To ensure that children followed our instructions and actually approximated in these tasks, we conducted a one-way repeated-measures ANOVA over the three ratio levels for the nonsymbolic as well as the symbolic approximate task (see Figure 8). Analyses showed a significant effect for the nonsymbolic task ($F(2, 126) = 18.05$, $p < .001$, $\eta^2 = .223$) as well as for the symbolic task ($F(2, 126) = 9.64$, $p < .001$, $\eta^2 = .133$). In both tasks, significant linear contrasts confirmed the significant results (nonsymbolic: $F[1,63] = 29.856$, $p < .001$, $d = .96$; symbolic: $F[1,63] = 17.713$, $p < .001$, $d = .69$). Thus, we conclude that children in fact recruited their ANS to obtain the results in both formats.

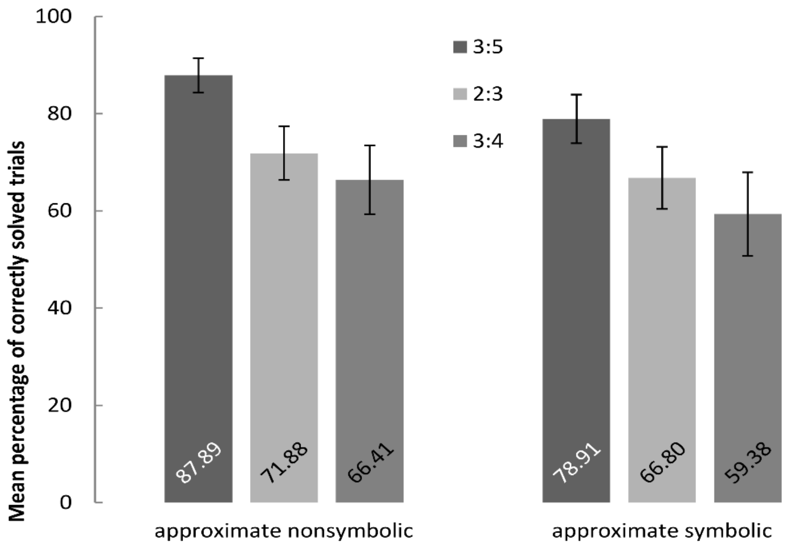

Figure 8. Study 1: Ratio effects in the approximate main tasks, displayed as percentage correctly solved problems for each ratio in both tasks

Second testing of exact symbolic problems. After a first review of the T1 performances, we felt that further improvement was needed beyond the shortening of the tasks (see Chapter 9.1). A problem with our exact symbolic problems was that contrary to the other three main tasks, the operands of the problems were visible simultaneously for several seconds. In both approximate as well as in the exact nonsymbolic task, operands that had to be added/subtracted disappeared after some seconds. This might result in a greater working memory load in these three tasks compared to the exact symbolic one. In line with this assumption was the high accuracy especially in the addition problems of the exact symbolic task (93% correct solutions, see Figure 7). To adapt exact symbolic problems which are highly practiced as soon as school starts (and for many children even earlier due to communication with older siblings, parents, or friends) to the first graders´ proficiency level, we replicated the exact same problems of the T1 task but adjusted the timing of visibility on our slides. Thus, at no time both operands could be seen simultaneously (see Figure 3) – this resembled the representation of the trials in each of the three other tasks. The exact task consisted of the identical four addition and four subtraction problems as described above. Contrary to before, from the onset of each slide, the first operand was visible for 4 seconds and then disappeared. Then the second operand appeared on its position on the other side of the plus/minus sign, also disappearing after 4 seconds. Plus/minus sign and equal sign were visible all the time.

The disappearance of the second operand was followed by a short delay of 2 seconds before the next slide appeared with a symbol that indicated that children were now to give their answer in their booklet.

The retested problems with adjusted working memory demands again revealed very good performance of the participants[6] (*M* = .94, *SD* = .115). Again, accuracy was greater for addition than for subtraction problems. However, this difference was not significant (addition: *M* = .96, *SD* = .113, vs. subtraction: *M* = .92, *SD* = .181, $t(62) = 1.793$, $p = .078$). Performance in the addition and subtraction problems in the retested tasks did not correlate ($r = .182$, $p = .157$), further confirming our impression that addition problems provide the more reliable measure of first graders´ arithmetic abilities.

9.2.2 Intercorrelations of the 2x2 tasks

Are first graders who are good at approximating nonsymbolic quantities also good at approximating symbolically presented numerosities? Or is processing of number symbols a more important and determining aspect in their performance? This would be mirrored in a positive correlation between both symbolic tasks. By calculating the correlations between each of the T1 tasks, we wanted to get an impression if one of the two factors (demanded accuracy vs. format) is crucial in first graders´ numerical processing and produces clear-cut lines between either symbolic and nonsymbolic processing (meaning intercorrelations only within each format, no matter the demanded accuracy), or between exact and approximate processing (meaning intercorrelations only within each level of demanded accuracy, no matter the format).

We found a significant correlation between the performances in the two approximate measures (see Table 4). Children who solved many approximate nonsymbolic problems correctly showed also high accuracy in the approximate symbolic task. Contrary, there was no intercorrelation between the two exact measures at the first date of testing. However, given the methodological problems with the exact symbolic task at T1, we also checked for correlations with the retested exact symbolic measure. There was a significant correlation between performance in the exact nonsymbolic task and the performance in the analogue retested exact symbolic task. Both approximation measures

[6] We assume that the excellent performance despite the increased working memory demands is due to the assessment happening notably later in the school year, and hence the grown experience of children with arithmetic problems of that kind. At T1, the results of this task might have been a little more moderate.

did not correlate with the exact symbolic measures (original or retested). Note, however, that the correlations between both approximation tasks and exact nonsymbolic performance was on the margin of significance. The latter might be a hint on the increasing impact of exact processing which during formal education starts to re-influence ANS representations (Kolkmann et al., 2013), but this finding would need further confirmation in our replication study (Study 2).

Table 4. Study 1: Intercorrelations of performances in the main tasks: Correlation coefficients between the four main tasks as well as the retested exact symbolic measures (each value displays the proportion of correct solutions)

		Approximate	**Exact**		
N = 62		Symbolic	Nonsymbolic	Symbolic	Retested symbolic
Approximate	nonsymbolic	**.378**** *p* = .002	.218 *p* = .089(*)	.107 *p* = .409	.174 *p* = .177
	symbolic		.239(*) *p* = .061	.142 *p* = .270	.086 *p* = .505
Exact	nonsymbolic			.142 *p* = .273	**.390**** *p* = .002
	symbolic				**.327**** *p* = .009

9.2.3 The influence of SES on exact and approximate arithmetic

Due to the need for retesting and the resulting loss of two participating classes, we were only able to conduct our complete testing sequence in a high-SES suburb of Cologne. This resulted in an imbalanced distribution of SES-categories, with an unproportional high number of children with 'good' status (see Table 5). Only five children were categorized as having the lowest status (note: due to the EGP classification system, '1' is the highest, '7' the lowest socioeconomic status), no children at all were classified into the next two categories, while almost half of the participants had the highest SES.

Table 5. Study 1: Distribution of SES-categories in our sample. Note that in the EGP classification system, low values stand for high socioeconomic background and high values for low SES (1 stands for the 'best' status)

	SES category							
	1	**2**	**3**	**4**	**5**	**6**	**7**	**all**
N	27	8	11	13	--	--	5	64

Correlations with SES. We conducted correlation analyses to test if one of the factors 'demanded accuracy' or 'format' is a criterion for a possible impact of SES on performing arithmetic. This would produce clear-cut lines with regard to which performances would intercorrelate and which would not. This in turn might allow tentative conclusions about the character of the relationship between SES and numerical skills. Negative (due to the inverted classification system, see above) correlations between SES and both symbolic tasks would argue in favor of the idea that low-SES children lack behind in symbolic number knowledge. Negative correlations between SES and both exact tasks, on the other hand, are in line with the idea of a mediating influence of language in the connection of SES and math performance.

Correlation analyses revealed that of the three original T1 tasks and the retested exact symbolic task, SES was only correlated with the performance in exact symbolic problems, and – even if only marginally significantly – with performance in the exact nonsymbolic task (see Table 6). The 'worse' children´s socioeconomic background was, the less exact problems were solved correctly, no matter if presented symbolically or nonsymbolically. The fact that the correlation between exact nonsymbolic performance and SES was only on the margin of significance might be due to the general weaker performance in this task than in the exact symbolic one on the one hand, and the mentioned imbalance in the SES distribution in our sample.

Table 6. Study 1: Correlation coefficients between the four main tasks with SES and WM

Task		SES	WM
Approximate	nonsymbolic	-.032 *p* = .801	.093 *p* = .486
	symbolic	.032 *p* = .806	.057 *p* = .673
Exact	nonsymbolic	-.228$^{(*)}$ *p* = .075	.121 *p* = .368
	retested symbolic	**-.321*** *p* = .011	**.282*** *p* = .032
N		62	58

We also checked for a connection between SES and formal math performance as assessed in the DEMAT 1+ at T3. Children had performed at an average level in this standardized measure (mean percentile rank of the overall score was 54, see Table 7). In line with the poor results in the subtraction problems of our tasks, even at the end of the school year, subtraction and chain problems (which also contained subtraction) remained the most difficult ones, according to their lowest percentile rank values.

Table 7. Study 1: Math performance at the end of the school year as measured in the subtests of the DEMAT 1+

Subtest	**mean percentile rank**
Quantities – Numbers	47.58
Number space	54.03
Addition	56.32
Subtraction	46.24
Decomposition and finding the missing number	56.92
Part-whole knowledge	51.24
Chain problems	45.08
Ineqations	52.06
Story problems	63.39
Overall score	54.03
N	62

Although the standardized mathematics measure (overall raw score of the DEMAT 1+) showed a relationship in the predicted direction (negative correlation: the higher a child's SES-category and hence the lower the actual socioeconomic status, the lower was the overall score), this correlation only was on the margin of significance (see Table 8). The DEMAT 1+ performance was furthermore positively correlated with working memory. Children who showed also higher working memory capacity also obtained a higher score in the math test.

Table 8. Study 1: Correlation of DEMAT with SES and WM (coefficients between the overall score of the DEMAT 1+ with SES and Working Memory)

	SES	**WM**
DEMAT 1+	-.219(*) *p* = .087	**.314*** *p* = .015
N	62	60

Finally, we also tested for the relationships between our tasks and children´s WM capacity as a control measure to ensure that intercorrelations did not go back to WM demands. For the WM measure, children were presented with digit strings up to seven numbers (see Table 9). Our measure was the latest trial in which a child remembered *both* strings of the respective trial correctly. On average, children managed to do so up to trial 3 (M = 3.183, SD = .813). Also the distribution of frequencies confirms that most children did not succeed in remembering both strings of a given trial with more than four numbers (see Table 9). The results of the correlation analysis show that only the exact symbolic arithmetic performance correlated with WM (see Table 6). Thus, we conclude that the low accuracy in the exact nonsymbolic task cannot be attributed to possible higher working memory demands.

Table 9. Study 1: Frequency table of children´s WM performance at the end of the school year

Trial	**Number of to-be-remembered digits**	**Latest completely correct trial for n children**
1	2	0
2	3	7
3	4	41
4	5	8
5	6	2
6	7	2
Missing		4
N		64

9.2.4 Prediction of formal math performance

To test if one or both of the approximate measures would serve as a predictor of mathematic skills at the end of the school year, we performed a hierarchical multiple regression analysis. Although we found no correlation between approximate nonsymbolic performance and the DEMAT 1+ measure at the end of the school year (and only a marginally significant one between symbolic approximation and DEMAT 1+), due to the great interest this format receives in the literature, we included both approximate measures in the final regression model (see also Lyons et al., 2014). First we controlled for age and working memory capacity**,** which were both significant predictors and together accounted for 17% of the variance (see Table 10). After that, we put in both approximation measures. We found that not the approximate nonsymbolic performance (as previously found by Gilmore et al., 2010) but the approximate *symbolic* task was a significant predictor, accounting for more or less all the rest of the explained variance in the model (6,1 % of the total of 23,8% explained variance,).

Table 10. Study 1: HMR of approximation on DEMAT: Hierarchical Multiple Regression Analyses of both approximate tasks on standardized math achievement (mean overall raw score of DEMAT 1+, N = 60), after controlling for Age and Working Memory

		Dependent measure: DEMAT 1+ (raw score)					
Step	**Regressor**	**R^2**	**R^2 Change**	***p* of R^2 change**	**Beta**	**T**	***p***
1	Age	.112		**.009**	.334	2.698	**.009**
	Age				.276	2.222	**.030**
2	WM (latest trial with both correct)	.170	.059	**.049**	.249	2.010	**.049**
	Age				.272	2.176	**.034**
	WM (latest trial with both correct)				.243	1.940	.057
3	Approximate non-symbolic	.176	.006	.525	.078	.640	.525
	Age				.289	2.379	**.021**
	WM (latest trial with both correct)				.309	2.582	**.013**
	Approximate non-symbolic				-.032	-.247	.806
4	Approximate symbolic	.238	.061	**.040**	.272	2.103	.040

9.3 Discussion

With our first study, we aimed to resolve the confounding of format and accuracy that we observed in current research. Mostly, approximation processes are only tested in a nonsymbolic format. This ignores one crucial statement of the Triple Code Model (Dehaene, 1992; Dehaene et al., 2003); that all approximation processes are carried out in the analogue magnitude code, no matter in which format a problem is presented to us – symbolic (Arabic numerals or verbal) or nonsymbolic. Exact processing, on the other hand, has rarely been tested in a nonsymbolic format; at least not with the explicit goal to compare children´s performance in this task to that in analogue symbolic problems, or with regard to its susceptibility to environmental influences. Jordan et al. (1992) seemed to be the first who did. They, however, only tested exact arithmetic, not approximate. Later, Mejias and Schiltz (2013) tested all the 2x2 combinations; not with arithmetic problems, but in the field of number representation. Both reported studies were conducted with preschoolers, and to our knowledge, there is no study up to date that tested the interrelations or the predictive capacity of symbolic vs. nonsymbolic approximate arithmetic; or the influence SES might have on each of the assessed skills at the beginning of formal education.

So, the exploration of the field in Study 1 was thought to give first hints how the four combinations of format and demanded accuracy might be interrelated in a sample of first graders with regard to their socioeconomic background. At the beginning of first grade, we tested exact arithmetic problems as well as approximate arithmetic problems each in a nonsymbolic and in a symbolic version. At the end of the school year, we presented the same children with a standardized math assessment (DEMAT 1+) and with a group-adjusted working memory measure.

Intercorrelations at the beginning of first grade. We found that there were interrelations between first graders´ performances in exact tasks, no matter their representational format, as well as between their results in the two approximate tasks. There was no significant correlation between exact and approximate performance at the beginning of first grade, but marginal ones between exact nonsymbolic performance and both approximation tasks. This seems to underline the division between ANS and exact number processing proposed in the Triple Code Model (Dehaene, 1992; Dehaene et al., 2003). At first glance, however, it seems to contradict the results that Mejias and Schiltz (2013) had obtained for significantly younger participants´ number representation. In 4 to 5 year-old preschoolers, they found the only correlation between the four tasks to exist between the exact nonsymbolic and the approximate nonsymbolic measure. They attributed this relationship to common perceptual processes in both tasks. One year later, however, all tasks were intercorrelated with the exception of nonsymbolic exact and symbolic approximate number representation. In our study, with children one more year older, we find neither of these configurations. There is no such connection between the formats as in Mejias and Schiltz´s youngest sample, nor does everything intercorrelate, like in their older participants. Instead, we found significant connections depending on the kinds of processes that are demanded in the tasks. Those processes that are *carried out* in the same system (ANS or one of the symbolic codes) correlate: both approximate measures were intercorrelated, and both exact ones were, too. It is possible that the more complex *arithmetic* account that we administered in our study is responsible for the clear-cut results? Number representation as tested by Mejias and Schiltz (2013) might be more trained and internalized already before starting school, and thus be 'simply an old hat'; arithmetic however may not be as trained and thus the ability to perform arithmetic has not yet been unified across the different systems. However, it has to be noted that there were marginally significant correlations of the nonsymbolic exact measure to both approximate performances, indicating an increasing importance of the ability to exactly process quantities.

Correlations with SES. As it had been reported by others before (Gilmore et al., 2010), we also found that children´s exact symbolic math performance varied according to their socioeconomic status, whereas their performance in approximate nonsymbolic arithmetic did not. Interesting, however, was our extension of these results: *exact* nonsymbolic performance indeed differed according to SES, *approximate* symbolic performance did not. Or, to put it differently: both exact arithmetic tasks (symbolic and nonsymbolic) were influenced by socioeconomic factors, both approximate arithmetic measures were not. This points to an interpretation in line with one of our hypotheses concerning the character of the relationship between SES and math performance. Only exact processing is impaired in low-SES children, no matter in which format the problems are presented to them. According to the TCM, such processing has to be carried out in one of the two language-based symbolic codes. Language has been found to be impaired in low-SES children, too. Thus a mediating role of linguistic abilities might characterize the connection between SES and math performance. A different account frequently proposed in literature argues that symbolic number knowledge is impaired in children receiving less math related support and activities at home. However, this should have resulted in worse performance of low-SES participants in each task containing and demanding the understanding of Arabic numerals. We on the other hand found that symbolic approximate arithmetic did not differ according to SES. Thus we tentatively conclude that a mediating role of linguistic skills could account for the relationship. It has to be kept in mind that a majority of the participants in our sample stemmed from rather high-SES families so that the low SES-categories were underrepresented in our analyses. This makes it on the one hand all the more remarkable that we even detected the reported correlations, but on the other hand the hypothesis should be tested in further research with a more balanced sample.

What approximation measures can predict formal math performance? Contrary to other studies (Gilmore et al., 2010; Pinheiro-Chagas et al., 2014), we found no predictive relationship between children´s performance in a nonsymbolic approximate arithmetic task and their formal math performance at the end of first grade. Instead, we found that relationship for *symbolic* approximate arithmetic.

On the one hand, by that we extended the mentioned results that reported nonsymbolic approximate arithmetic to predict math achievement. From our results, it seems that the ability to approximate arithmetic problems *in general* might be a fundamental basis for the development of formal mathematical skills, and that this ability is not restricted to nonsymbolic problems. This confirmed our expectation of symbolic approximate arithmetic to predict formal math performance at the end of first grade.

However, on the other hand it is curious that we did not observe a predictive relationship between nonsymbolic approximation and formal math skills (but see Lyons et al., 2014, Soltész, Szűcs, & Szűcs, 2010, for similar results). One possible explanation that comes to mind would be a ceiling effect in the nonsymbolic approximate arithmetic task. Because no symbolic number knowledge is required and only the seemingly basic ANS-processing is recruited, children in first grade might simply solve almost each problem correctly, leading to too little variance to produce a measureable effect size (correlation). However, our data show that performance in this task was not better than in the symbolic exact problems (see Table 2); and that also the respective *SD*s were comparable. This cannot account for the missing predictive capacity of this performance for formal math.

As another alternative explanation, one could also speculate that the task, although administered several times before in a similar presentation (Barth et al., 2005; Gilmore et al., 2010; Pinheiro-Chagas et al., 2014), was not valid and did not measure ANS acuity. But this is also unlikely, as we observed the classical signature of approximation, the ratio effect, to a similar degree in both approximation tasks (see Figure 8).

Therefore, we rather speculate that this missing finding might go back to a power problem in our testing. We had to exclude several children from our analyses because the length of our testing had to be adjusted, resulting in a relatively small sample size. Furthermore, in a group setting like that we used, there is naturally more noise in the data than in individual or even computerized testing sessions. However, why should these methodological issues only apply to the nonsymbolic approximate task and not to the symbolic approximate one? One reason might be a generally weaker connection between formal math skills and ANS acuity measured nonsymbolically, than with ANS acuity measured symbolically. In other research it has been found that this is exactly the case when it comes to *comparison* tasks: Nonsymbolic comparison tasks are a weaker or more unstable predictor of math performance than symbolic comparison (de Smedt et al., 2013, Schneider et al., 2016). For approximate *arithmetic*, however, only nonsymbolic tasks have been tested so far (and to our knowledge, a relationship with formal math performance has been robustly found). So it cannot be ruled out that symbolic approximate arithmetic also shows a stronger relationship with formal math performance than nonsymbolic approximate arithmetic. This does not necessarily mean that the latter has no connection at all to our DEMAT 1+ measure. But it is possible that the effect is smaller so that we would have to test more participants to find a statistically significant correlation for nonsymbolic approximate arithmetic (see also Chen & Li, 2014).

We checked this explanation with a post-hoc power analysis using the free software G*Power (Faul, Erdfelder, Lang, & Buchner, 2007). The test revealed that in fact with our sample size of 60 children included in this analysis (not all children were present at DEMAT 1+ assessment), there was only a 33% chance to detect a correlation of this size. Thus, we assume that a relationship between nonsymbolic approximate arithmetic performance and DEMAT 1+ in this study was not confirmed due to the small sample size. We expect to observe it in the adjusted replication (Study 2).

To sum it up, with our first study, we have provided first evidence that approximate arithmetic, independently of the format it is presented in, is robust against socioeconomic influences at the beginning of first grade. Furthermore, we also reported some indication concerning the character of the relationship between SES and exact math performance. Given the validity of the Triple Code Model, decelerated linguistic development should lead to poorer performance in processes that are carried out in one of the two symbolic codes. This applies for exact processing. So, we suspected that if differences in linguistic abilities between low- and high-SES children mediate the impact that socioeconomic factors have on numerical abilities, performances in both exact tasks should vary with SES, while the approximate measures that could be processed in the analogue magnitude code on the contrary should not be affected. This is the pattern that we found (although it has to be noted that the correlation of nonsymbolic exact arithmetic with SES was only marginally significant). So, according to our findings, we tentatively hypothesize that a mediating role of linguistic abilities in this connection is plausible. However, these results have to be treated with caution, because of the small sample size and additionally the disproportionately high number of high-SES children in it. Results might look differently when testing a more diverse sample. The proposed possible mediating role of linguistic abilities should of course be tested directly if our pattern of SES affecting only exact skills holds true also in our second study. Furthermore, we extended current research looking for the connections between ANS-bound processing and later formal math skills. We found that symbolic approximate arithmetic measured at the beginning of first grade was predicting a significant proportion of variance in formal math performance assessed at the end of first grade. The missing correlation between the analogue but nonsymbolically presented approximate arithmetic task and formal math performance probably goes back to a generally weaker connection and the resulting reduced power in our small sample. We thus aim to test this assumption in Study 2 with more participants and improved measures.

10 Study 2

Study 2 was for the most part a replication of Study 1, but with some adjustments and improvements. This time, the four main tasks (nonsymbolic exact, nonsymbolic approximate, symbolic exact, symbolic approximate) included only addition problems, because in Study 1 we had found a significant decrease in performance from addition to subtraction problems. Furthermore, there was no correlation between addition and subtraction performance in three of the four main tasks. We reduced the number of problems to ten problems per task, because in Study 1 teacher feedback was given that our testing was too extensive. To measure a more direct relationship between working memory and performance in our main tasks, the adapted working memory test was shifted to the beginning of the school year. Furthermore, three additional measures were assessed. They tackle questions of the mediating function of numerical ordering, the domain specificity of the relationships between performance in our tasks and formal math performance, and the effect of canonical vs. noncanonical representation of arithmetic tasks. The background for taking up these measures are described in the following, description of the tasks themselves follows in Chapter 10.1.

Numerical ordering ability (NO). NO refers to the knowledge which number comes before/after another number (Attout & Noël, 2014, see also Chapter 6.2) and was for example measured by Lyons and Beilock (2011) by letting participants judge if a string of numbers was in ascending order or if the actual ordinal positions of the numbers were interchanged. NO had been found to mediate the relationship between different ANS measures in part or completely, in different age groups and for symbolic as well as nonsymbolic tasks. Lyons and Beilock (2011) found in their adult sample that the relationship between the ANS measure (nonsymbolic comparison) and exact arithmetic was completely mediated by numerical ordering ability. A similar correlation was reported by Attout and Noël (2014) for second graders (from the second half of first grade), but not found for kindergarten children (see also VanMarle et al., 2014). However, NO has not yet been tested regarding its mediating capacity for approximate arithmetic tasks and we have reason to believe that approximate arithmetic should not only be connected with formal math performance via NO (see Chapter 6.2).

Reading Comprehension. At the end of the school year, besides the standardized test of math achievement (DEMAT 1+), we also administered a standardized measure of children´s reading comprehension (ELFE 1-6). Gilmore et al. (2010) found that reading ability of first graders correlated with their general math performance at the end of the school year, but not with their performance in non-symbolic approximate problems. To also check for the domain specificity of the to-be-found interrelations, we thus include reading comprehension as measured by the ELFE 1-6 as a control variable.

Canonical and noncanonical format. Another added question taken up in this study concerned the canonical vs. noncanonical representation of the trials in the approximate tasks. Canonical means that the trial fits the typical spatial organization of mathematical problems: The mathematical problem, in this case the operation of the approximate arithmetic trial, is presented on the left side of the screen, whereas the 'result' (the quantity to compare the approximate result of the operation to) is on the right side. Less typical and therefore noncanonical representation is given when the operation 'happens' on the right side and the result is displayed on the left. McNeil et al. (2011) had conducted a study with a similar aim like our studies (exploring the environmental influence on approximate arithmetic). They presented children with a symbolic as well as a non-symbolic approximate addition task and found that, in both tasks, kindergartners with high socioeconomic background performed better than children with middle socioeconomic background. These results are in contrast to those of Gilmore et al. (2010) or of our results obtained in Study 1. Neither in Gilmore´s nor in our study had the approximate arithmetic measures been found to be influenced by SES. McNeil et al. (2011) suggested that if environmental factors (which also comprise formal education) are responsible for possible differences between children of varying socioeconomic backgrounds, this should also be mirrored in differential outcomes in trials in a canonical vs. noncanonical representation. The more practiced mathematical skills are, the larger the difference in favor of canonical trials should be. The authors found that only children with high socioeconomic status displayed such an effect, whereas there was no difference between the results for both representation forms in the middle status sample (despite their general lower performance level compared to the high status group). They suggest that children of higher SES have been exposed to mathematical content at home to a higher degree than middle-SES children.

10.1 Method

Participants. In Study 2, participants were 144 first graders (72 girls). They were recruited from six classes in four elementary schools in the metropolitan area of Cologne, Germany. We only included participants in our analyses who had been present on each of the four testing sessions. Due to this condition, twenty children were excluded. Also, four children had to be taken out of analysis because the experimenter observed that these children did not follow instructions (e.g. not starting with solving the problems when they were supposed to, continuing working on the previous problem when time was up, trying to count the quantities of the nonsymbolic approximation task). Thus, 120 subjects (59 girls) with a mean age of 6 years and 8 months (*SD* = 4.5 months) remained in our sample.

Two of the schools were located in a low- to middle-class community, two schools in a middle- to upper-class community. In addition to this frequently used local definition of children´s SES, we gave out a questionnaire about the parents´ education and current profession. Again, we intended to categorize the individual SES according to the EGP classification system (Erikson et al., 1979), to analyze the influence of socioeconomic status on a more fine-grained level. However, in the course of testing it became clear that not enough parents returned the questionnaire. So we followed common practice and operationalized socioeconomic status by the area the respective schools were located in. Three of the classes were from high-SES suburbs in Cologne (Bayenthal and Müngersdorf), the other three from low-SES areas (Kalk and Bickendorf). Note that consequently in Study 1 and 2, children´s SES has been assessed differently and possible differing results have to be treated with caution.

Material. In the following, adjustments and changes in the reused tasks are described. After that, the tasks that were used additionally to the ones already introduced for Study 1 are presented.

Exact tasks

The symbolic and nonsymbolic exact tasks again included identical problems. Based on the findings of our pilot study (Study 1), we made some adjustments to the tasks. First of all, we decided to leave out subtraction problems. In Study 1, accuracy in subtraction problems was rather low. Furthermore, missing correlations between addition and subtraction performance in each but the exact nonsymbolic task suggested that at that early stage of schooling, subtraction was not a suitable measure of calculation proficiency in our sample. Because of this assumption, and to further increase reliability of our measure, this time we presented the first graders in both exact tasks with *ten* two-addend addition problems instead of four addition and four subtraction problems. The smallest addend was 3 (this addend only occurred once in the first problem), the largest 9, with results ranging from 8 to 15. Addends within the subitizing range were excluded and the position of the larger addend was counterbalanced across the problems. The problems are depicted in Appendix C.

Symbolic exact task

The exact task consisted of the addition problems as described above. As in Study 1, problems were presented in Arabic numerals via a Microsoft PowerPoint presentation, each problem on its own slide. But this time, from the onset of each slide, the first addend was visible only for 4 seconds and then disappeared. Then the second addend appeared on its position on the other side of the plus sign, also disappearing after 4 seconds. Plus sign and equal sign were visible all the time. The disappearance of the second addend was followed by a short delay of 2 seconds before the next slide appeared with a symbol that indicated that children were now to give their answer in their booklet.

Nonsymbolic exact task

Problems were the same as in the symbolic exact task and the procedure was the same as in Study 1, only with adjusted onset times. Because the task had been found to be rather difficult for participants in Study 1, we slightly slowed down the sequence of each trial. Each dot took 1 second (instead of 0.5) for its path into the container.

Approximate tasks

We administered both the nonsymbolic and the symbolic approximate addition task. Like in Study 1, both tasks contained the same ten trials, but this time only addition problems were included. In each trial, children had to compare the result of a two-addend addition problem to a reference quantity to indicate if the result or the reference quantity was the larger one. Procedure was the same as in approximate addition trials of Study 1. At the end of each trial, participants had to indicate which of the two children had more candies/marbles. In half of the trials, 'Tim' was the right answer, 'Lisa' in the other half. In six trials, the larger number was associated with the addition problem, in four trials with the reference quantity.

In the ten addition problems that formed the ten trials (together with their respective reference quantities), the smallest addend was 11, the largest 35, with results ranging from 24 to 60. The ten reference quantities ranged from 18 to 65. We realized three different ratios of the resulting quantities that had to be compared. Four comparisons had to be made in the easiest ratio 3:5, three comparisons in the ratio 2:3 and three in the ratio 3:4. In each task, half of the trials presented the problems in a canonical manner (operation on the left side), the other half presented them in a noncanonical form (operation on the right side). Both tasks were again presented via a Microsoft PowerPoint Presentation. The problems are depicted in Appendix D.

Symbolic approximate task

Except for the differing operands and the absence of subtraction problems, the symbolic approximate task remained unchanged from Study 1.

Nonsymbolic approximate task

Besides the differing operands and the absence of subtraction problems, we made some further perceptual adjustments in the nonsymbolic approximate task to ensure that participants made their decisions based on quantity. Marbles varied in individual size, that is, the complete array of an addend could either consist of small marbles (0.18 cm in Microsoft PowerPoint) *or* large marbles (0.25 cm). Also the complete stimulus area (area on which the marbles were sprinkled) could be in one of two sizes (small 3 x 5 cm, or large 3.5 x 6 cm in Microsoft PowerPoint). Thus, dot size, array size and density varied between the quantities. Following Gilmore et al. (2010), to prevent children from basing their answers on continuous problem variables, dot size, array size and density of the marble pattern were positively correlated with the larger quantity in half of the trials, and negatively correlated with the larger quantity on the other half (see also Barth et al., 2005).

Working Memory (WM) Task

The WM Task remained unchanged (see method section of Study 1) except for the shortening by one trial. So, the longest sequence children had to remember contained six digits.

Numerical ordering ability (NO) task

Numerical ordering has been found to be a possible mediator of the relationship between nonsymbolic approximate measures and mental arithmetic performance in adults (Lyons & Beilock, 2011). To assess this relationship also in learners without much prior knowledge, we assessed this ability by letting participants work through three pages of a test booklet. Each page contained ten strings of numbers. The task for the children was to decide if the numbers of a string were in the right order ('becoming larger' from front to end) or not. They were to mark the icon of a smiling face if the order was right, the icon of a sad face if the order was wrong. On the first page, one string always consisted of four numbers that ranged from 2 to 9. On the second page, each string only contained three instead of four numbers, but this time numbers were higher, thus posing more difficulty to children: The three numbers of a string on page 2 always resembled a string on page 1 (always leaving out one number per string) but ranged in the number space from 10 to 20. For example, one string on page 1 contained the numbers '3 4 7 9', and the corresponding problem on page 2 was '13 14 17'. So, the absolute differences between the respective numbers were equal, although the numbers on page 2 were larger. On page 3, the first numbers of the three-number strings were identical to the first numbers on page 2, but the ratio of the first to the second and the second to the third number resembled the ratios realized on page 1, resulting in numbers up to 54. So we can disentangle if children’s numerical ordering ability depended on the ratios between the different numbers (like found in number comparison basing on the ANS) or on the absolute differences between the numbers (like found for exact addition in the problem size effect). Problems are depicted in Appendix E.

Maths Performance: DEMAT 1+

See method section of Study 1.

Reading Comprehension: ELFE 1-6

The ELFE 1-6 (Lenhard & Schneider, 2006; Ein Leseverständnistest für Erst- bis Sechstklässler; A reading comprehension test for first- to sixth graders) is a standardized test designed for the German primary school system (and the first two grades of secondary education which belong to elementary school in some states). The subtests of the paper-pencil version are 1) word comprehension (72 items, choosing the correct word for an image among four graphically and phonemically similar alternatives); 2) sentence comprehension (28 items, choosing the right among five alternative parts of a sentence that fits best into a given sentence) and 3) text understanding (20 items, children read a short text and a question referring to this text. They then choose the right answer among four alternatives. Some questions demand only the finding of isolated information, some request anaphoric references over several sentences, and others demand building inferences to be answered correctly, thus constituting varying degrees of difficulty). The retest reliability of the paper-pencil version is r = .92.

Procedure. Due to our experiences in Study 1, we split our testing sessions for each measurement point (see Table 11). So we had two sessions on two separate days (T1a and T1b) at the beginning (first quarter) of the school year in each class. At the end of the term (T2) we again tested children on two separate days (T2a and T2b). Each testing session took 45 to 60 minutes. In the first session (T1a), we tested children's working memory, before the exact symbolic and then the approximate symbolic task were administered. In the second testing session (T1b), the exact nonsymbolic and after that the approximate nonsymbolic task were given. At the end of this session, we tested numerical ordering ability. We decided to always administer the exact measure before the approximate one because some results suggest that there might be a unidirectional influence of approximation on exact calculation (Park & Brannon, 2013; Hyde et al., 2014). At the end of the school year, in the third testing session (T2a) the standardized math performance measure, the DEMAT 1+, was given. In the fourth and last assessment time, we tested children's reading comprehension with the ELFE 1-6.

Table 11. Study 2: Timeline of testing. The 2x2 tasks printed in italics are referred to as the 'main tasks'

	T1: beginning of first grade		**T2:** end of term	
	T1a	T1b	T2a	T2b
administered measures	working memory (WM)	*exact nonsymbolic task*	DEMAT 1+	ELFE 1-6
	exact symbolic task	*approximate nonsymbolic task*		
	approximate symbolic task	numerical ordering (NO)		

10.2 Results

Analogous to Study 1 and along our research questions and hypotheses, the first part of the result section will be about the performance of our participants in the T1 tasks and the interrelations of the four main tasks. After that, the influence of children´s SES on these performances as well as on the T2 measures (DEMAT 1+ and ELFE 1-6) will be reported. An additional analysis follows in which we reexamine the intercorrelations separately for children with low vs. with high SES. Lastly, we will present the results of the hierarchical multiple regression analysis of what predicts formal math performance best, this also including the question of the possible mediating influence of NO.

10.2.1 Performance and intercorrelations in the 2x2 tasks

Intercorrelations between different performances are important indicators to analyze which aspects of mathematical processing determine children´s performance at a given point in time. Before reporting the intercorrelations, we will give some descriptive statistics of the performances in the T1 tasks.

In the approximation tasks, children again solved almost every problem (see Table 12). In the both the nonsymbolic and the symbolic task there were five children who missed at least one problem. The same was the case for the exact nonsymbolic problems (six children who did not solve each problem), whereas there was a substantial number of 21 children who left out some of the trials in the exact symbolic task. Consequently, the *SD* in the number of solved exact symbolic problems was more than three times higher than in the other three main tasks (see Table 12). To ensure comparability between the exact tasks, we thus decided to use the percentage of correctly solved problems relative to the number of solved problems as our measures of performance for all four main tasks.

Table 12. Study 2: Performances in the main tasks of T1: approximate nonsymbolic and symbolic; exact nonsymbolic and symbolic. Depicted are number of solved problems and percentage correct of solved problems

		Number of solved problems	Percentage correct
approximate	nonsymbolic	9.933 *(.310)*	72.097 *(16.418)*
	symbolic	9.917 *(.46)*	68.424 *(16.734)*
exact	nonsymbolic	9.908 *(.467)*	51.322 *(28.683)*
	symbolic	8.942 *(1.519)*	55.642 *(30.799)*

As in Study 1, children again performed above chance (more than 50% correct solutions) in each of the four main tasks (see Table 12); and as had to be expected because of the Study 1 results, there was a significant advantage in nonsymbolic approximation compared to symbolic approximation (M = 72.1, SD = 16.42 vs. M = 68.42, SD = 16.73, $t(119) = 1.97$, $p_{1\text{-tailed}} = .025$, $d = .36$). The difference between the performances in the exact tasks also had the same direction as in Study 1, but just missed significance (nonsymbolic M = 51.32, SD = 28.68 vs. symbolic M = 55.64, SD = 30.8, $t(119) = 1.65$, $p_{1\text{-tailed}} = .051$).

Signature of the ANS in the approximation tasks. Both approximate tasks produced the typical ratio effect, that is, the larger the numerical distance between the to-be-compared quantities gets, the less correct solutions are found (see Figure 9). The one-way repeated-measurement ANOVA over percentage correct solutions with the three ratio levels as repeated factor showed a significant effect for the nonsymbolic task ($F(2, 238) = 4.197$, $p = .016$, $\eta^2 = .034$). Also the linear contrast was significant ($F[1,119] = 8.334$, $p = .005$, $d = .34$). The analogous ANOVA for the symbolic task was significant, too ($F(2, 238) = 12.203$, $p < .001$, $\eta^2 = .093$), as was the according linear contrast ($F[1,119] = 18.859$, $p < .001$, $d = .48$). We thus conclude that also the participants of our Study 2 actually approximated the results of our approximation tasks as it was intended.

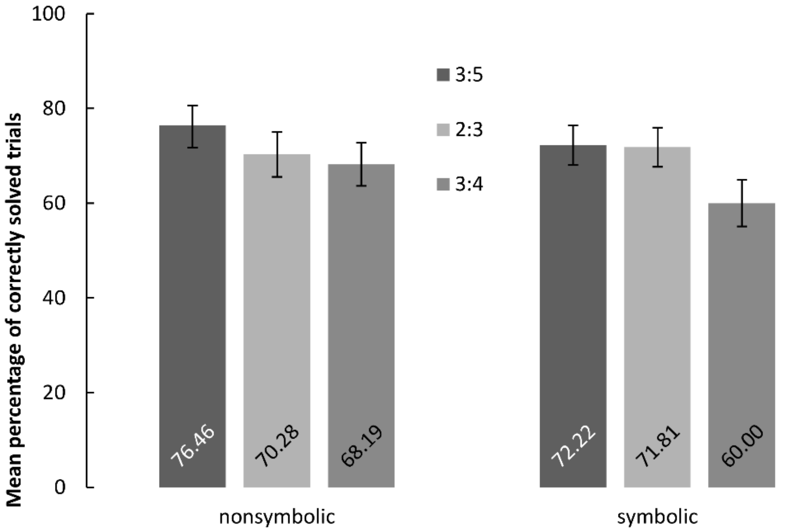

Figure 9. Study 2: Ratio effect in the approximate main tasks. Percentage correct solutions/solved problems separately for each ratio in each approximate task

Canonical vs. noncanonical representation of approximate problems. If approximation is already strongly connected to the learned procedures of formal calculation in first graders, they should show better performance when solving problems that follow the representational sequence they are taught in school. This canonical sequence goes from the left to the right. In our approximation tasks canonical trials would be those with the operation on the left side, and the possible result – here, the quantity to compare the assumed result with – on the right side of the screen (McNeil et al., 2011). We calculated a 2 (SES: high vs. low) x 2 (canonicity: canonical vs. noncanonical sequence) mixed-design ANOVA with percentage of correctly solved (non-) canonical problems as dependent variable separately for the symbolic approximate and the nonsymbolic approximate task. We were interested in a possible main effect of canonicity in each task, as well as in a possible interaction of canonicity with SES. McNeil et al. (2011) had found that preschool children with a higher SES showed poorer performance in noncanonically presented problems compared to canonical ones, while same age children with lower SES did not produce that pattern. They had hypothesized that after starting elementary school, this interaction could either become stronger because children are more exposed to formal arithmetic, or alternatively the interaction might disappear. The latter might be the case because children also encounter the demand to produce an exact result to canonical arithmetic problems at school. This in turn would diminish the advantage of canonical problems that should only be approximated. The performances in the canonical and noncanonical problems of each task are depicted in Table 13 separately for high- and low-SES participants.

Table 13. Study 2: Accuracy in canonical and noncanonical trials of the approximation task. Performance (proportion correct) separately for both SES groups

	Low SES (N = 56)		High SES (N = 64)	
Approximation Tasks	**Canonical**	**Noncanonical**	**Canonical**	**Noncanonical**
Nonsymbolic	.665 (.207)	.642 (.205)	.797 (.187)	.763 (.181)
Symbolic	.669 (.219)	.698 (.219)	.683 (.156)	.691 (.208)

Neither in the nonsymbolic nor in the symbolic approximate task did we find a main effect of canonicity (nonsymbolic: $F(1, 118)=1.661$, $p = .2$; symbolic: $F(1,118) = .869$, $p = .353$), nor a significant interaction of SES and canonicity ($F(1, 118) = .071$, $p = .791$ vs. $F(1, 118) = .296$, $p = .587$).

Performance in the two control measures WM and NO assessed at T1. In this study, we assessed WM capacity and NO, both at the beginning of our study to secure the causal direction of to-be-found relationships between the measures and formal math performance. As described above, in the NO task the problems were presented on three pages with increasing difficulty, so the three pages are analyzed separately. A look at the data in Table 14 indicates that from page to page, the number of solved problems increased ($F(2, 238) = 26.552$, $p < .001$, $\eta^2 =.182$) whereas the percentage of correct solutions decreased ($F(2, 232) = 37.86$, $p < .001$, $\eta^2 =.246$). Still, general accuracy was higher than 50% on all three pages, thus exceeding chance level of the dichotomous answering possibility.

In line with this finding of increasing solution frequency and decreasing accuracy, also d´ (the dimensionless sensitivity index from Signal Detection Theory SDT) decreased from page to page ($F(2, 210) = 37.321$, $p < .001$, $\eta^2 =.262$). In the SDT, d' describes the ability to discriminate between signal-present trials (here: the correct order) and signal-absent trials (wrong order). That is, the index d' reflects the relation between hits (number strings in correct order that have been marked by the children with a smiling face) and false alarms (FA, number strings in wrong order that have been marked with the smiling face; d'= z(Hits) – z(FAs)). The response bias c here measures participants' general tendency to mark problems as correctly ordered (c= -0.5 x (z(Hits) + z (FAs))).

The response bias c was negative on each page, indicating a rather liberal answering behavior (see Table 14). The finding of the lowest sensitivity on page 3 indicates that judging the ordinality in this task does not only depend on the relations between the numbers of a sequence and hence is qualitatively different from approximation. Page 3 presents strings of numbers in a higher number range, but mirrors the ratios of page 1. Simply recruiting the ANS to perform multiple comparisons within a string should thus lead to comparable performance on this page, which clearly is not the case.

However, the significant trade-off between attended sequences and correct answers might also be a sign that the high numbers on page 2 and 3 were too difficult to process for first graders (although, note that operands were not much higher than in the symbolic approximation task (see Appendix D) in which children performed above chance level). We thus decided not to analyze the three pages separately any further in the following. Because performance (d´) was highly correlated between the three pages (r_{12} = .494, p < .001; r_{13} = .377, p < .001; r_{23} = .507, p < .001), in the following analyses we use overall d´ as our NO measure.

Table 14. Study 2: Performance in NO: Performance in the Numerical Ordering task in terms of solved problems, percentage correct of answered problems, and SDT measures for the 120 participants (SD in brackets)

	Numerical Ordering Task			
	Page 1	Page 2	Page 3	All pages
Answered problems	7.45 (3.167)	8.375 (2.685)	9.142 (1.812)	24.967 (6.49)
Percentage correct	72.543 (26.506)	61.949 (24.607)	50.664 (25.269)	60.28 (21.084)
Hits	77.227 (32.426)	73.72 (32.188)	60.868 (34.821)	69.014 (27.192)
Correct rejections	66.101 (35.315)	51.387 (30.731)	41.764 (30.316)	51.38 (26.669)
Sensitivity d´	2.509 (2.951)	1.439 (2.119)	.0995 (2.568)	.791 (1.827)
Answering bias c	-.339 (1.242)	-1.043 (.998)	-.618 (1.103)	-.455 (.826)

In the WM task, children were presented with digit strings up to six numbers. We shortened the WM task from Study 1 as there had been very few children who managed to remember the seven-digit string (see Table 9) and because in Study 2, we tested WM capacity at the beginning of term (instead of at the end as in Study 1).

On average, children succeeded in remembering *both* strings of a given trial correctly only up to trial number 2 (M = 1.85, SD = .932). This trial included three digits. The difference in performance to their peers in Study 1 is shown in Table 15 – there are 12 children in our sample who did not manage to remember the two-digit strings of the first trial two times in a row.

Table 15. Study 2: Frequency table of children´s WM capacity at T1

Trial	Number of to-be-remembered digits	Latest completely correct trial for n children
1	2	24
2	3	56
3	4	26
4	5	2
5	6	0
None		12
Missing		---
N		120

Intercorrelations of the measures. The intercorrelations between the four main tasks are less clear-cut than in Study 1. Each of them correlated significantly with each of the others (see upper panel of Table 16). The correlation between both exact measures (r = .539, see Table 16) resulted in the highest coefficient and indicated a significantly stronger relationship between the exact measures than between the exact symbolic performance and both approximate measures (with symbolic approximate: r = .328, $t_{difference}(117)$ = 2.511, $p_{(1\text{-}tailed)}$ = .06, q = .26; with nonsymbolic approximate: r = .306, $t_{difference}(117)$ = 2.769, $p_{(1\text{-}tailed)}$ = .003, q = .29). However, it did not differ significantly from the correlation coefficients between exact nonsymbolic and approximate nonsymbolic performance (r = .409, $t_{difference}(117)$ = 1.487, $p_{(1\text{-}tailed)}$ = .069) or – most surprisingly – between exact nonsymbolic and approximate symbolic abilities. That is, exact nonsymbolic arithmetic was to a comparable degree correlated with symbolic approximation (r = .406, $t_{difference}(117) = 1.538$, $p_{(1\text{-}tailed)}$ = .063) as with exact symbolic processing. This is curious because in these two tasks neither the operation is the same (exact vs. approximate processing) nor the format (nonsymbolic vs. symbolic).

Table 16. Study 2: Correlation coefficients between the various measures of T1

		Approximate	**Exact**		**WM**	**NO (d´)**
N = 120		symbolic	nonsymbolic	symbolic		
Approxi-mate	nonsymbolic	**.244**** *p*=.007	**.409***** *p* < .001	**.306**** *p*=.001	.163$^{(*)}$ *p*=.075	**.195*** *p*=.033
	symbolic		**.406***** *p* < .001	**.328***** *p* < .001	.130 *p*=.158	**.296**** *p*=.001
Exact	nonsymbolic			**.539***** *p* < .001	**.381***** *p* < .001	**.413***** *p* < .001
	symbolic				**.288**** *p*=.001	**.483***** *p* < .001
WM						**.374***** *p* < .001

Partial correlations controlling for exact nonsymbolic performance

		Approximate symbolic	**Exact** nonsymbolic	**Exact** symbolic	**WM**	**NO (d´)**
Approxi-mate	nonsymbolic	.093 *p*=.155		.111 *p*=.113	.008 *p*=.463	.031 *p*=.367
	symbolic			.141$^{(*)}$ *p*=.061	-.029 *p*=.376	**.154*** *p*=.046
Exact	symbolic					**.339***** *p* < .001
WM						**.257**** *p*=.002

The described pattern shows that among the main tasks, the performance in the exact nonsymbolic task was correlated the strongest with each of the other measures. As interrelations between exact nonsymbolic processing and approximation had been indicated already in Study 1, we suspect that this task measures a superordinate numerical skill, the dominance of which is responsible for the strong interrelations between all of the main tasks in this study. As a means to anyhow get a look at the possible differential importance of format or demanded accuracy (which we had detected in Study 1), we thus decided to perform the correlation analyses with the influence of exact nonsymbolic performance kept constant. These analyses of partial correlations (see lower panel of Table 16) revealed that after taking out the common variance with nonsymbolic exact performance, only the correlations between tasks including Arabic number symbols remained significant (approximate symbolic arithmetic, exact symbolic arithmetic, and WM with NO; still marginally significant: approximate symbolic arithmetic with exact symbolic arithmetic). Together, the results indicate that the ability to exactly manipulate quantities is an important factor in first graders´ numerical competency independently of symbolic number knowledge: The symbolic exact performance did not intercorrelate with the other measures to an equal degree as the nonsymbolic one, and furthermore the correlations between symbolically presented tasks existed independently of nonsymbolic exact performance.

Surprisingly, after partialling out the influence of nonsymbolic exact performance, also the correlation between both approximate measures was not significant anymore. This might be a hint that at this developmental stage, the obviously dominant ability to manipulate quantities in an exact manner already reinfluences the fundamental ANS (see for example Kolkman et al., 2013; or Mejias & Schiltz, 2013).

The correlations of the main task performances with NO and WM are less surprising (see upper panel of Table 16): As would be expected for a possible mediator, NO correlated with each other measure. However, correlations were significantly stronger with both exact measures than with the approximate nonsymbolic measure (exact symbolic with NO vs. approximate nonsymbolic with NO: $t_{difference}(117) = 3.024$, $p_{(1\text{-}tailed)} = .001$, $q = .33$; exact nonsymbolic with NO vs. approximate nonsymbolic with NO: $t_{difference}(117) = 2.382$, $p_{(1\text{-}tailed)} = .009$, $q = .24$). Exact symbolic performance also exceeded the approximate symbolic measure in its correlation to NO (exact symbolic vs. approximate symbolic: $t_{difference}(117) = 2.020$, $p_{(1\text{-}tailed)} = .022$, $q = .22$), but not the exact nonsymbolic performance ($t_{difference}(117) = 0.92$, $p_{(1\text{-}tailed)} = .179$). This is in line with the possibility that NO mediates the relationship between ANS and formal math performance respectively exact arithmetic (Lyons & Beilock, 2011).

Working Memory was correlated with performance in both exact tasks, but not (respectively only marginally) with the approximate ones, resembling results obtained in adults by Kalaman and LeFevre (2007) who found that exact arithmetic did depend on WM significantly stronger than approximate arithmetic (see also Attout et al., 2014).

10.2.2 The influence of SES on exact and approximate arithmetic

As explained above, in this study we changed the SES measure from the EGP classification to an operationalization by the local area of the schools in which we did our testing. This resulted in a dichotomous categorization of low and high SES. In the first group there were 56 children, in the latter 64 children.

Table 17. Study 2: Correlation between all measures and SES: Correlation coefficients and significance levels between the performance in the T1 tasks (percentage correct of solved problems), the control and the T2 measures and socioeconomic status (operationalized as location of the school)

Task		**SES**	**N**
Approximate	nonsymbolic	**.384 ***** $p < .001$	120
	symbolic	.017 $p = .858$	120
Exact	nonsymbolic	**.243**** $p = .007$	120
	Symbolic	**.382***** $p < .001$	120
Control	WM	**.407***** $p < .001$	120
	NO (d´)	**.239**** $p = .009$	120
T2	DEMAT 1+	**.45***** $p < .001$	109
	ELFE 1-6	**.567***** $p < .001$	105

As expected, both standardized end-of-term measures correlated significantly and positively[7] with SES (see Table 17). That is, children with higher SES obtained more points in the standardized math test as well as in the reading comprehension test. As in Study 1, also both exact measures showed a significant relationship with children´s socioeconomic status. Again, the performance in the approximate symbolic task was not correlated with children's socioeconomic background. Surprising, however, is the high and unexpected correlation of the approximate nonsymbolic task with SES – this is the measure usually found not to be influenced (Gilmore et al., 2010; but see McNeil et al., 2011).

Looking at the data as plotted in Figure 10 (and see also Table 18), however, draws a clearer picture. It is visible that in the subgroup of low-SES children it did not matter in which format a problem was presented. They showed comparable performance in the symbolic like in the nonsymbolic approximation problems. Same for exact arithmetic: Low-SES children showed equal (poor) performance in both of the exact tasks.

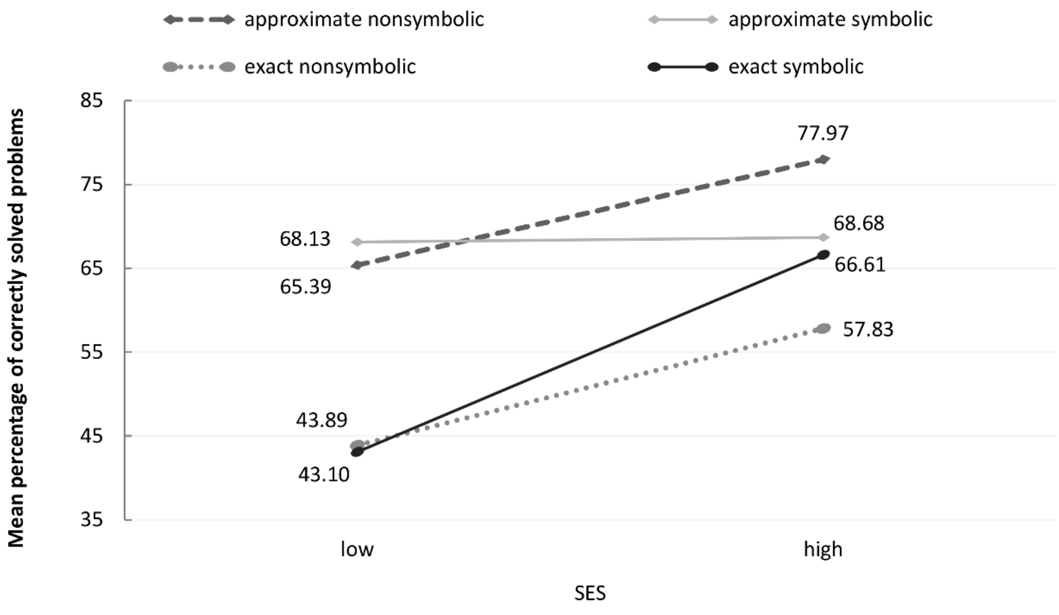

Figure 10. Study 2: Performances in the four main tasks for low- vs. high-SES participants

[7] Note that in this study, with area of school we have a different SES measure than in Study 1. The participants with lower SES were coded with '0', the high-SES children with '1', so that – contrary to Study 1 – a positive SES value actually stands for higher SES.

However, this looks differently in the subsample of high-SES children. Within each of the representational codes to perform the operation in (that is, analogue magnitude vs. one of the symbolic codes; or simply: approximate vs. exact), high-SES participants differed in their performance depending on the format in which the problems were presented. In approximate problems, they performed better when the problems were displayed nonsymbolically. In the exact tasks, on the other hand, the symbolic representation seemed to be easier for them. We checked the differences by performing two separate 2 x 2 mixed-design ANOVAs for the percentage of correctly answered trials with SES as the between-subject and format (nonsymbolic vs. symbolic) as the within-subject factor for the approximate tasks on the one hand, and the exact tasks on the other.

For the approximate tasks, analysis revealed a main effect of SES: Children with high SES showed significantly better performance ($F[1, 118] = 7.977$, $p = .006$, $\eta_p^2 = .06$), while there was only a marginally significant main effect of format ($F[1, 118] = 3.346$, $p = .07$, $\eta_p^2 = .03$). However, there was also a significant interaction ($F[1, 118] = 11.305$, $p = .001$, $\eta_p^2 = .09$). Bonferroni-adjusted post-hoc t-Tests with a conducted alpha-level of $p = .025$ revealed that only in the high SES children, the performance differed between nonsymbolic and symbolic problems with significantly more nonsymbolic problems answered correctly ($t(63) = 4.289$, $p < .001$, $d = .55$), while this was not the case in the low SES participants ($t(55) = -.94$, $p = .351$).

For the exact tasks, the ANOVA also showed a main effect of SES ($F[1, 118] = 17.514$, $p < .001$, $\eta_p^2 = .13$), but no main effect of format. The interaction term this time was only on the margin of significance ($F[1, 118] = 3.404$, $p = .068$, $\eta_p^2 = .03$). Again checking via Bonferroni-adjusted post-hoc t-Tests showed that there was no significant difference between the performance in both formats in low SES children ($t(55) = -.195$, $p = .846$), but again between both performances in high SES participants, with significantly more symbolic problems solved correctly ($t(63) = 2.627$, $p = .011$, $d = .39$).

Table 18. Study 2: Performances in the four main tasks for low- vs. high-SES participants (SD in brackets)

		SES		
		low	**high**	**all**
appro-ximate	nonsymbolic	65.387 *(16.355)*	77.969 *(14.162)*	72.097 *(16.418)*
	symbolic	68.130 *(18.740)*	68.681 *(14.906)*	68.424 *(16.734)*
exact	nonsymbolic	43.889 *(30.192)*	57.826 *(25.809)*	51.322 *(28.683)*
	symbolic	43.104 *(30.660)*	66.611 *(26.628)*	55.641 *(30.799)*
N		56	64	120

What can explain the differential advantages of format in this subgroup? How can it be that in approximation, the nonsymbolic format is easier to solve for high-SES children, when in exact arithmetic they show better performance in the symbolic representation? It seems that high-SES children were better to perform a specific (approximate vs. exact) operation, when the problems were already presented in the 'according format' (analogue representation for the ANS, symbols/numerons for exact processing) and did not have to be transcoded mentally to fit the adequate code (see the TCM, Dehaene, 1992; Dehaene et al., 2003). That these might be the kinds of problems that children might encounter in school and in everyday contexts more often is actually indicated in the confounding of demanded accuracy and format in current research. This might hint to more experience and practice of classical mathematical problems presented to these children. This experience seems to be lacking in low-SES children, because they performed equally in both formats within a given code. However, the only task in which they did not perform worse than their high-SES peers is the approximate symbolic task. This task actually was the one that they were most successful in (but note that direct comparisons between approximate and exact performances are not possible due to the differing answering format). It thus seems to be the measure in which the high-SES children of our sample have not yet raced ahead due to more experience, making it the least susceptible to detrimental environmental influences – at least at this level of formal schooling.

10.2.3 Additional analysis of the intercorrelations according to SES

In Chapter 10.2.1 it was reported that children of different SES backgrounds did not produce a comparable pattern of performances in the four main tasks. Instead there were tendencies of interactions between SES, the format the problems were presented in, and the demanded accuracy: While low SES children were nonsignificantly but slightly better in approximate symbolic than nonsymbolic arithmetic, in high-SES children the pattern was reversed (and significant). A similar (though again nonsignificant) interaction can be seen in the exact tasks – but this time the low-SES participants being slightly better in nonsymbolic than in symbolic exact problems, and high-SES children the other way round. This observation might be a hint to explain the diverging findings in Study 1 (in which intercorrelations as well as correlations with SES displayed rather clear-cut lines between approximate vs. exact problems, irrespective of the format) and Study 2 (see Chapter 10.2.1): There was a different distribution of SES in the participants of Study 2 than in Study 1. We had to shorten the tasks during the testing in Study 1 which resulted in the exclusion of classes (see Chapter 9.2.1). Due to this and to the new SES measure and in Study 2, in the sample of Study 2 there was a larger proportion of children with low SES. With regard to the descriptive interaction of SES and format mentioned above, this might be a reason why we did not find as clear-cut results in the more mixed SES sample of Study 2. We might well obtain findings like in Study 1 when only looking at the high-SES participants of the current Study. Thus, to test if the larger group of low-SES children was responsible for the diverging results, we again performed the analysis of intercorrelations, this time separately for the group of low-SES children (n = 56) and the group of high-SES children (n = 64) in Study 2.

And indeed, the correlation matrix obtained with high-SES children shows a pattern very similar to the results in Study 1 (see Table 19). The highest correlation of $r = .48$ was found between the two exact measures. The two approximate measures were also positively intercorrelated ($r = .29$). As in Study 1, exact symbolic arithmetic did not correlate with any of the approximate tasks. However, there was also a not entirely new but definitely more pronounced finding compared to Study 1: We found significant correlations between exact nonsymbolic performance and both approximation measures in this subsample. In Study 1, the according coefficients had *just* missed significance. Both findings combined confirm the important role that exact manipulation of quantities plays in first graders´ numerical skills, and that it is an important factor in the subsample of high-SES participants.

However, also in this subgroup it is not the only dominating one. This is shown by the still significant partial correlation ($r = .231$, see Table 19) between both approximate tests. So, although the ability to calculate and manipulate quantities with high precision (that is, exactly) plays an important role in children´s performance in all four main tasks, there are specific interrelations between tasks taxing the ANS vs. tasks requiring exact processing and thus are performed in one of the two symbolic codes of the TCM (Dehaene, 1992; Dehaene et al., 2003).

Table 19. Study 2: Intercorrelations of performances in the main task for high-SES children

		Approximate	**Exact**	
N = 64		symbolic	nonsymbolic	symbolic
Approximate	nonsymbolic	**.290*** $p = .020$	**.260*** $p = .038$	.048 $p = .707$
	symbolic		**.298*** $p = .017$	.228(*) $p = .070$
Exact	non-symbolic			**.480***** $p < .001$
Partial correlations controlling for exact nonsymbolic performance				
Approximate	nonsymbolic	**.231*** $p = .033$		-.091 $p = .24$
	symbolic			.101 $p = .213$

Taking a look at the correlation matrix for low-SES participants (see Table 20), however, reveals the source of the pattern in the collapsed sample: In low-SES children, performances in both approximate arithmetic tasks did not (or just marginally significant) correlate. Instead, we find the high correlations between the approximate measures and exact symbolic, but the even higher correlations with exact nonsymbolic performance. Partialling out the common influence of exact nonsymbolic arithmetic in this subgroup leaves only the connection between the approximate and the exact symbolic task, referring again to the upper mentioned importance of symbolic number knowledge, this time specifically in the group of low-SES children.

Table 20. Study 2: Intercorrelations of performances in the four main tasks for low-SES children

		Approximate	**Exact**	
N = 56		symbolic	nonsymbolic	symbolic
Approximate	nonsymbolic	.232^(*) *p*=.086	**.430**** *p*=.001	**.305*** *p*=.022
	symbolic		**.505***** *p* < .001	**.443**** *p*=.001
Exact	non-symbolic			**.512***** *p* < .001
Partial correlations controlling for exact nonsymbolic performance				
Approximate	nonsymbolic	.019 *p*=.445		.109 *p*=.211
	symbolic			**.249*** *p*=.032

Taken together, the results indicate that not only general performance varies with children´s SES, but also the pattern of interrelations between the numerical abilities. While there are strong correlations between the approximate measures as well as between the exact measures in high-SES children, also a transfer across both operational codes was found in this group. This connection was visible in a strong correlation between both approximate measures and performance in the exact nonsymbolic arithmetic problems. This might be an indicator that in addition to ANS acuity, the ability to perform operations in an exact manner has become an important factor in the development of various numerical abilities already after spending a few weeks in school.

The low-SES participants show a different picture. In this subgroup, there were also high correlations between the exact nonsymbolic performance and each of the other tasks. When partialling out this common factor, however, there still was a correlation between the approximate symbolic and the exact symbolic task. This indicates that symbolic number knowledge is one important factor in low-SES children´s performances (see Starkey & Klein, 2008). Furthermore, the missing link between both approximate tasks in low-SES children might point to the idea of an iterative development of approximate and exact skills (Mejias & Schiltz, 2013, Noël & Rouselle, 2011). Given the assumption that low-SES children do not show a qualitatively different, but rather only a decelerated development (Jordan et al., 2008; Jordan & Levine, 2009), one would expect that in the course of development, also low-SES children would 'train' their ANS and connect its recruitment in both the nonsymbolic and symbolic format.

All in all, correlation analyses suggest that in first graders the ability to mentally manipulate quantities in an exact manner is an important element integrating various numerical performances. While in low-SES children, symbolic number knowledge seems to be an additional integrative factor (visible as partial correlations between symbolic tasks); in high-SES children intercorrelations rather seem to be based on a common code in which the according operation has to be performed (exact vs. approximate, visible in the partial correlation between both approximate measures).

10.2.4 Prediction of formal math performance

We performed a hierarchical multiple regression analysis to test if both approximate measures would predict standardized math performance at the end of the school year. Like in Study 1, we controlled for age and WM first. As third and fourth step, we put both approximate measures in the equation, before finally taking NO into account. This analysis showed that of the 52% explained variance in this model (see Table 21), WM was the strongest predictor (18%). Age did not explain a significant proportion. Like in Study 1, symbolic approximation explained a significant proportion of variance (14%), but contrary to the first study, this time also nonsymbolic approximation was a significant regressor and explained additional 8%. Also the last regressor put into the equation, NO, resulted in a significant gain of explained variance (further 11%) beyond the predictive power of the previously included factors.

The latter contradicts the findings of Lyons et al. (2014) or Attout et al. (2014) who both found no predictive relationship between NO and formal math performance in grades 1 and 2. In a previous study with adults, Lyons and Beilock (2011) had found that NO completely mediated the correlation between a nonsymbolic ANS measure and exact calculation performance. They concluded that numerical ordering consisted mainly in both symbolic and nonsymbolic ANS acuity as well as general ordering ability. To test if our results with first graders ground on this relationship, we switched the order in which the regressors were put into the equation. So we aimed to see if NO is the mediating influence responsible for both ANS measures´ predictive power regarding formal math performance (DEMAT 1+) at the end of first grade.

Table 21. Study 2: HMR of approximation and NO on math performance, Model 1: Hierarchical Multiple Regression Analyses of Numerical Ordering and both approximate tasks on standardized math achievement (mean overall raw score of DEMAT 1+, N = 108), after controlling for Age and Working Memory

		DEMAT 1+					
Step	**Regressor**	**R^2**	**R^2 Change**	***p* of R^2 change**	**Beta**	**T**	***p***
1	Age	.006		.427	-.077	-.797	.427
	Age				-.013	-.143	.887
2	WM	.186	.180	**<.001**	.429	4.823	**<.001**
	Age				.007	.084	.933
	WM				.389	4.763	**<.001**
3	approximate symbolic	.329	.142	**<.001**	.380	4.694	**<.001**
	Age				-.024	-.315	.753
	WM				.353	4.529	**<.001**
	approximate symbolic				.312	3.953	**<.001**
4	approximate nonsymbolic	.406	.077	**<.001**	.290	3.663	**<.001**
	Age				-.066	-.936	.352
	WM				.197	2.529	**.013**
	approximate symbolic				.194	2.569	**.012**
	approximate nonsymbolic				.280	3.895	**<.001**
5	NO	.516	.110	**<.001**	.391	4.813	**<.001**

Taking NO into account *before* the two approximate tasks were put into the equation sharpened the picture. NO explained 20% variance beyond WM, while the predictive power of both the approximate symbolic and the nonsymbolic performance were reduced. However, while there was a high percentage loss of explained variance compared to the first model for the approximate symbolic measure (from 14% down to 6%), the percentage of variance explained by nonsymbolic approximation almost remained the same (from 7,7% to 7.2%, see Table 22).

Table 22. Study 2: HMR of NO and approximation on math performance, Model 2: Hierarchical Multiple Regression Analyses of Numerical Ordering and both approximate tasks on standardized math achievement (mean overall percentile rank of DEMAT 1+, N = 108), after controlling for Age and Working Memory

		DEMAT 1+					
Step	**Regressor**	**R^2**	**R^2 Change**	***p* of R^2 change**	**Beta**	**T**	***p***
1	Age	.006		.427	-.077	-.797	.427
2	Age				-.013	-.143	.887
	WM	.186	.180	**<.001**	.429	4.823	**<.001**
3	Age				-.059	-.759	.449
	WM				.214	2.483	**.015**
	NO	.386	.200	**<.001**	.495	5.817	**<.001**
4	Age				-.037	-.494	.622
	WM				.228	2.766	**.007**
	NO				.400	4.621	**<.001**
	approximate symbolic	.444	.058	**<.001**	.257	3.272	**.001**
5	Age				-.066	-.936	.352
	WM				.197	2.529	**.013**
	NO				.391	4.813	**<.001**
	approximate symbolic				.194	2.569	**.012**
	approximate nonsymbolic	.516	.072	**<.001**	.280	3.895	**<.001**

This is in line with the hypothesis that the relationship between both approximate arithmetic measures and formal math performance should not be completely mediated by Numerical Ordering ability. We tested this assumption with mediation analyses using a bootstrap estimation approach with 1000 samples (Preacher & Hayes, 2008). It showed that children´s performance in the approximate nonsymbolic task was not a significant predictor of numerical ordering ability ($b = .017$, $t = 1.62$; $Se = .022$, $p = .108$), and consequently the reduction from the indirect/total effect ($b = .25$, $t = 4.74$, Se = .41, $p < .001$) to the still significant direct effect of symbolic approximation on math performance (*b* = .21, *t* = 4.56, *Se* = .04, *p* < .001) was not a significant mediation (95% CI = -.033, .122) (see Figure 11).

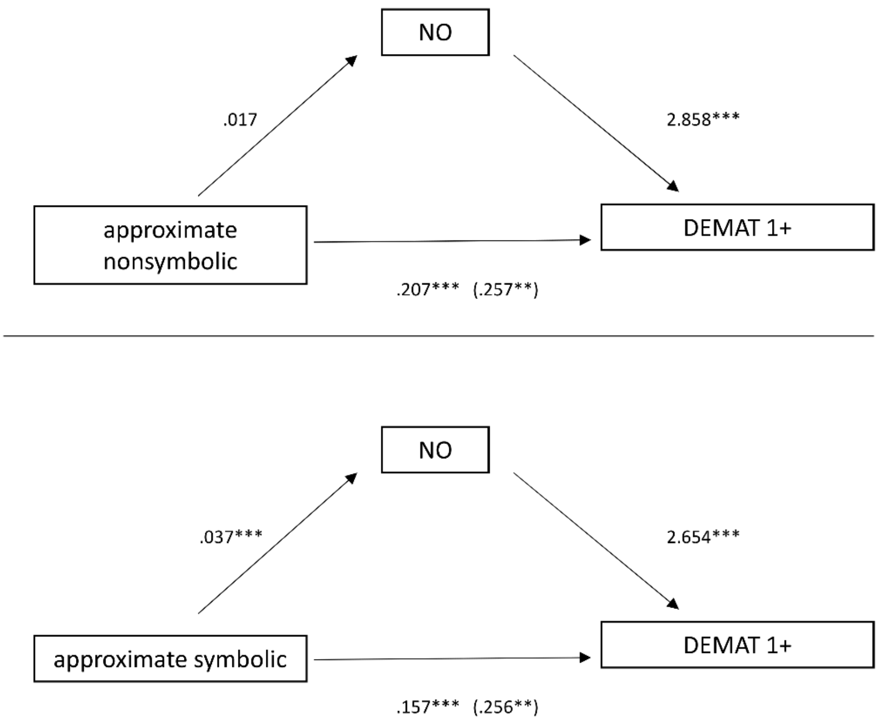

Figure 11. Study 2: Mediation Analyses for the relationship between approximate arithmetic and formal math performance (coefficients without brackets are the respective direct effects, the coefficient in brackets is the indirect effect, that is, without regarding the mediating variable in the model. Note that the coefficients are unstandardized as recommended by Hayes, 2013): No significant mediation by numerical ordering ability for nonsymbolic, significant but incomplete mediation for symbolic approximation

On the contrary, when testing a possible mediation of the relationship between approximate symbolic performance and DEMAT 1+ scores through NO, we found that symbolic approximation significantly predicted NO (*b* = .037, *t* = 3.74; *Se* = .001, *p* < .001). NO, in turn, was a significant predictor of DEMAT 1+ performance (*b* = 2.65, *t* = 5.97; *Se* = .445, *p* < .001). In this case, the total effect (*b* = .26, *t* = 4.87; *Se* = .053, *p* < .001) was reduced significantly when taking NO into account as a mediator (*b* = .16, *t* = 3.24; *Se* = .049, *p* = .002; 95% CI = .034, .163).

To secure that the connection between approximate arithmetic and formal math performance is a domain specific relationship, we computed hierarchical multiple regression analyses with reading comprehension as the dependent variable. Putting in the same regressors as for the DEMAT 1+ produced the results depicted in Table 23. Together, the regressors explained 33% of the variance in reading comprehension. Crucially, the only significant predictors were Age, WM, and NO. WM had the highest predictive capacity (16%), followed by NO (10%). Beyond that and contrary to the pattern in formal math assessment, both approximate measures did not explain further variance in reading comprehension.

Table 23. Study 2: HMR of NO and approximation on reading comprehension: Hierarchical Multiple Regression Analyses of Numerical Ordering and both approximate tasks on standardized reading comprehension assessment (subtest word understanding of ELFE 1-6, N = 110), after controlling for Age and Working Memory

		ELFE 1-6					
Step	**Regressor**	**R^2**	**R^2 Change**	***p* of R^2 change**	**Beta**	**T**	***p***
1	Age	.040		**.035**	-.201	-2.134	**.035**
	Age				-.140	-1.601	.112
2	WM	.205	.164	**<.001**	.410	4.699	**<.001**
	Age				-.174	-2.104	**.038**
	WM				.270	3.007	**.003**
3	NO	.302	.097	**<.001**	.342	3.846	**<.001**
	Age				-.196	-2.361	**.020**
	WM				.249	2.773	**.007**
	NO				.327	3.694	**<.001**
4	approximate nonsymbolic	.321	.019	.090	.142	1.714	.090
	Age				-.188	-2.258	**.026**
	WM				.251	2.793	**.006**
	NO				.305	3.303	**.001**
	approximate nonsymbolic				.127	1.488	.140
5	approximate symbolic	.325	.004	.407	.073	.832	.407

10.3 Discussion

Study 2 was an improved and adjusted replication of Study 1. Compared to the first exploration, we realized our testing in a sample of more diverse socioeconomic backgrounds and administered adjusted tasks. We assessed two additional cognitive control variables (reading comprehension and numerical ordering) and tested a larger sample. In the following we will discuss the results ordered by our research questions and hypotheses. Like in Study 1, our participants showed better performance in the nonsymbolic approximate than in the symbolic approximate problems. This effect was reversed in the exact tasks (better performance in symbolic than nonsymbolic one). The general performance level was lower in Study 2 than in Study 1, especially in the approximate nonsymbolic and the exact symbolic task. We suspect that this goes back to the more balanced distribution of SES in our second study compared to the first study, in which there had been an overproportion of high-SES children.

Intercorrelations at the beginning of first grade. Not only the general performance level, but also the pattern of correlations between the four main tasks was different in Study 2. Contrary to the findings in Study 1, performances in *each* of the four main tasks were intercorrelated. Partial correlations revealed that this seemed to go back to a communal factor captured by our task of exact nonsymbolic arithmetic. While in Study 1 there had already been some indication (marginal significant correlations) that the exact nonsymbolic task might integrate all numerical abilities tested in our main tasks, this became more pronounced in Study 2. When this influence was partialled out of the intercorrelations between the measures, the only remaining significant correlations were those between tasks containing Arabic numerals. On the one hand this indicates that the importance of the ability to manipulate quantities in an exact manner is independent from the format in which a problem is presented to the learner. On the other hand it suggests that aside from that, understanding of Arabic symbols has become an important ability in first graders.

To test if these results – which were quite different from what we had found in Study 1 – were due to the larger proportion of low-SES participants in Study 2, we computed the correlation analyses separately for low- and for high-SES children. This revealed that indeed only in the low-SES subgroup, both exact measures correlated highly with both approximate measures. Curiously, in this subgroup the approximate measures did not correlate with each other. Partialling out the influence of exact nonsymbolic arithmetic (which again had obtained the highest correlations coefficients with the other three main tasks) left only a significant connection between the approximate symbolic and the exact symbolic task (as we found in the collapsed sample). On the contrary, in the high-SES sample, there was a significant positive correlation between both approximate measures, as well as between the exact ones. Exact symbolic performance did not correlate with the approximate measures, but exact nonsymbolic performance once again did. Partialling out the common influence of exact nonsymbolic processing did not extinguish the intercorrelation between both approximate measures in this group, indicating that a genuine connection between symbolic and nonsymbolic approximation processes existed in these children. The findings resemble those found in Study 1, although there the correlation between nonsymbolic exact arithmetic and both approximate measures had only been marginally significant. The obtained patterns thus suggest that the relations between numerical abilities differ between high- and low-SES first graders. In both groups, the ability to exactly manipulate quantities seems to be a common factor underlying and unifying all measured performances. For the performance of low-SES children however, a second factor that might play a role seems to be symbolic number knowledge (involvement of Arabic numerals was the common characteristic of the two tasks that were still interrelated when controlling for exact nonsymbolic arithmetic). In the high-SES group on the other hand, besides that common factor of exact mental manipulation of quantities, the required accuracy of the to-be-performed operation seemed to be determining for children´s performance (demanding only approximate processing was the common characteristic of the two tasks that were still interrelated when controlling for exact nonsymbolic arithmetic).

Correlations with SES. The necessary additional SES-split in the analysis of intercorrelations reported above already stressed the importance and the complexity of the impact which socioeconomic factors have on numerical skills. We calculated a correlation matrix of children´s main task scores with SES. It revealed that with the exception of the approximate symbolic task, performance in each of the tasks differed with children´s SES. Thus, children with high SES obtained better results in both exact tasks as well as in NO; they showed higher WM capacity and, surprisingly, also better performance in nonsymbolic approximate arithmetic.

Like in Study 1, the existing correlation of exact nonsymbolic arithmetic with SES (contrary to the results of Jordan et al., 1992) and the (again) missing correlation between approximate symbolic arithmetic and SES at first seem to point to the idea that the influence of socioeconomic factors on numerical abilities is mediated by linguistic skills. Language is a fundamental basis of both symbolic representational codes which are supposed to be the modules for exact numerical processing of any kind (Dehaene, 1992; Dehaene et al., 2003). In the first study, this was mirrored in significant correlations between SES and the exact arithmetic tasks and missing correlations between SES and the approximate arithmetic tasks. Surprisingly, however, in Study 2, also approximate nonsymbolic arithmetic was influenced by SES.

We thus think that in our participants, the process of exact skills reinfluencing and refining the ANS (Kolkman et al., 2013) has already started and became statistically measurable in the larger sample of Study 2. However, symbolic approximation (still) seemed to be robust against detrimental socioeconomic factors. This indicates that approximating symbolically presented problems is the least trained competency among the four main tasks. Approximation with nonsymbolic quantities is familiar from early childhood on, and the typical exact symbolic calculation is a skill which almost every child is eager to acquire already before and certainly when starting school. Also exact processing of nonsymbolic quantities should probably be practiced more and more in everyday activities, for example when sharing items with increasing accuracy. This missing experience with the symbolic approximation task opens the door for counterproductive interferences especially in high-SES children: McNeil et al. (2011) assumed that the more practiced a learner is in *formal* mathematical problems, the more he/she internalizes the requirement to produce an exact result. Solving approximation problems that due to their symbolic format activate this requirement thus might pose difficulties specifically to high-SES children. In favor of this argument is also our finding that – unlike McNeil´s participants – our high-SES group does not show greater accuracy in canonical than in non-canonical trials of the approximation tasks. McNeil at al. (2011) explained this (to them still hypothetical) result with canonically presented problems activating the demand to produce an exact result, thus hindering approximation performance.

The results indicate that actually approximate and exact processing become integrated abilities with formal education (see also Kolkmann et al., 2013). As McNeil et al. (2011, p.64) put it, “children’s knowledge of exact arithmetic shapes their ability to use their approximate number system to reason arithmetically”. The pattern indicates that already at the beginning of first grade, children with high SES seem to have started to refine their ANS, which is visible in the intercorrelations of the main task performances as well as in the fact that nonsymbolic approximation was affected by SES. Symbolic approximate arithmetic on the other hand seemed to be the ability influenced the least by socioeconomic factors. Children of high SES were better in tasks that did not require an unusual direction of transcoding but in which the problems were already presented in a format that fit the representational code in which the operation had to be performed.

What approximation measures can predict formal math performance? In Study 2, we found that after controlling for age and WM, both approximate measures explained a significant proportion of the variance (symbolic: 14%; nonsymbolic: 8%) but that numerical ordering, contrary to previous findings in this age group (Lyons et al., 2014; Attout et al., 2014), also showed substantial predictive power (11%). So we changed order of predictors and found that after controlling for numerical ordering first (explaining 20%), both approximate measures still predicted a significant proportion of the variance in formal math performance at the end of first grade. However, only the symbolic approximation measure was markedly reduced in its explanative power (down to 6%), whereas nonsymbolic approximation remained more or less unaffected in its predictive capacity (still 7%). Mediation analyses confirmed that numerical ordering was partially mediating the connection between symbolic approximation and formal math performance, but did not significantly mediate the predictive relationship of nonsymbolic approximation to the DEMAT 1+ score. These results are in line with our hypothesis that symbolic approximate arithmetic should be suited to predict later math performance and that numerical ordering ability should not completely mediate this relationship. Approximate arithmetic differs from the frequently employed comparison task to measure the ANS. For the relationship between this task and formal math performance, a complete mediation by NO had been found in adults. However, to compare two quantities and name the larger one, mainly ordinal knowledge is needed (but see Vogel et al., 2015). In approximate arithmetic, different characteristics and principles have to be understood (Cowan, 2003) which suggested that ordinal understanding alone would not account for the relationship between approximation and formal math performance.

11 Discussion for Studies 1 and 2

In our Studies 1 and 2 we tried to disentangle the implications and effects of two factors that are important in early arithmetic: Demanded accuracy (if the operation has only to be performed approximately or if an exact result is demanded) and the format a problem is presented in (nonsymbolically or symbolically).

Dehaene´s Triple Code Model clearly states that all approximation processes, even if the problems are presented in Arabic numerals, are carried out on the representational level of the ANS (or the analogue magnitude code). However, in current research mostly nonsymbolic tasks to measure the ANS are applied. The few studies that employ symbolic ANS measures almost exclusively administer comparison tasks. However, recent research indicates that approximate arithmetic might be a more direct predictor of formal math performance (e.g. Pinheiro-Chagas et al., 2014). This task format has rarely been tested symbolically, although already preschoolers are able to perform simple arithmetic in a symbolic format as long as they only have to do it approximately (Gilmore et al., 2007). Thus, one general goal was to explore the potential of such a task in the field of prediction and intervention (see Experiments 3 and 4 for the latter), but also with regard to the underlying ability´s possible robustness against socioeconomic influences, and its interrelations to children's mastery of different manifestations of arithmetic problems.

With regard to this general goal, we administered arithmetic tasks incorporating all 2x2 combinations of the two upper mentioned factors. Thus, in the four main tasks, addition problems (in Study 1 we also tested the use of subtraction problems, but only administered addition problems in Study 2 to increase reliability) were presented nonsymbolically and symbolically with the requirement of producing an exact result; and furthermore as symbolic and nonsymbolic approximate arithmetic problems. In both studies we also assessed children´s working memory with a group-adjusted measure adapted from the HAWIK IV Zahlen Nachsprechen subtest (Petermann & Petermann, 2007). At the end of term, (i.e., 9 months after the first testing) we administered the DEMAT 1+ to assess formal math performance. In our second study, we furthermore included a measure for numerical ordering ability to determine the character of the relationship between approximate and formal math skills more closely. We also incorporated a reading comprehension test to secure that the predicted relationships were specific to the numerical domain and not an outlook on general academic performance.

We aimed to find (first) answers to the following questions. (1) How is performance in these tasks interrelated at the beginning of formal education? (2) Which of these tasks are or are not affected by children´s SES? (3) Can both or one of the approximate arithmetic measures function as a predictor for formal math performance at the end of first grade? (4) What role does numerical ordering ability play in the predicted relationships?

Performance and interrelations. In both studies, children performed above chance in all main tasks. Ratio effects were found in both approximation tasks, ensuring us that the children did not rely on other factors than numerosity to solve the problems, but actually approximated in these tasks.

Intercorrelations between children´s performances in the main tasks were quite clear-cut in Study 1. Besides marginally significant correlations between nonsymbolic exact performance and both approximate measures, significant coefficients were only found between the approximate measures as well as between the exact ones, indicating the distinctiveness of the codes in which operations (approximate vs. exact) have to be performed and underlining the upper mentioned proposition of the TCM. Despite methodological improvements and adjustments (see Chapter 10.1 for a detailed description of the alterations), the results of Study 2 are less clear. Performances in all main tasks were now intercorrelated, with the highest coefficients obtained whenever the nonsymbolic exact arithmetic performance was part of the correlation. This in parts replicates the (only marginally significant) correlation between nonsymbolic exact and both approximate tasks in Study 1. So, we assume that in first graders the ability to mentally manipulate quantities in an exact manner is one important factor that integrates the numerical abilities assessed in our study (see for example Resnick, 1992). This ability must be regarded as independent from symbolic skills, though, as in Study 1 only the *nonsymbolic*, not the symbolic exact task demonstrated significant correlations with the other main tasks. In Study 2, these specific correlations were not the only, but certainly the highest ones.

Contrary to Study 1, we found significant intercorrelations between the exact *symbolic* task and both approximate measures in Study 2. To get to the ground of these differences we analyzed correlations separately for low and high-SES children. In our second study and partly due to the changed measure of SES, there had been a considerably larger proportion of children with low SES. Indeed, the correlational patterns obtained in the high-SES subsample more closely resembled the results of Study 1. The correlation of nonsymbolic exact arithmetic with the approximate measures remained significant, however, but the correlation between symbolic exact processing and approximation vanished. In the subsample of low-SES participants on the contrary, both exact measures correlated with each other but also with both approximate measures, which in turn did not intercorrelate. The finding of a correlation between both approximate measures only in high-SES participants (and not in low-SES participants) suggests that high-SES children use their ANS more consistently over different formats of problems. This is further confirmed by partial correlations. In both subsamples of Study 2, we controlled for the common variance of nonsymbolic exact processing to see which relationships between numerical abilities would hold beyond that factor in each group. In the high-SES group, the remaining correlation was the one between both approximate measures. In the low-SES group, however, only the correlation between both symbolic tasks remained significant. So, we conclude that on the one hand, for both subgroups it is the ability to exactly manipulate quantities that connects performances in all tasks (see Resnick, 1992). On the other hand, beyond that, high-SES children seem to have sharpened the distinct codes in which an operation has to be performed, leading to remaining partial intercorrelations between task performances processed within a specific code. Low-SES children do not seem to have made this development (yet). Instead, in this subsample and as has been frequently suggested (Jordan et al., 2008; Jordan & Levine, 2009), rather the symbolic number knowledge seems to be playing a role in connected performances.

Impact of SES. In our first study, the effects of SES on arithmetic performances seemed to have stuck to the clear-cut lines already reported for the intercorrelations in the first study. Both exact measures were significantly or marginally significantly correlated with SES, whereas this correlation was entirely absent for the two approximate measures. That socioeconomic factors impeded exact but not approximate processing irrespective of the format in which the problems were presented might be an insight in the character of the relationship between SES and numerical abilities. Linguistic abilities have already been demonstrated to be impaired in low-SES children and to mediate the detrimental influence of SES on executive functions (Noble et al., 2005). So, also the relationship between SES and math performance might be mediated by language skills. Exact processing should thus suffer from low SES no matter if presented symbolically or nonsymbolically, as both representational codes which are capable of exact processing are strongly based on human language (Dehaene, 1992; Dehaene et al., 2003). Approximate processes on the other hand are assumed to be carried out in the language-independent analogue magnitude code and hence would remain unaffected by SES. This pattern (exact performances are, approximate performances are not influenced by SES) is what we found in Study 1.

For the second study, already the results of the separate correlation analyses reported above indicate a more complex configuration of SES´s effects on children´s performances. Here, not only the exact tasks but also the nonsymbolic approximation measure correlated with SES. Contrary to Study 1, this does not fit any of the two competing hypotheses that we formulated. It thus seems that at the beginning of first grade, there are actually no clear-cut lines between the factors 'demanded accuracy' and 'format' anymore. This is in line with and partially mirrors the results of Mejias and Schiltz (2013) who found that right before starting school, children´s exact and approximate number representation was influenced by SES in a symbolic as well as a nonsymbolic format.

But why then was symbolic approximation robust against the influence of SES in our tasks? Looking at Figure 10 suggests that this correlation was missing not because low-SES children performed particularly well in this task, but rather because high-SES children performed remarkably poor. One reason for this might be the missing experience with this kind of task. This in turn facilitates interference of the implicit requirement felt by high-SES children to retrieve an exact result when encountering an arithmetic problem presented to them in Arabic numerals. This idea receives further support by our additional finding that canonicity of the trial did not elicit differential performance in our participants – a finding that McNeil et al. (2011) explained with the same reason (that is, a canonical problem would strongly challenge an exact solution).

This suggests that the increasing practice that especially high-SES children obviously obtain in the abilities tested in the other main tasks[8] may in specific circumstances also

[8] While approximation can be performed early in life and consequently typically starts with analogue magnitudes, neither in school nor at home is there a lot of enforcement to approximate when Arabic numerals are involved. Also exact processing with nonsymbolic quantities might be trained to a stronger degree, because in every day context, when children encounter quantitative requirements like

pose an obstacle to a successful problem solving. This is the case in unfamiliar problems which require a flexible solution process that does not fit learned procedures. It seems thus that symbolic approximation might be the 'last' ability in which high-SES children have not yet raced ahead compared to their low-SES peers at the beginning of first grade.

Prediction of later formal math performance with ANS measures. In both studies and in line with the expectation that approximate arithmetic in first grade should predict formal math skills, we found the symbolic approximate arithmetic task to explain a significant proportion of variance in children´s DEMAT 1+ score at the end of first grade. However, in Study 1, the nonsymbolic approximate arithmetic task was not a significant regressor in our model. As alternative explanations seemed unlikely, we suspected that due to our small sample size there was a problem of power in this study. For comparison measures, it has already been stated that connections between nonsymbolic tasks and exact math performance seem to be weaker or more unstable than between symbolic tasks and math skills. As a reason for the weaker relationship, sometimes methodological issues in the administration of the task or in the comparability of different studies are discussed. However, DeSmedt et al. (2013) point out that this should also apply to the symbolic comparison measure, which it does not. It thus might also be that the representations and processes tested with nonsymbolic comparison tasks are not as critical for the development of school-relevant competencies (DeSmedt et al., 2013). This might explain why we found the effect of symbolic approximate arithmetic as a predictor in the first study, but not of nonsymbolic arithmetic. A power analysis confirmed that with our sample size, we only had a one-third chance of detecting the correlation between nonsymbolic approximate arithmetic and formal math performance.

Regression analyses in the second study with doubled sample size also argue in favor of this idea. This time, *both* approximation measures predicted significant proportions of variance in formal math performance after controlling for age and working memory. Symbolic approximation accounted for almost double as much explained variance as nonsymbolic approximation. Numerical ordering, which frequently has been proposed to play an important role in connecting the educationally acquired formal math skills with ANS-based processes, also was a significant regressor. This is in line with our ex-

sharing or deciding for alternatives based on quantity, they perform these operations on an increasingly exact level as soon as they are capable of it (Baroody, Wilkins, & Tiilikainen 2003).

pectations that in first grade, approximate arithmetic in both formats should be suitable predictors for later math achievement in school, and that the symbolic approximate arithmetic should be an even better one. Numerical ordering ability partially mediated this relationship for symbolic approximation. The results support the notion that the transfer of the principles (that can be understood at the respective developmental level and are used in both the ANS and the formal math tasks) onto a symbolic context models the relationship between ANS measure and math assessment. While in preschoolers these were symbolic abilities of number representation and counting (vanMarle et al., 2014); after the onset of formal education numerical ordering becomes more and more important to represent quantities with increasing accuracy (Lyons et al., 2014). In first grade, however, this mediation is not fully developed yet, indicating that there is still more to understand in the numerical domain for participants (Cowan, 2003).

Considering the results of Studies 1 and 2, we conclude that not (only) symbolic number knowledge can account for differences in numerical performance according to socioeconomic factors. This is indicated by the finding that low-SES children perform poorer than high-SES children in symbolic as well as nonsymbolic exact arithmetic. Beyond that, more important for first graders´ numerical performance in the different tasks seems to be the ability to exactly process quantities *nonsymbolically* – without the need of explicit symbolic number knowledge. However, taking out this common factor of variance left only the intercorrelation between symbolic tasks in low-SES children and the intercorrelation between the approximate measures in high-SES children. Exact processing has to be carried out in one of the two symbolic codes of the TCM. That exact nonsymbolic arithmetic showed such tight connections to all of the other measures again stresses the importance of the codes depending on language. It thus is worthwhile for future research to follow this thought and test the possibility that linguistic abilities mediate the influence of SES on numerical abilities more directly.

However, although our results indicate a possible mediating role of language, they also show the complexity of this influence: Differential patterns of intercorrelations exist in high- and low-SES first graders. In the low-SES subsample, in fact symbolic number knowledge seems to be another integrating factor in all performances (see also Jordan et al., 2008). In high-SES children, this role seems to have vanished, instead, here the intercorrelations of performances carried out within a common code of the TCM are more pronounced. In line with these findings are also the diverging results obtained in the two studies using the computer game “Number Race” reported above: Originally, Wilson et al. (2006) found that the different approximation tasks included in this game only benefitted symbolic approximate skills in low-SES first graders. Obersteiner et al. (2013), however, obtained beneficial effects with the same game in both symbolic and nonsymbolic approximation performance in a sample of the same age. An explanation for this divergence might also be a differential composition of participants. In the study of Obersteiner et al. (2013), children of different SES participated, while in Wilson et al.´s work only low-SES children took part (2006). The diverging results match our assumption that first graders of higher SES might have sharpened their approximate number system through the acquired exact processing skills, leading to a stronger connection between approximation in both formats and hence also explaining the benefit in both formats in the study of Obersteiner et al. (2013).

It thus seems that the two global factors that are of interest in this thesis, that is, demanded accuracy and format, can be applied to differentiate between low- and high-SES children´s math performance. For all first graders, the ability to exactly manipulate quantities correlated with their other numerical skills, demonstrating an increasing sharpening of representations and arithmetic processing like suggested by Mejias and Schiltz (2013) or Kolkman et al. (2013). While knowing number symbols can pose an important difficulty to those who are not trained in it (low-SES children), for those who are, the processes that are demanded seem to come to the fore. Probably, sufficient practice in recognizing number symbols allow to focus on the operations and to integrate ones knowledge and understand the common basis of operations independent from format. When we follow the assumption that low-SES children display a decelerated, but not qualitatively different development of numerical skills (Jordan et al., 2008; Jordan & Levine, 2009), the findings can be interpreted as an iterative influence between exact and approximate skills as soon as formal education starts (see also Resnick, 1992). It should be expected that in the course of development, also the low-SES first graders tested here might use the ability to exactly manipulate quantities (which already has been found to influence all their performances) to calibrate their ANS acuity. This should lead to an integration of performances in approximate tasks like ours, irrespective of format.

12 Experiment 3

Experiment 3 was designed to test if approximately calculating symbolic and/or non-symbolic arithmetic problems which contain the possibility to exploit a specific mathematic principle would benefit children´s ability to recognize and exploit this principle in a more formal, exact context afterwards. Additionally, we aimed to test which kind of knowledge – procedural or conceptual – would profit from referring to the principle in an approximation task. Furthermore, we wanted to determine the respective possible effects in comparison to the impact of an explicit and direct instruction about the principle. Thus we should get a first impression of the actual ecological value of any effect we hope to find. If approximation of any kind benefits principle knowledge, it is of great interest if this benefit is equivalent or even bigger than the effect teachers achieve with conventional instruction that includes an explicit explanation of the principle.

We decided to test second graders because we assumed that in order to foster conceptual knowledge, there probably already to be a basis of this knowledge to some degree. That is, children should have heard about the principle in question. We chose the additive law of commutativity for our experiment. Its core property, the order-irrelevance principle, is ubiquitous in everyday situations[9]. The commutativity principle states that in binary operations of addition or multiplication, the order of the operands does not affect their sum or product (cf. a + b = b + a; see Cowan, 2003). Here, we wanted to test if experiencing the commutativity principle in approximate calculation first will foster the *spontaneous* exploitation of the principle in exact arithmetic problems. To avoid frequent methodological artefacts (see below) and to assess conceptual and procedural knowledge independently of each other, we took a new approach to assess both kinds of knowledge in exact arithmetic calculation (Gaschler, Vaterrodt, Frensch, Eichler, & Haider, 2013): first, we never informed the children about the existence of commutative problems. Second, we used two different task types in order to assess their procedural and conceptual knowledge separately. Both tests were presented in a school-like situation. The so-called *computation task* was aimed at assessing procedural knowledge. The *judgment task* served to measure children's conceptual knowledge.

[9] Note again that the theoretical background and the description of the majority of the tasks for Experiments 3 and 4, as well as the findings of Experiment 4 have been prepublished with the permission of the Dean of Research of the University of Cologne (Hansen et al., 2015).

In Studies 1 and 2, we have obtained diverging results of what to be the best predictor of mathematical skills at the end of the first year of formal schooling. In the first study, approximate symbolic performance predicted the formal math performance at the end of the term. In our second study, and in line with previous literature, it was both the approximate symbolic, but additionally also the approximate nonsymbolic task. The latter already has been found to be a suitable intervention to strengthen children´s performance in simple arithmetic tasks (Hyde et al., 2014). So we decided to test approximate symbolic as well as approximate nonsymbolic problems as possible inductions for the understanding of mathematical principles. It has to be noted that in our studies, symbolic approximate arithmetic did consistently not correlate with SES. This makes it a promising composition to foster arithmetical understanding of children with diverging backgrounds at the beginning of formal education.

12.1 Method

The main goal of the current experiments was to examine if approximately calculating the results of symbolic problems that link to a specific mathematical principle can alter children's ability to spontaneously spot and use this arithmetic principle in exact arithmetic problems. A second goal was to test whether alluding to a principle in approximation affects only procedural or additionally also the conceptual understanding of the principle.

According to current practice in school, precursory knowledge is usually not actively used to foster the acquisition of formal knowledge. The longer children have attended school, they instead tend to increasingly separate their precursory mathematical knowledge acquired in real world contexts from formal mathematical understanding (see e.g., Bönig, 2001; LeFevre, Greenham, & Waheed, 1993; Verschaffel, de Corte, & Lasure, 1994). One way to avoid this phenomenon might be to explicitly rely on such precursory mathematical knowledge when introducing new arithmetic concepts in school (e.g. DeCaro & Rittle-Johnson, 2012; Obersteiner et al., 2013). For instance, several studies provide evidence that relying on children's ability of approximate calculation also facilitates their exact calculation competencies (Cantlon et al., 2009; Charras et al., 2012; Gilmore et al., 2007, 2010). Recently, Hyde et al. (2014) trained children with nonsymbolic approximate addition and number comparison problems. Subsequently, they let children work through an exact symbolic addition task. Compared to two different control conditions with other training tasks, the short approximate calculation training significantly improved the children´s performance in the subsequent exact symbolic addition task.

These findings raise the question of whether the activation of such precursory mathematical knowledge can also enhance the understanding of abstract mathematical principles like, for instance, equivalence problems or the commutativity principle. As reported in Chapter 3.2, first results come from Sherman and Bisanz (2009; see also Obersteiner et al., 2013). They investigated the effect of concrete, nonsymbolic material on the understanding of equivalence problems in formal arithmetic. First, they instructed second graders to solve nonsymbolic equivalence problems and afterwards symbolic equivalence problems. In a second condition, participants received the reverse order. The results revealed that solving nonsymbolic problems first facilitated the performance in symbolic problems, whereas symbolic problems did not affect the performance in nonsymbolic problems.

Most studies on commutativity assess children's knowledge about this principle by asking them to solve an arithmetic problem first and then to describe their strategy (Baroody & Ginsburg, 1983; Cowan & Renton, 1996; Canobi, 2005). For example, Canobi, Reeve, and Pattinson (1998) told children to solve addition problems, interspersed with commutative ones. After a child had solved a problem, the interviewer asked how she/he “worked out the answer”, and prompted her/him when necessary. For instance, children who counted were asked, “What was the first number you said as you started counting?” They assumed that the children had used their conceptual knowledge of commutativity if they reported solving a problem by referring to a related, immediately preceding problem, for instance, "I saw that 2 + 7 had the same numbers as 7 + 2 (the preceding problem), so I knew the answer to 2 + 7 was 9 as well" (Canobi et al., 1998). This combined assessment of procedural and conceptual knowledge enables researchers to investigate if a child only applies the strategy (procedural knowledge) or if he/she additionally understands why the strategy applies (Baroody, 1987; Baroody, 2003; Cowan & Renton, 1996; Hiebert & LeFevre, 1986; LeFevre et al., 2006). However, it is unclear whether asking children to explain their solution strategies might trigger the use of shortcut strategies during the investigation. It is conceivable that children look at the problems more attentively and select strategies more flexibly when they are asked to verbalize their procedures. Consequently, conclusions concerning the question of whether a child is able to spontaneously use her/his knowledge about a certain mathematical principle might vary depending on the tests that were applied. On that note, Schneider and Stern (2010) called for assessing procedural and conceptual knowledge in the context of arithmetic development multifaceted and independently of each other (see also Schneider, Rittle-Johnson, & Star, 2011).

Participants. Participants were 140 second graders from three different elementary schools in Cologne. All children had permission to take part in our study and were assigned to three experimental conditions (37 children in the nonsymbolic induction condition, 50 in the symbolic induction condition, and 53 in the instruction condition). In this experiment, amongst others we tested conceptual knowledge about the additive law of commutativity. In the so-called 'judgment task', children had to mark problems that would need no computation (that is, commutative problems, see description of the task below). There was a group of children who did not understand the instructions and marked each problem presented in the task – these children (n = 15) were excluded from analysis (three in the nonsymbolic condition, 11 in the symbolic condition, and one in the instruction condition) along with children who were excluded for other reasons (for example because they were classified as special-needs students or observed to not follow the instructions). This was the case for eight children in the nonsymbolic condition, six in the symbolic condition, and four in the instruction condition. Thus, 107 participants (53 girls) with a mean age of 8 years and 1 months (SD = 9 months) remained in our testing (26 children in the nonsymbolic induction condition, 33 in the symbolic induction condition, and 48 in the instruction condition).

Material. In the computation task, children received two subsets of problems; one subset contained commutativity problems, the other did not. Time to work through each subset was limited so that it was very unlikely for the children to solve all problems. Children were explicitly told that it was impossible to solve all problems of a subset within the given time to prevent a loss of motivation. Importantly, this time limit enabled us to compare the number of solved problems between the two subsets. If children rely on the timesaving commutativity-based shortcut (that is, writing down the solution of a commuted problem without calculating anew), they should solve more problems per time in the subset containing commutative problems compared to the one that lacks such shortcut options.

The logic for assessing conceptual knowledge was similar: in the judgment task, children received commutative and noncommutative problems without any further information. They were simply asked to mark those problems which they believed required no calculation in order to get to the result. If children possess conceptual knowledge, they should be able to figure out that this only applies to commutative problems (Haider et al., 2014). We assume that the judgment task taps children's metacognitive knowledge of the commutativity principle which according to Flavell (1976) is an essential component of conceptual understanding. Thus, this task allowed us to assess conceptual knowledge about commutativity without informing children about the existence of commutative problems.

Before children received these two exact arithmetic tasks, depending on their respective condition they either received an explicit instruction about commutativity, or were administered their respective induction task (symbolic approximate arithmetic or nonsymbolic approximate arithmetic, both were similar to those used by Gilmore et al., 2007). In the following, both inductions, the procedural and the conceptual task are described in greater detail.

Induction Tasks. Each of the two induction tasks was primarily used to trigger the exploitation of commutativity in exact arithmetic problems. Therefore, it contained 14 pairs of two-digit addition problems that were either commutative (i.e., the order of the addends in the first problem was reversed in the second problem) or noncommutative. The induction tasks are similar to the approximation tasks used in Studies 1 and 2. In each of the 14 trials, images of two children (Tim and Lisa) were shown together with a large candy (symbolic approximation task) below each child, or two collections of marbles (nonsymbolic task). The two quantities were interspersed with a plus sign. In the symbolic task, the respective candy symbolized the number of candies of Tim and Lisa. To this end, each candy contained a symbolic addition problem composed of two addends larger than 10 (e.g. Tim's candy contained "35 + 31" and Lisa's candy "31 + 35"; see Figure 12). The addends ranged between 13 and 97 leading to results between 36 and 140 (see Appendix F for all problems). Children were asked to answer if both Tim and Lisa possess the same number of candies, or whether Tim or Lisa has more candies. They were explicitly told to estimate, and not to engage in exact calculation.

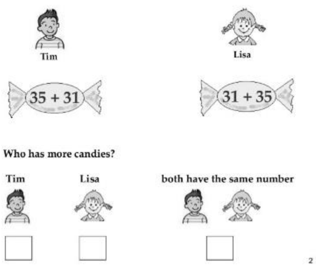

Figure 12. Experiment 3: Example of a commutative trial (symbolic approximation task)

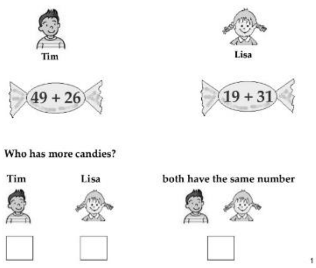

Figure 13. Experiment 3: Example of a noncommutative trial (symbolic approximation task)

The nonsymbolic approximation task was marked by the complete absence of Arabic numerals. Instead, when in a trial of the symbolic task it says “23+25” on Tim´s candy, in the according trial of the nonsymbolic task one marble collection under Tim´s image would consist of 23, the other of 25 marbles. The addends in the trials were identical in the symbolic and the nonsymbolic task.

Seven out of the 14 trials contained commutative problem pairs: the candies/marbles of Lisa and Tim consisted in identical addends in different order (“23 + 45” and “45 + 23”). The remaining seven trials were noncommutative. For each noncommutative trial, like in Studies 1 and 2 we used one of the three ratios 3:5, 2:3 and 3:4 for the two results to be compared Feigenson, Dehaene, & Spelke, 2004; Gilmore et al., 2007, 2010). For instance, the two problems “38 + 28” and “42 + 57” lead to the totals of 66 and 99, resulting in a ratio of 2:3. Using discriminable ratios should minimize the risk that participants mistakenly judge the two addition problems as commutative when they are indeed noncommutative. In four noncommutative problems, Tim possessed the larger amount of candies whereas Lisa did in the remaining three. In addition, in half of all the problems (commutative and noncommutative), the larger addend was the first one. This should discourage using heuristic shortcut strategies (e.g., comparing only the first addend of the two problems). Time limits and the large addends ensured that children did not calculate the results of these problems.

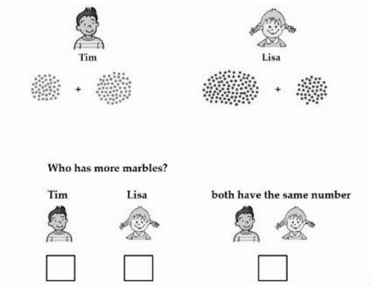

Figure 14 Experiment 3: Example of a noncommutative trial (nonsymbolic approximation task)

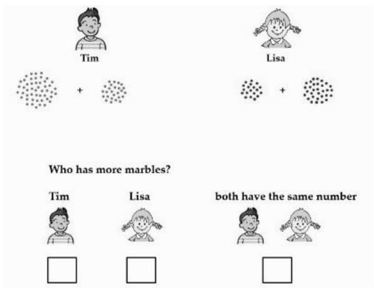

Figure 15 Experiment 3: Example of a commutative trial (nonsymbolic approximation task)

Computation Task. The computation task assessed procedural knowledge and was composed of two subsets, the commutative and the noncommutative subset that contained 30 problems each. Both subsets were presented as small booklets of five pages with six problems on each page (see Table 24 for an example and Appendix G for all problems). In the commutative subset, two out of the six problems per page were commutative to the immediately preceding problem. This was the only difference between the commutative and the noncommutative subset. In both subsets, the problems consisted of two different addends between 1 and 9 (maximum result was 17). We included

"1" as an addend (one problem within each subset), as well as the possibility to repeat the same addend in a problem (e.g., 4 + 4; four problems within each subset) to increase the pool of possible problems.

Table 24. Experiment 3: Examples of the problems presented in Experiments 3 and 4 in the computation task (commutative and noncommutative subsets) and (in Experiments 3, 4 b and c) the judgment task

computation task				**judgment task**	
commutative subset		noncommutative subset		2 + 7 + 9	○
				9 + 5 + 4	○
Exp. 4 a	Exp. 4 b and c	Exp. 4 a	Exp. 4 b and c	**2 + 6 + 5**	○
				6 + 5 + 2	○
2 + 3 =	3 + 5 + 4 =	3 + 2 =	5 + 3 + 4 =	8 + 7 + 5	○
6 + 5 =	**4 + 9 + 8 =**	5 + 6 =	8 + 9 + 4 =	**3 + 5 + 6**	○
5 + 6 =	**4 + 8 + 9 =**	9 + 2 =	6 + 7 + 8 =	**6 + 5 + 3**	○
4 + 3 =	6 + 2 + 5 =	3 + 4 =	5 + 2 + 6 =	2 + 9 + 5	○
9 + 7 =	**9 + 7 + 2 =**	7 + 9 =	2 + 7 + 9 =	**6 + 7 + 9**	○
7 + 9 =	**2 + 7 + 9 =**	8 + 8 =	9 + 4 + 5 =	**9 + 6 + 7**	○

Problems in bold indicate the commutative pairs of the respective task.

Judgment Task. The judgment task (Haider et al., 2014) also consisted of 30 three-addend-addition problems distributed over three pages (problems are depicted in Appendix H). Among the ten problems per page, three problems were commutative to their respective precursor problem. The first page of the judgment task was for training. Participants were instructed to first compute the solutions of the problems on this page. Afterwards, they were instructed to mark those problems which they believed needed no calculation to obtain the correct solution (e.g. Canobi et al., 1998, for similar instructions). For the remaining two pages, children were told to do only the latter; not to calculate before. Therefore, problems on these pages were presented without equal signs. Instead, a circle to the right of the problems could be marked (see Table 24). Bermejo and Rodriguez (1993) found that among 6–7 year-olds, less than 10% needed to actually calculate results to discriminate commutative from noncommutative expressions. So, we are confident that children – provided they have conceptual knowledge about commutativity – are able to recognize commutative problems in this task.

Explicit Instruction. In this condition, children did not work through an induction before being administered the computation and the judgment task, but instead received a direct instruction about the principle of commutativity. They were told that together with the experimenter, they would repeat and practice certain problems that they already had come across in school, the so-called “exchange problems” (*Tauschaufgaben* in German). The experimenter explained that sometimes there are problems in which the numbers are identical and only the order of the operands is exchanged. An example of a three-addend addition problem was written on the board together with its commuted counterpart and pointed out that only the order had changed. It was also stated by the experimenter that no matter of how the addends were ordered, the result would always remain the same, as it always was the case in exchange problems. One would not have to calculate the second problem of a pair of exchange problems anew, but could just carry over the result of the first problem. Thereby, one would save a lot of time!

Procedure. The participants received all problems as paper-pencil tests in the classroom. An experimenter introduced all tasks to the whole group (of up to 25 children). Three to four additional experimenters observed small subgroups of up to five children within the larger group during the entire experiment. Children of different classes were distributed randomly to the different experimental conditions. The experiment started with the induction task or the explicit instruction (see above), followed by the computation task and ending with the judgment task.

Each induction task began with a training sheet containing one approximation trial of the respective representation (candies or marbles, that is symbolic or nonsymbolic). The experimenter explained this problem exemplarily and solved the example together with the children. Once the children signalled that they had understood this instruction, they were asked to work through the 14 trials of the induction task and to solve as many 'candy' or 'marble' problems as possible. Time was limited to 1.5 minutes (enough time to solve all problems without calculating the results).

The computation task also started with a short instruction. Children were told to solve the problems as quickly and as accurately as possible. A warm-up phase with six addition problems followed (all were noncommutative). Children were given one minute to calculate these problems (i.e., sufficient time to solve all six warm-up problems). After this short training, a second instruction followed. Children were informed that for the next two subsets, it would be impossible to solve all problems during the period of time given for each subset. The instruction also stressed that they should work through the problems page by page and from top to bottom. They also were told to work only with a pencil. The time limit for each subset was set to 3 minutes. After having finished the first subset, children paused for 1 minutes and then received the second subset without further instruction. By providing the same time limit for both arithmetic subsets and by keeping the difficulty of the problems comparable over both subsets, we thus assessed the use of the commutativity shortcut. A commutativity benefit should show in more problems per time being solved in the subset containing commutative problems compared to the subset not containing such shortcut options. On the contrary, in case of more problems solved in the second, noncommutative subset, a general practice effect rather than the exploitation of commutativity would be evident.

The last task was the judgment task which began after a one minute break. Participants should, without a fixed time limit, solve as many problems as possible of the ten addition problems on the first (training) page. They were reminded in the instruction that they should work through the page from top to bottom again. The experimenter then explained that some of the problems just calculated could alternatively have been solved without calculation. Children were asked to look for such problems and to mark them when they felt they could get the answer without actually calculating the result. Again, participants had 2 minutes for that. After this training page, participants received the remaining two pages of this task. They were told not to solve the problems, but only to mark those that need no calculation. Again, 2 minutes were given for that.

12.2 Results

There were 26 children in the nonsymbolic, 33 in the symbolic condition, and 48 children in the condition that received an explicit instruction about commutativity. The result section will start with a description of the performance in the induction tasks, before turning to the computation task performance and the effect that the different inductions might have on it. After that, we will turn to children´s achievements in the judgment task and possible differences between the conditions in their conceptual knowledge, before finally reporting the integration of both kinds of knowledge.

12.2.1 Performance in the induction tasks

Each of the two induction tasks (the third condition included no induction but direct instruction) contained 14 trials, seven of which were commutative. There was no difference in the general number of solved trials between the conditions ($F[1, 57] = .151$, $p = .7$) and neither in accuracy ($F[1, 56] = 2.7$, $p = .11$). Also in this experiment, there thus seem to be no motivational differences in processing the approximation problems of two different formats.

Table 25. Experiment 3: Performance in the two inductions: General performance and accuracy relative to solved trials separately for the commutative and the noncommutative trials of the respective induction task in the two conditions

			Percentage correct relative to solved trials		
N	**Condition**	number of solved trials	complete task	only commutative	Only non-commutative
26	**nonsymbolic**	10.769 *(3.777)*	80.673 *(12.824)*	62.028 *(27.823)*	95.293 *(8.329)*
33	**symbolic**	10.424 *(3.052)*	87.170 *(21.558)*	85.281 *(30.546)*	88.499 *(21.058)*
59	**collapsed**	10.576 *(3.364)*	84.307 *(18.382)*	75.258 *(31.380)*	91.493 *(16.915)*

A 2 x 2 mixed-design ANOVA for the percentage of correctly answered trials with Condition as the between-subject and kind of trial (commutative vs. noncommutative) as the within-subject factor revealed a main effect of kind of trial: Noncommutative problems were solved more accurately in each group ($F[1, 56] = 22.395$, $p < .001$, $\eta_p^2 = .286$). However, there was also a significant interaction ($F[1, 56] = 15.156$, $p < .001$, $\eta_p^2 = .213$). Bonferroni-adjusted post-hoc t-Tests revealed that only in the nonsymbolic induction task, significantly more noncommutative trials were answered correctly ($t(24) = -5.379$, $p < .001$, $d = 2.196$), while this was not the case in the symbolic induction ($t(32) = -.673$, $p = .506$, $d = .238$).

Signature of the ANS in the approximation tasks. In both induction conditions, a 2 (condition) x 3 (ratio) mixed-design ANOVA showed a significant main effect of ratio ($F[2, 86] = 6.653$, $p = .002$, $\eta_p^2 = .134$). There was no interaction of the two factors ($F[2, 86] = .093$, $p = .911$). Thus, an effect of ratio was found that did not differ between the two conditions (see Figure 16). This indicates that in both tasks children actually had estimated their answers.

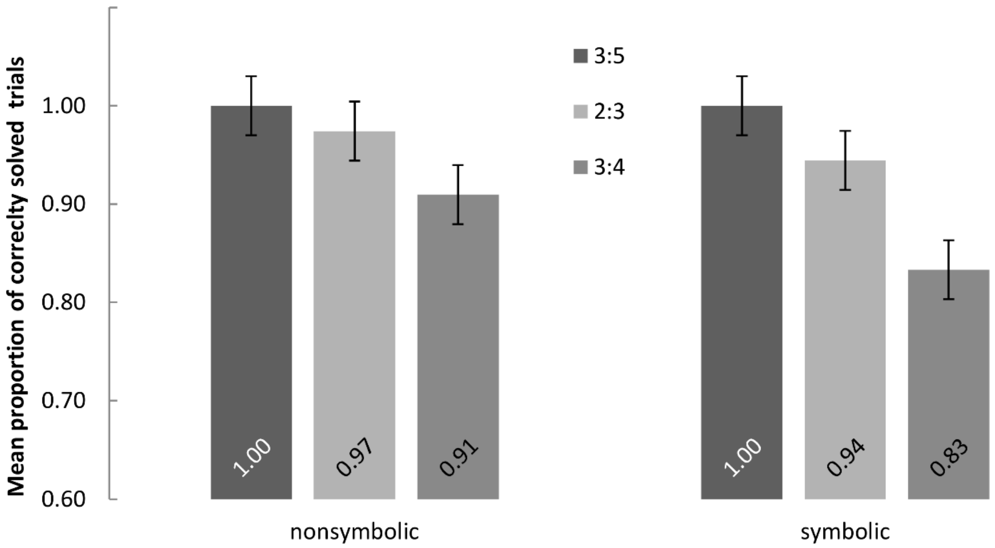

Figure 16. Experiment 3: Ratio effect in both inductions: Proportion correct solutions/solved problems separately for each ratio in each approximate task

12.2.2 Performance in the computation task (procedural knowledge)

In the computation task, we use the percentage of solved problems as our performance measure because we aimed at assessing the exploitation of commutativity as a shortcut. This exploitation would be reflected in transferring the result of one problem to its commuted counterpart, no matter if it was solved correctly. Thus, to assess commutativity usage independently of children´s general calculation proficiency, the number of solved problems (instead of percentage of correct solutions) is the more adequate measure.

On average, children solved more than half of the problems presented in the procedural knowledge task. In the collapsed sample, there was a small advantage for the commutative subset (see Table 26).

Table 26. Experiment 3: Procedural knowledge differentiated by induction: Mean percentage solved for each of the two subsets of the Procedural Knowledge Task (SD in brackets) separately for each induction condition, as well as the weighted difference between commutative and noncommutative problems (positive Difference values indicate use of commutativity)

N	Condition	Commutative Subset	Control Subset	Difference
26	**nonsymbolic**	54.615 *(16.655)*	55.00 *(14.974)*	-.115
33	**symbolic**	52.525 *(18.277)*	49.697 *(16.102)*	.848
48	**instruction**	58.333 *(17.592)*	54.722 *(19.091)*	1.083
107	**collapsed**	55.639 *(17.629)*	53.24 *(17.275)*	.72 *(2.903)*

Differences in procedural knowledge by induction

A one-way repeated measurement ANOVA of solved problems[10] with Subset as the within-subjects factor revealed a significant advantage for the commutative subset in the collapsed sample (F[1, 106] = 6.573, p = .012, η_p^2 = .058). According to our expectations, there was also a significant commutativity effect in the instruction condition (t(47) = 2.266, $p_{1\text{-}tailed}$ = .014, d = .661) and in the symbolic condition (t(32) = 1.786, $p_{1\text{-}tailed}$ = .0418, d = .632), but, surprisingly, not in the nonsymbolic condition (t(25) = -.275, $p_{1\text{-}tailed}$ = .393, d = .11).

Notable is the missing effect in the nonsymbolic condition. Although research provides an inconsistent picture so far, in our second study we had found nonsymbolic approximate arithmetic to be predictive of exact mathematical competencies (see also Gilmore et al., 2010). Hyde et al. (2014) were able to foster simple exact symbolic arithmetic with a nonsymbolic approximate addition task. Our current results however indicate that nonsymbolic approximation was not a suitable induction for an arithmetic principle like commutativity. Interesting is the comparison of the commutativity effect (and thus the beneficial influence of the induction) in the symbolic condition with the effect of an explicit instruction about commutativity. An according t-Test of the Difference measure (number of solved problems in the commutative subset – number of solved problems in the noncommutative subset) remained nonsignificant (t(79) = -.336, p =.738, d = .076), indicating that the mean profit from commutativity was of comparable size after an explicit instruction (M = 1.083, SD = 3.312) and after a symbolic approximate induction (M = .848, SD = 2.729). However, due to the unequal sample sizes in both conditions,

[10] We chose to use number of solved problems as dependent variable rather than mean solution time per problem as the latter is more vulnerable to outliers.

we assumed a power problem. Subsequent analysis with G*Power revealed an only 9% chance of detecting an according effect in the Difference measure. Thus, this finding of a comparable benefit through explicit instruction and symbolic approximate arithmetic has to be treated with caution. Nevertheless, it seems that of inductions that are based on approximation, to induce an arithmetic principle like commutativity, symbolic approximation is better suited than nonsymbolic approximation to foster at least procedural knowledge of that principle.

12.2.3 Performance in the judgment task (conceptual knowledge)

Conceptual commutativity knowledge was measured by the so-called judgement task as introduced by Haider et al. (2014). A concept of the additive law of commutativity was measured without explicit asking, but by assessing knowledge about the constraints of the principle more indirectly. Children should recognize that the only problems that – when looking from top to bottom at the page – needed no actual calculation, were problems that had a commuted counterpart. Children who do not understand the principle might rather think that each problem has to be calculated anew, or simply mark problems that for some reason seem easy to them. So we calculated the SDT measures of hits and false alarms (FA) to compare the according percentages, as well as the sensitivity measure d´ and a possible response bias c between the conditions.

Table 27 shows the results for the collapsed sample as well as separately for each of the three conditions.

Table 27. Experiment 3: Conceptual knowledge demonstrated in the judgment task: Mean rates of hits and false alarms (FA), sensitivity and response bias in the judgement task for the collapsed sample as well as separately for each condition (SD in brackets)

N	Condition	Hits	FA	Sensitivity d´	Response Bias c
26	**nonsymbolic**	74.359 *(27.98)*	33.654 *(29.952)*	2.13 *(1.928)*	-.234 *(1.187)*
33	**symbolic**	75.253 *(27.361)*	27.652 *(30.898)*	2.509 *(2.13)*	-.038 *(1.211)*
48	**instruction**	72.222 *(29.24)*	27.604 *(34.495)*	2.521 *(2.169)*	.037 *(1.416)*
107	**collapsed**	72.676 *(28.135)*	29.089 *(32.158)*	2.422 *(2.088)*	-.052 *(1.294)*

Data show that children actually are sensitive to the principle of commutativity. In the whole sample, the rate of hits is more than double as high as the false alarms rate. The sensitivity index d´ is a dimensionless measure and thus only interpretable in relation to the according values in the separate conditions. The negative c value indicates a general tendency to answer liberally, that is, to mark relatively many of the problems in general in the two induction conditions.

Differences in conceptual knowledge by induction. In each condition, conceptual knowledge was demonstrated by a significantly greater percentage of hits than false alarms, as was shown by a 3 (Condition) x 2 (Hits vs. False alarms) mixed-design ANOVA with number of marked problems as dependent variable ($F[1, 104] = 148.024$, $p < .001$, $\eta_p^2 = .587$). However, there was no significant effect of condition ($F[2, 104] = .239$, $p = .788$) or – more importantly – a significant interaction of the factors condition and 'hits vs. false alarms' ($p = .772$). This result is mirrored in an ANOVA of the sensitivity index d´. Although descriptively, sensitivity is highest in the symbolic and in the instruction condition (see Table 27), the according ANOVA remains nonsignificant ($F[2, 104] = .332$, $p = .718$). Thus, no reliable differences between conceptual knowledge in the three conditions and hence the possible influence of the different inductions can be deduced.

12.2.4 Relation between procedural and conceptual knowledge.

An additional possibility how a potential impact of our inductions might become visible is the integration of procedural and conceptual knowledge. Although conceptual knowledge might not have profited directly from the interventions tested here, it is possible that one or more of the inductions serve to integrate both kinds of knowledge, which might indicate an increasingly abstract concept of commutativity. The parameters hit rate, false alarm rate and sensitivity (d´) were included as indicators of conceptual knowledge, whereas the number of solved problems separately in each subset of the computation task as well as the resulting Difference value were put into the correlation matrix as procedural measures. An integrated concept of commutativity should be indicated by a significant correlation between the difference score and the sensitivity score d’. Table 28 presents the correlations for the three experimental groups separately, as well as collapsed across all conditions. As can be seen, in all experimental conditions there was no significant correlation between the difference score and d’. Even when collapsing the three conditions in order to increase power, no statistically significant correlation between d´ and the difference score was detected.

Table 28. Experiment 3: Integration of procedures and concepts: Correlation coefficients between procedural and conceptual knowledge in Experiment 3

N	Condition	judgment task	computation task Difference
26	**nonsymbolic**	d´	.188 *p* = .358
33	**symbolic**	d´	.016 *p* = .930
48	**instruction**	d´	.184 *p* = .210
107	**collapsed**	d´	.145 *p* = .137

12.3 Discussion

Our results provide strong hints that symbolic approximation is more suitable to induce an arithmetic principle like commutativity for a later exact context than nonsymbolic approximation. In our study, children who encountered an induction with nonsymbolic approximate arithmetic containing commutative trials showed no advantage in an exact symbolic task also interspersed with commuted problems afterwards. After symbolic approximation on the other hand, children actually showed a commutativity effect that was of comparable size as the effect after a direct instruction of the principle. So it seems that also symbolic approximation can be used to foster children´s skills in exact arithmetic. Hyde et al. (2014) found in their sample of first graders nonsymbolic arithmetic to be a suitable induction for exact symbolic addition problems. We extended this research question to arithmetic *principles* and found that in second graders, an approximation training with interspersed problems that allow the use of that principle could even induce the exploitation of the additive law of commutativity. This was only the case for *symbolic* approximation, though. However, neither of the inductions nor an explicit explanation of the principle led to a specific advantage in increasing *conceptual* commutativity knowledge. This is in line with several studies that also failed to demonstrate the effectiveness of diverse kinds of inductions to increase conceptual knowledge (Fyfe, Rittle-Johnson, & DeCaro, 2012; Matthews & Rittle-Johnson, 2009; Sherman & Bisanz, 2009).

The commutativity effects found in this experiment were mostly of only medium size. One explanation might be that second graders already are so proficient in solving addition problems that there was little need for them to exploit shortcut possibilities. Calculating a problem anew might not mean a notable disadvantage in solution time to them. So in the following experimental series (Experiments 4 a – 4 c), we tested different age ranges with the symbolic approximate arithmetic induction, starting with children at the beginning of first grade. We have demonstrated that at this age, children already perform above chance in a symbolic approximate arithmetic task (Studies 1 and 2, see also Gilmore et al., 2010), and consistently found no detrimental influence of SES on this measure. Thus, it seems promising to test it also at the beginning of formal education as an induction for principles that children already have some precursory knowledge of.

13 Experiment 4

We replicated Experiment 3 in order to further investigate the effect of symbolic approximation on procedural and conceptual knowledge of commutativity. Experiment 3 had indicated that symbolic approximation can trigger at least the procedural exploitation of an arithmetic principle. We thus repeated our testing with the symbolic approximate arithmetic induction in three age groups to narrow the age range in which symbolic approximation might be a suitable means of teaching[11].

To that aim, in Experiment 4 a we tested children who had just started school with slightly simplified versions of the symbolic approximate arithmetic induction as well as the computation task used in Experiment 3. In this experiment, we did not include the conceptual measure (the judgment task) because we did not want to overburden participants and furthermore assumed that in children who had just started school there would be no sufficient basic understanding to be fostered by an indirect induction. Nevertheless extending the testing of Experiment 3, we realized two conditions, one starting with the induction (the symbolic approximation task), and the other with the computation task. Thereby, we were able to test if the positive influence of symbolic approximate addition problems presenting the principle of commutativity is actually a unidirectional one, or if exact commutative problems might in turn also have a fostering effect on subsequent approximation.

In Experiment 4 b, we tested children at the end of first grade and used the same tasks as in Experiment 3, including the measure of conceptual commutativity knowledge. Besides to the two conditions already described for Experiment 4 a, we also instantiated an additional control condition which did not receive exact commutativity problems at all during the course of the experiment. This enabled us to disentangle the effects of the symbolic approximation task and exact commutativity problems on conceptual knowledge (see Chapter 13.2.1 for the more detailed description of the procedure).

Experiment 4 c was identical to 4 b except for the age group tested and some adjusted time limits. This time our participants were third graders, in order to on the one hand test if our symbolic approximation induction is also 'working' with children who had sufficient opportunity to practice the principle in question, and on the other hand to ex-

[11] The experimental series described in this chapter has been prepublished with the permission of the Dean of Research of the University Cologne (see Hansen et al., 2015).

plore if conceptual knowledge might be fostered by our induction in children who should already have a more stable foundation of conceptual commutativity understanding.

13.1 Experiment 4 a: Start of first grade

The main goal of Experiment 4 a was to investigate whether children who had not yet received any formal instruction about the commutativity principle would benefit from symbolic approximate arithmetic problems with respect to spontaneously spotting and applying commutativity-based shortcut options in exact arithmetic problems. For this purpose, we investigated first graders who had attended school for approximately four months and had not yet learned about commutativity in school. Half of the children started with the symbolic approximation task and then received the exact arithmetic problems (*approximation-first group*). The remaining children were administered to the reversed order of tasks; that is, they solved the exact arithmetic problems (computation task) first and then worked through symbolic approximation task (*computation-first group*). If the symbolic approximation task triggers the exploitation of commutativity in the exact arithmetic problems like we found in Experiment 3, children in the approximation-first group should show a larger commutativity benefit than the computation-first group.

13.1.1 Method

Participants. Sixty-eight (43 girls) first graders with a mean age of 6 years and 8 months (*SD* = 5.3 months) who had been attending school for four months took part in our study. We recruited them from one elementary school in a middle socio-economic status suburb of Cologne. All children had permission to take part in our study and were assigned evenly to the two experimental conditions. Thirty-six children (24 girls) participated in the approximation-first group, 32 children (19 girls) in the computation-first group.

Procedure. The procedure and time limits were basically the same as in Experiment 3. Due to the younger age of the participants, however, for Experiment 4 a, we shortened the symbolic approximation task from 14 to 11 trials (7 of which were commutative). Again, the addends ranged between 13 and 97 leading to results between 36 and 140. Furthermore, contrary to Experiment 3, the computation task here only contained two-addend instead of three-addend problems. In this youngest sample, we did not yet apply the judgment task because we did not expect children to have enough conceptual knowledge to foster it with approximation.

As a control condition we had another group of children solve the computation task first (computation-first group). By this we could ensure that the fostering effect of symbolic approximation problems was unidirectional instead of an unspecific warm-up of commutativity (against this argument, however, see also the differential impact of nonsymbolic and symbolic problems in Experiment 3).

13.1.2 Results

We excluded the data of children who did not follow instructions, for example by missing to start working on the task or trying to calculate the symbolic approximation problems (two children in the approximation-first and two in the computation-first condition). Furthermore, children were excluded who solved less than three problems in one of the computation subsets (i.e., two standard deviations below the mean). This concerned two children in the approximation-first and two children in the computation-first group. Thus, 32 children remained in the approximation-first condition, and 28 in the computation-first condition. We will report the influence of condition on the computation task first and then test if the influence is bidirectional. The latter would be mirrored in differential performance in the symbolic approximation task according to condition.

To test for the effect of the symbolic approximation task on the computation task, we conducted a 2 x 2 mixed-design ANOVA with Condition as the between-subject, Subset as the within-subject factor and with number of problems solved[12] within the given time as the dependent variable. We only found a significant interaction between Condition and Subset ($F[1, 58] = 6.31$, $p = .015$, $\eta_p^2 = .098$, for all other effects $F < 1$). Planned contrasts indicated that the approximation-first group exhibited a substantial commutativity effect ($F[1,58] = 6.54$, $p = .013$, $d = .33$), whereas the computation-first group did not ($F = 1.1$, $p = .30$, see the left panel of Figure 17). The findings did not change when only using correctly solved problems.

In order to investigate the effect of exact computation on the symbolic approximation task, we conducted a corresponding analysis for the symbolic approximation problems.

[12] Again, we chose to use number of solved problems as dependent variable rather than mean solution time per problem. We are aware that this makes the comparison between the age groups difficult but it is sufficient for our main goal of within age-group comparison.

The 2 (Condition) X 2 (Problem Type: commutative vs. noncommutative 'candy problems') mixed-design ANOVA with the proportion of correctly answered problems as the dependent variable did not yield any significant effect (all *F*s < 1, see also the left panel of Figure 18).

The results indicate that both conditions did not differ with regards to the symbolic approximation problems and thus the exact arithmetic tasks did not seem to affect the approximation performance (see Table 29). Furthermore, and in line with our results in the symbolic approximation condition of Experiment 3, there was no difference in accuracy for commutative and noncommutative 'candy problems' in the symbolic approximation task (Sherman & Bisanz, 2009).

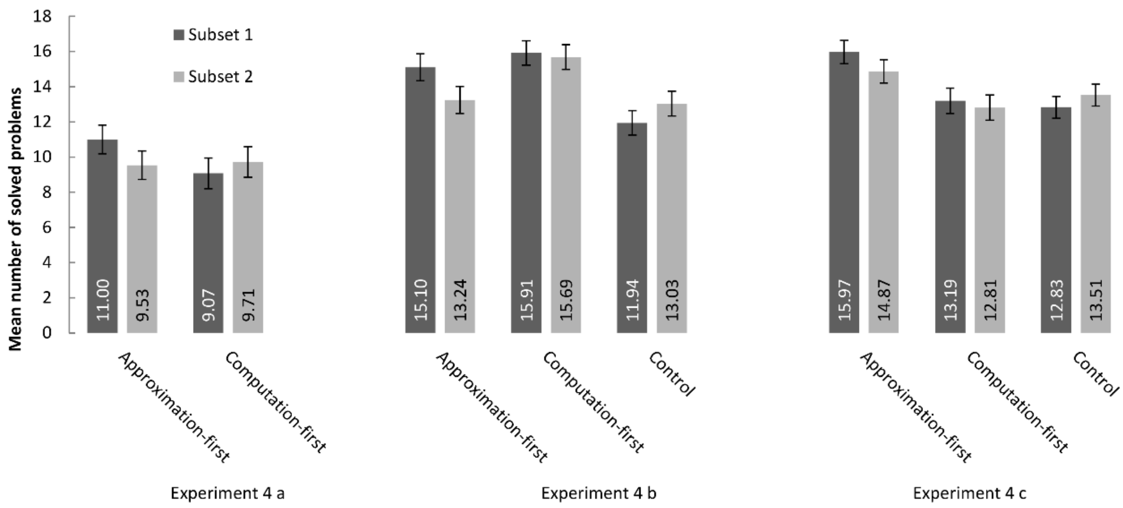

Figure 17. Experiments 4 a – c: Numbers of solved exact arithmetic problems. Within experiments, numbers of solved problems are depicted as a function of subset and condition. In the approximation-first and the computation-first conditions, Subset 1 refers to commutative problems, and Subset 2 to noncommutative baseline problems. In the control condition, subsets 1 and 2 both only contained noncommutative problems. Error bars reflect within-participants confidence intervals based on the MSe of the Condition X Subset Interaction (Loftus & Masson, 1994)

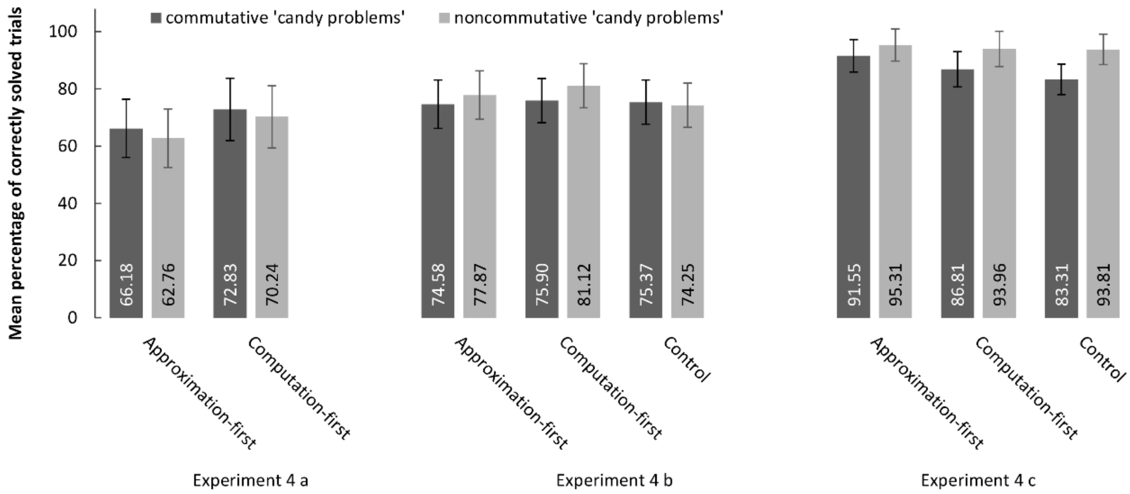

Figure 18. Experiments 4 a – c: Percentage of correctly answered trials in the symbolic approximation task. Within experiments, commutative and noncommutative problems are depicted separately for the approximation-first, the computation-first and the control group. Error bars reflect within-participants confidence intervals based on the MSe of the Condition X Subset Interaction (Loftus & Masson, 1994)

Table 29. Experiment 4 a: Performance of young first graders in the symbolic approximation task

N	Condition	*M* solved	% correct	% correct commutative	% correct non-commutative
32	**approximation-first**	6.81 *(2.46)*	65.15 *(30.84)*	66.18 *(40.56)*	62.76 *(34.91)*
28	**computation-first**	9.11 *(2.02)*	71.63 *(25.55)*	72.83 *(29.25)*	70.24 *(32.27)*

Mean number of solved symbolic approximation problems, rates of correct answers in general as well as separately for commutative and noncommutative 'candy' problems are depicted for both conditions (SD in parentheses). Participants had 1.5 minutes to solve 11 symbolic approximation problems.

13.1.3 Discussion

Experiment 4 a revealed that an induction phase with commutative and noncommutative symbolic approximate arithmetic problems increased even first graders' ability to spot and use the commutativity-based shortcut in exact addition problems. By contrast, and refining our results of Experiment 3, we did not find the reverse effect from exact addition problems to symbolic approximation problems[13].

Thus, symbolic approximation problems can not only enhance exact symbolic arithmetic performance, they can also trigger the subsequent use of an arithmetic principle by

[13] Note that our symbolic approximation task for Experiments 3 and 4 a - c was mainly constructed as an induction rather than as a measure of its own. Therefore, results concerning performance in the symbolic approximation task should be treated with caution (see also the discussion section).

children in less familiar, abstract addition problems (see also Sherman & Bisanz, 2009). It is important to note that this was the case even though our participants in this experiment had very little experience in formal addition and never had received any classroom instruction about the commutativity principle. Apparently activating precursory commutativity knowledge (i.e., knowledge about the order-irrelevance principle acquired before formal instruction in school) can help children to apply this knowledge to exact arithmetic problems. This confirmation of the findings in Experiment 3 again raises the question of whether approximate calculation might activate existing precursory *conceptual* knowledge in children who actually have heard about commutativity in school recently (which is the case in first grade in Germany), or if it influences the performance on a more general path by promoting flexibility in problem solving.

13.2 Experiment 4 b: End of first grade

The second experiment of this series aimed at testing the question of whether our symbolic approximation task either only influences procedural knowledge or, alternatively, can also affect the activation of conceptual knowledge of the commutativity principle. This time we tested slightly older first graders who had attended school for approximately nine months. In contrast to the group of first graders in Experiment 4 a, these children had already received classroom instruction about the commutativity principle. Thus, participants of Experiment 4 b should not only be trained in solving addition problems, but also should have (at least) some formal conceptual knowledge about commutativity. Note that the second graders who participated in our Experiment 3 did not show any specific conceptual benefit from the symbolic approximation induction, at least not one that goes beyond the effect of the nonsymbolic approximation induction, or direct instruction about the principle. However, it might be that – although they might not be making use of the principle consistently – these second graders already had too much conceptual commutativity knowledge to be fostered by a rather indirect induction like ours. So we decided to test children of an intermediate age group, that is, at the end of first grade. To assess conceptual knowledge, we again used the above mentioned *judgment task* (Haider et al., 2014; Appendix H) as an additional measure.

Experiment 4 b contained the three tasks already described for Experiment 3: (1) the symbolic approximation task, (2) the computation task, and (3) the judgment task. We expected to replicate the finding that activating precursory commutativity knowledge enhances the exploitation of commutativity during exact calculation (i.e., in the computation task). If our induction with the symbolic approximation task in this age group also activates conceptual knowledge, we should observe an impact of the symbolic approximation task on the number of correctly marked commutative problems in the judgment task.

As in Experiment 4 a, we realized the *approximation-first group* in which children started with the symbolic approximation task, and the *computation-first group* in which children began with the computation task. We also introduced an additional *control group.* Here, children started with the two computation subsets. However, none of the two subsets did contain any commutative problems. Participants were then given the symbolic approximation task and (like the approximation-first group) finished with the judgment task. The control condition had two functions: first, it helped us to assess the magnitude of the general practice effect when receiving two consecutive sets of arithmetic problems without the commutativity shortcut option. The computation-first group might have made use of the commutativity shortcut, but the benefit might have been occluded by a general practice effect. The control condition should allow us to differentiate the general speedup from problem subset 1 to problem subset 2 from the effect specific to the usage of commutativity knowledge. The second function was to measure the direct influence of the symbolic approximation problems on conceptual knowledge. In the approximation-first group, the effect of symbolic approximation problems on conceptual knowledge might be moderated by additionally encountering commutativity while calculating the problems of the computation task. Thus, there might be a direct and/or an indirect effect. However, as the control group received only noncommutative computation problems and, furthermore, worked through the symbolic approximation problems immediately before the judgment task, it exclusively measured the impact of the symbolic approximation task on the judgment task performance (conceptual knowledge).

13.2.1 Method

Participants. In Experiment 4 b, 131 first graders participated (55 girls, mean age of 7 years and 3 months, *SD* = 4.9 months). We recruited children from three elementary schools situated in middle socio-economic status suburbs of Cologne. Forty children (15 girls) were assigned to the approximation-first group, 45 children (17 girls) in the computation-first group, and 46 children (23 girls) were tested in the control group. All children had permission to join our study.

Procedure. The procedure was similar to Experiments 3 and 4 a. This time each participant received four booklets, one for each the symbolic approximation task and the judgment task, and the two subsets of the computation task. This time, we again used the symbolic approximation task with 14 trials as well as the three-addend computation task.

The *approximation-first group* started with the symbolic approximation task. Children were instructed to solve as many 'candy problems' as possible. The time limit was 2 minutes. A short break (2 minutes) followed. Afterwards the computation task started. All children began with the commutative subset and then received the noncommutative subset without any further instruction. The time limit was set to 4 minutes per subset (1 minute more than in Experiment 4 a, as children here received three-addend problems). As in Experiment 3, the judgment task was presented after a one minute break. Participants received 2 minutes to solve as many problems as possible of the ten addition problems on the first (training) page. Almost all children completed the ten problems before reaching the time limit. Children were asked to look for and mark those problems they felt they could get the answer without actually calculating the result. Again, participants had 2 minutes for that. After this training page, participants received the remaining two pages of this task. They were told not to solve the problems, but only to mark those that need no calculation. The time limit here was set to 3 minutes.

The *computation-first group* received the same three tasks in a different order. They started with the computation task, followed by the judgment task. Lastly, they received the symbolic approximation task.

The *control group* also started with the computation task. However, children here received solely noncommutative problems in *both* subsets. That is, they did not encounter any commutative problems in the computation task at all. Afterwards, children were given the symbolic approximation task and lastly worked through the judgment task.

13.2.2 Results

As in Experiment 4 a, the independent variables were experimental condition (task order) and task format: symbolic approximation task (Problem Type: commutative vs.

noncommutative 'candy problems'), computation task (Subset: subset 1 vs. subset 2[14]), and the judgment task. Again, our main dependent variable was the number of completed problems in the two subsets of the computation task. As dependent measures in the judgment task, we measured hits (correctly identified commutative problems), false alarms (incorrectly marked problems), as well as sensitivity index d' and response bias c from SDT (see Chapter 10.2.1). In our study, a response bias is liberal if a child marks many commutative and noncommutative problems as problems that need no calculation. The response bias is conservative, if a child only marks very few commutative and noncommutative problems as needing no calculation (Stanislaw & Todorov, 1999).

Again, children who did not follow instructions or who solved remarkably few of the arithmetic problems (less than two SDs below the group means) were excluded from further analyses. In addition, we also excluded children who marked each problem in the judgment task (11 children in the approximation-first group, ten children in the computation-first group, and 11 children in the control group). This led to 29 remaining children in the approximation-first group, 35 children in the computation-first group, and 35 children in the control group. As in Experiment 4 a, we first present the results of the computation and the symbolic approximation task, followed by the results of the judgment task, each also with regard to potential differences between the conditions. Finally, we describe the relationship between procedural and conceptual knowledge.

Performance in the computation task (procedural knowledge). The middle panel of Figure 17 depicts the number of problems solved in each of the two subsets for each of the three conditions. A 3 (Condition) x 2 (Subset) mixed-design ANOVA with number of solved problems as dependent variable yielded a significant main effect of Condition ($F[2, 96] = 5.00$, $p = .009$, $\eta_p^2 = .094$), as well as a significant interaction between Condition and Subset ($F[2, 96] = 7.97$, $p = .0006$, $\eta_p^2 = .142$). There was no main effect of Subset ($F[1, 96] = 1.27$, $p = .27$). The main effect of Condition was due to children in the computation-first group solving more problems than the participants in the control condition (revealed by Scheffé Test, $p = .009$, $d = .77$). This difference was unexpected. It might be due to a sampling error, even though children from different classes were randomly assigned to the experimental conditions.

[14] As the control group did not receive commutative problems in the first or second subset, we only refer to 'commutative' and 'noncommutative subset' when reporting on the approximation-first or computation-first group.

More importantly, planned interaction contrasts (Condition X Subset) revealed that the approximation-first group and the control condition differed significantly in the number of problems solved in the first vs. second subset ($F[1, 96] = 15.93$, $p < .001$; $d = .818$). The comparison between the computation-first and the control condition just failed the level of significance ($F[1, 96] = 3.49$, $p = .065$, $d = .381$). In addition, also the interaction contrast between the approximation-first and the computation-first condition was significant ($F[1, 96] = 4.89$; $p = .029$; $d = .453$). Thus, the expected Condition X Subset interaction again indicates that the approximation-first group benefited much more from the commutative problems than the computation-first condition. An additional analysis with only correctly solved problems as the dependent variable did not change the results.

Performance in the induction (symbolic approximation task). The 3 (Condition) X 2 (Problem Type: commutative vs. noncommutative 'candy problems') mixed-design ANOVA with the proportion of correct responses in the symbolic approximation task as the dependent variable revealed no significant effects (each $F < 1$; see middle panel of Figure 18). This finding indicates that all conditions performed equally well in the symbolic approximation problems. Practice on calculation problems apparently did not affect the performance in the symbolic approximation task (see Table 30). Again it has to be kept in mind that for the Experiments 3 and 4 a- c, the symbolic approximation task was constructed as an induction of commutativity, not as an instrument to measure approximation competencies. As it comprised only few commutative vs. noncommutative candy problems, we have to be cautious regarding its reliability.

Table 30. Experiments 4 b and c: Performance in the symbolic approximation inductions: Performance of older first graders and third graders in the symbolic approximation task

	Condition	***M* solved**	**% correct**	**% correct commutative**	**% correct non-**
Experiment 4 b					
N					
29	**approximation-first**	11.86 *(2.66)*	76.16 *(22.09)*	74.58 *(34.34)*	77.87 *(19.61)*
35	**computation-first**	12.14 *(2.60)*	78.95 *(20.81)*	75.91 *(31.80)*	81.12 *(23.55)*
35	**control group**	11.23 *(2.71)*	75.49 *(27.52)*	75.37 *(34.71)*	74.25 *(25.96)*
Experiment 4 c					
N					
31	**approximation-first**	13.39 *(1.48)*	93.37 *(9.04)*	91.55 *(14.26)*	95.31 *(8.66)*
26	**computation-first**	13.81 *(0.57)*	90.38 *(14.70)*	86.81 *(23.88)*	93.96 *(9.19)*
35	**control group**	13.43 *(1.4)*	88.56 (*16.33)*	83.31 *(28.40)*	93.81 *(10.62)*

Mean number of solved symbolic approximation problems, rates of correct answers in general as well as separately for commutative and noncommutative 'candy' problems are depicted for both conditions (SD in parentheses). Participants had 2 minutes to solve 14 symbolic approximation problems.

Performance in the judgment task. For the conceptual knowledge task, we first computed the hit rate (proportion of correctly identified commutative problems) and false alarm rate (proportion of incorrectly marked noncommutative problems) for each child individually. In addition, we computed the sensitivity index d' and the response criterion c. As can be seen from Table 31, the mean d' values did not differ very much between conditions. The corresponding one-way ANOVA with Condition as independent and d' values as dependent variable revealed no significant effect ($F < 1$). However, a closer look on the hit and false alarm rates in Table 31 also showed that the *general* frequency of marking problems in the judgment task varied considerably between the groups. Therefore, we also compared the conditions' mean response criteria. The response criterion c reflects participants' response bias with negative c scores indicating liberal, positive scores and a conservative response bias. The one-way ANOVA showed a substantial effect of Condition on participants' response criterion ($F[2, 96] = 3.72$, $p = .028$, $\eta_p^2 = .072$). A Scheffé Test revealed that the response criterion differed significantly between the approximation-first and the control group ($p = .033$, $d = -0.72$). Children in the approximation-first condition responded more liberally than children in the control condition. Thus, overall, the findings of the judgment task suggest that our symbolic approximation task, albeit it affected procedural knowledge, did not influence the conceptual commutativity knowledge of older first graders either (Sherman & Bisanz, 2009).

Table 31. Experiment 4 b: Performance of older first graders in the judgment task

N	Condition	Hits	False alarms	Sensitivity d´	Response Bias c
29	**approximation-first**	.78 *(.24)*	.49 *(.32)*	1.76 *(1.76)*	-.56 *(1.21)*
35	**computation-first**	.69 *(.35)*	.42 *(.29)*	1.47 *(1.68)*	-.24 *(1.47)*
35	**control group**	.58 *(.35)*	.28 *(.31)*	1.63 *(2.37)*	.32 *(1.23)*

Rate of hits and false alarms as well as the sensitivity index and the response criterion for each condition are depicted (SD in brackets).

Relation between procedural and conceptual knowledge. As indicators of procedural knowledge, we used the number of solved problems in the commutative and the non-commutative subset, as well as the difference scores between the two subsets. We included hit rate, false alarm rate, and the sensitivity scores (d') as measures of conceptual knowledge. An integrated concept of commutativity should be indicated by a significant correlation between the difference score and the sensitivity score d'. The control condition was excluded from this analysis since these participants did not receive any commutative problems in the computation task. Thus, no measure of procedural knowledge exists for this condition. Table 32 presents the correlations for the two experimental groups separately, as well as collapsed across both conditions. As can be seen from Table 32, both experimental conditions showed only small and non-significant correlations between the difference score and d’. Even when collapsing both conditions in order to increase power, no substantial correlation was detected.

Table 32. Experiment 4 b: Integration of procedures and concepts: Correlation coefficients between procedural and conceptual knowledge

N	Condition	judgment task	computation task Difference
29	**approximation-first**	d´	-.09 p = .628
35	**computation first**	d´	.08 p = .658
64	**collapsed**	d´	.01 p = .968

Correlation coefficients between procedural and conceptual knowledge for first graders, depicted separately for all participants and for the three conditions of Experiment 4 b. The difference score between the commutative and noncommutative Subset of the computation task indicates use of commutativity with positive values.

13.2.3 Discussion

Experiment 4 b yielded two main results: first, we could replicate our findings of Experiment 3 and 4 a. In comparison to the control group, both the approximation-first and the computation-first conditions demonstrated at least some procedural knowledge of commutativity as measured in the computation task. However, the approximation-first condition profited significantly more from commutative problems than the computation-first condition. Thus, Experiment 4 b also indicated that the symbolic approximation task facilitated the use of the commutativity shortcut when solving exact arithmetic problems.

Second, this benefit was restricted to procedural knowledge – as it was in our Experiment 3 and also in the study of Sherman and Bisanz (2009; see also Rittle-Johnson, 2006). Conceptual knowledge was not enhanced. If at all, presenting the symbolic approximation task first slightly liberalized first graders’ response criterion in the judgment task. One possible explanation seems to be that the experience of not having to produce an exact result to mathematical problems gave the children the impression that this can apply to more or less any arithmetical problem. However, we did not find this effect in the control condition in which symbolic approximation was administered directly before the judgment task without the intermediary computation. So, there seems to be no or at least no direct causal link from isolated symbolic approximation to judging more liberally afterwards.

Finally, we did not find a substantial correlation between the application of the commutativity shortcut and conceptual knowledge in the approximation-first condition or in the computation-first group, as reflected in children's sensitivity in the judgment task. On the one hand, this result suggests that our symbolic approximation task is not suitable to boost the integration of both knowledge types in first graders. On the other hand, this missing correlation replicates our results of Experiment 3 with second graders, and is in line with the findings of Canobi et al. (1998; 2002) or Haider et al. (2014) who found first measureable integration of procedural and conceptual commutativity knowledge in second respectively third graders. Therefore, to maximize the chance of finding a possible beneficial influence of our symbolic approximation task on integration of conceptual and procedural knowledge in Experiment 4 c, we tested whether our symbolic approximation task might affect conceptual knowledge in third graders.

13.3 Experiment 4 c: Third grade

The main goal of Experiment 4 c was to replicate our findings with third graders. In addition, we tackled the question of whether our symbolic approximation task would enhance conceptual knowledge when participants possess more basic conceptual knowledge about commutativity, and, furthermore, if it is suited to boost the integration of procedural and conceptual knowledge about commutativity. Haider et al. (2014) found first signs of an integrated concept among third graders. Therefore, we surmised that if our symbolic approximation task affects conceptual knowledge about commutativity, we should be able to find a similar effect within this age group.

13.3.1 Method

Participants. One hundred and six (58 girls) third graders with a mean age of 8 years and 6 months (*SD* = 10 months) were recruited from three primary schools located in different middle socio-economic status suburbs of Cologne. Thirty-six (20 girls) children participated in the approximation-first group, 32 (15 girls) children in the computation-first group, and 38 (23 girls) in the control group.

Materials and Procedure. Except for adjusted time limits, materials and procedure were identical to Experiment 4 b. That is, we assigned children to one of three groups: the *approximation-first*, *computation-first*, and *control condition*. The time limit in the symbolic approximation task was unchanged; the other time limits were adopted for the third graders according to time demands estimated based on Haider et al. (2014). That is, we granted 3 minutes to solve each of the two computation subsets and 2 minutes for the judgment task.

13.3.2 Results

Employing the reported exclusion criteria, the data of five children in the approximation-first group, of six children in the computation-first group, and of three children in the control group were excluded from further analyses. This led to 31 remaining children in the approximation-first group, 26 in the computation-first group and 35 in the control group.

Performance in the computation task (procedural knowledge). A 3 (Condition) X 2 (Subset) mixed-design ANOVA with the number of solved problems as dependent variable revealed a significant main effect of Condition (F[2, 89] = 3.53, p = .034, η_p^2 = .073), as well as a significant Condition X Subset interaction (F[2, 89] = 3.89, p = .024, η_p^2 = .08). Figure 17 shows that the effect of Condition was due to the approximation-first group solving more problems than the other two conditions. As in Experiment 4 b, we assume that this is a sampling effect.

Planned interaction contrasts (Condition X Subset) only revealed a significant difference between the approximation-first and the control group (F[1, 89] = 7.64, p = .007, d = .587; Fs < 1.5 for the other two contrasts). As in the previous experiments, the results did not change when we restricted our analysis to only correctly solved problems.

Performance in the induction (symbolic approximation task). The corresponding 3 (Condition) X 2 (Problem Type: commutative vs. noncommutative 'candy problems') mixed-design ANOVA for the symbolic approximation task only revealed a main effect of Problem Type (F[1, 89] = 9.26, p = .003, η_p^2 = .094). This finding was due to more correctly solved noncommutative problems. Thus, third graders did not show a positive influence of the computation problems on the symbolic approximation task either – at least not with a symbolic approximation task set up as an induction rather than as a sensitive measure.

Performance in the judgment task. Table 33 shows the hit and false alarm rate, as well as the sensitivity scores d' and the response criterion c. A one-way ANOVA on d' scores yielded no significant effect of Condition ($F < 1$). Thus, the three conditions did not differ regarding their sensitivity in the judgment task.

Again, the data of hit and false alarm rates suggests that children in the approximation-first group tended to respond more liberally than children in the other two conditions. However, the one-way ANOVA with the response criterion c as the dependent variable showed no significant differences between the conditions ($F < 1$).

Table 33. Experiment 4 c: Performance of third graders in the judgment task

N	Condition	Hits	False alarms	Sensitivity d´	Response Bias c
31	**approximation-first**	.75 *(.21)*	.38 *(.32)*	1.85 *(2.32)*	-.24 *(.98)*
26	**computation-first**	.72 *(.31)*	.28 *(.32)*	2.47 *(2.81)*	.11 *(1.01)*
35	**control group**	.71 *(.27)*	.30 *(.30)*	2.30 *(2.06)*	-.002 *(1.08)*

Altogether, also third graders benefited from the symbolic approximation task when asked to solve arithmetic commutative and noncommutative problems. This finding is less clear than in the former two experiments, as the interaction contrast between the approximation-first and the computation-first conditions was not significant. However, only the approximation-first group differed significantly in their commutativity effect from the control group, and only in this group children solved significantly more problems in the first (the commutative subset), compared to the second subset (the noncommutative subset) at all ($F[1, 89] = 5.46$, $p = .023$, $d = .239$ in the approximation-first group; $F < 1$ for the computation-first group). As in the former experiments, this positive effect of the symbolic approximation task was restricted to procedural knowledge only. Even though third graders should possess more conceptual knowledge than first graders, we did not succeed in fostering their conceptual commutativity knowledge by means of symbolic approximation problems.

Relation between procedural and conceptual knowledge. As we did in Experiment 4 b, we also analyzed if procedural and conceptual commutativity knowledge is related. A positive correlation between these types of knowledge would indicate the formation of an increasingly abstract concept. Table 34 depicts the correlation coefficients for the approximation-first and the computation-first group, as well as collapsed across both conditions. As can be seen from Table 34, the correlations between the difference scores (procedural knowledge) and the sensitivity scores d' (conceptual knowledge) are small and non-significant among third graders as well.

Table 34. Experiment 4 c: Integration of procedures and concepts: Correlation coefficients between procedural and conceptual knowledge

N	Condition	judgment task	computation task Difference
31	**approximation-first**	d´	.14 *p = .441*
26	**computation first**	d´	.09 *p = .669*
57	**collapsed**	d´	.10 *p = .475*

Correlation coefficients between procedural and conceptual knowledge for third graders, depicted separately for all participants and for the three conditions of Experiment 4 c. The difference score between the commutative and noncommutative Subset of the computation task indicates use of commutativity with positive values.

13.3.3 Discussion

The pattern of results in Experiment 4 c is quite similar to that found for the younger children in Experiments 4 a and 4 b: first, children who started with the symbolic approximation task were more likely to spot and apply the commutative shortcut in the computation problems than children who received the computation task at first. By contrast, the reverse effect, a potential impact of solving arithmetic problems on approximation, was not observed. Second, third graders' conceptual knowledge was not altered by the symbolic approximation induction. While the d'-scores of third graders were higher than those of first graders, reflecting a higher familiarity with the commutativity principle of addition, symbolic approximation problems had no impact on later judging whether or not arithmetic problems could be solved without computation. The liberalization effect found for the approximation-first condition in Experiment 4 b was only descriptively replicated in Experiment 4 c. Consequently, as the effect does not seem to be very robust, we do not provide further speculations. However, it might be worthwhile to take up this issue in future research. Again, we did not find any sign of a better integration of procedural and conceptual commutativity knowledge in the approximation-first condition. Thus, our data once again suggests that the symbolic approximation induction solely enhanced the application of the commutativity shortcut when solving arithmetic problems.

14 Discussion for Experiments 3 and 4

In four experiments (Experiment 4 consisted in a series of three studies), we explored if symbolic approximate arithmetic that contains the possibility to detect the additive law of commutativity can increase the subsequent spontaneous usage and understanding of commutativity in exact symbolic arithmetic problems as encountered in school. There is now growing consensus that children with little numerical experience are able to master nonsymbolic or symbolic approximate addition problems with large addends as long as no exact calculation is required (see our Studies 1 and 2, or Barth et al., 2005; Gilmore et al., 2007; Piazza, 2010). In addition, some recent findings suggest that elementary school children can benefit from approximate nonsymbolic problems in their subsequent performance on exact symbolic arithmetic problems (Hyde et al., 2014). However, we know of only a few studies that investigated potential effects of activating early numerical competencies on the understanding of more abstract arithmetic principles as, for instance, inversion or commutativity (Sherman & Bisanz, 2009). Here, we focused on the question of whether activating precursory commutativity knowledge (Resnick, 1992) through symbolic approximate arithmetic problems will boost the exploitation of commutativity-based shortcuts in exact arithmetic problems.

Our experiments yielded three main results: first, the symbolic approximate arithmetic task in fact increased the probability for children to apply the commutativity shortcut in symbolic exact arithmetic problems, while nonsymbolic approximate arithmetic did not (Experiment 3). Furthermore, we tentatively conclude that this influence seemed to be unidirectional since there was no comparable effect of solving arithmetic problems on the symbolic approximation task (Experiments 4 a – c). Second, not even in our oldest participants had symbolic approximation a positive effect on conceptual knowledge. Third, the positive effect of commutativity-related symbolic approximation on spotting and applying commutativity-based shortcut options in symbolic exact arithmetic problems was already observed in children who had not yet received any classroom instruction about the commutativity principle in school. Our findings thus suggest that letting children explore the mathematical principle of commutativity in approximation problems activated procedural precursory mathematical knowledge (Gilmore & Spelke, 2008; Resnick, 1989). This activation was sufficient to trigger the use of the shortcut during calculation. That is, when confronted with the principle in the context of symbolic approximation problems, children might have realized or might have been reminded that an important strategy in arithmetic is to attend to the addends and to compare them within and between problems (Siegler & Booth, 2005). If addends are identical, the results of the problems are also identical. Applying a shortcut strategy like that differs from understanding the abstract mathematical concept of commutativity. It does not necessarily refer to the cardinality principle and it also does not necessitate metacognitive awareness of that one is no longer calculating when solving the problems either. The shortcut might simply be recognized as a helpful and labour-saving strategy when children are asked to calculate problems that follow the commutativity principle. This assumption might explain why we found reliable transfer from symbolic approximation to exact computation problems (procedural knowledge of commutativity), but not from symbolic approximation to judging arithmetic problems (conceptual knowledge of commutativity). Thus, we conclude that our approximation induction mainly triggered procedural knowledge and increased the flexibility of applying different strategies.

Our findings are in line with those of Matthews and Rittle-Johnson (2009) who activated procedural knowledge via instruction and subsequently found a strategy transfer to unfamiliar problems, but no gains in conceptual knowledge. Also Sherman and Bisanz (2009) or Fyfe et al. (2012) did not find any transfer from procedural to conceptual knowledge. Therefore, it seems justified to conclude that letting children explore a principle on the representational level of the ANS is effective for boosting the use of correct computational strategies and principle-based shortcuts that facilitate respectively supersede computation. However, the activated knowledge does not seem to be conceptual in the sense of an explicit representation of the abstract mathematical principle of commutativity (see also Sherman & Bisanz, 2009). It might be that the exploration of commutativity activates implicit knowledge representations rather than explicit conceptual knowledge about the abstract principle, or that the activation of the strategy after our induction is simply not sufficient for the children to become consciously aware of it (Siegler & Stern, 1998; Siegler, 2000). As an alternative interpretation of our results, one could argue that the impact of our symbolic approximation induction was simply due to an unspecific warm-up effect in the approximation-first conditions. The computation-first conditions in our experimental series 4 always started with the commutative subset of the computation task, whereas the approximation-first conditions received this subset after having solved the symbolic approximation problems. Thus, an unspecific warm-up effect seems plausible and would also explain the missing effect of our induction on conceptual knowledge as both groups would have been 'warmed-up' at this point. However, two arguments speak against this alternative explanation: first, an unspecific warm-up effect should have increased the overall number of problems solved in the approximation-first group compared to the computation-first condition. Obviously, this was not the case in our experiments: the conditions in Experiment 4 a did not differ at all in the overall number of arithmetic problems solved. In Experiment 4 b, the computation-first group solved more problems, and only in Experiment 4 c, it was the approximation-first group. Second, in Experiment 3 we had compared the effects of two commutativity-based induction tasks on subsequent strategy-use in second graders. If such an induction phase serves as a general warm-up, one should expect no difference between these groups. However, the exploitation of the commutativity shortcut in the computation task did not occur after the nonsymbolic approximate arithmetic induction, although this task should also have provided ample opportunity for warm-up and, more importantly, contained the same commutative problems. Thus, the alternative assumption of an unspecific warm-up effect seems rather unlikely to explain the current findings.

Why then is conceptual knowledge unaffected by the symbolic approximation induction? And why is there little integration of conceptual and procedural knowledge? So far, there seems to be no consensus on how to foster the development of abstract mathematical concepts. Our results of Experiments 3 and 4 suggest that experience with formal instructions in school does not seem to 'do the job': Neither the first graders in our Experiment 4 b, who had explicitly been taught the commutativity principle some months before participating in our experiment, nor the second and third graders of Experiment 3 and 4 c, who had received such instruction one, respectively two years before and afterwards spent much time on practicing basic arithmetic, showed any stable relationship between their strategy use and their ability to recognize the commutative problems in the judgment task, nor did they perform at ceiling level in these tasks. This is especially noteworthy given the findings of McNeil and Alibali (2000). They found that focusing on practice and correctly applied procedures during the initial learning of a mathematical principle led to less direct benefits in conceptual knowledge in third- and fourth-grade children than when conceptual guidance was involved. Interestingly, some weeks later the procedural conditions of the experiments had caught up in their conceptual understanding! So it is possible that a conceptual gain from the application of specific strategies needs much more time than what was given in our experiments (see also Dean & Kuhn, 2007). Another account for our results could be that our judgment task was not reliable and/or the according instruction might have been misleading. However, Haider et al. (2014) tested the reliability of the instrument and found satisfying split-half reliability coefficients between .78 and .83 for elementary school children. Also, the second argument – the instruction to the judgment task might have been misleading – does not seem to apply. First, when instructing the children, there was no indication that they did not understand the instruction. Second, there actually were some children in each sample who displayed perfect sensitivity in the judgment task and the number of these children increased from first to third graders. Of course, it is possible that some children drew on different ideas and concepts in trying to master the judgment task, but only relying on the principle in question – the additive law of commutativity – would result in the right answers and thus be measured as conceptual knowledge. Our material did not incorporate any other shortcut option that would result in a comparable benefit like exploiting commutativity. Third, Haider et al. (2014) collected data of adult students. These participants showed near perfect knowledge in the judgment task. Given these arguments, it seems justified to conclude that conceptual knowledge might emerge at a later point in development and could develop independent of procedural knowledge. The missing correlation in the current experiments was indeed due to the fact that children in the current study who were able to correctly mark all commutative problems did not show large benefits of commutativity in the computation task. Vice versa, children who showed large benefits of commutativity during calculation were not necessarily able to correctly mark the commutative problems in the judgment task. Thus, it seems that, at least in our study, the competencies assessed in the computation and the judgment tasks are more or less independent.

Our findings additionally showed that first graders in Experiment 4 a, who never had been taught the commutativity principle in a formal context before, already benefited from the approximation task. This adds to the findings of Fyfe et al. (2012) who studied the interplay of exploration and instruction in second and third graders. The authors found that the explicit instruction of a novel principle led to a higher usage of procedural knowledge when children could explore the task material without any feedback beforehand (as compared to when feedback was provided during exploration). Note that this was only the case for children who demonstrated some strategy knowledge before the exploration. So, it seems plausible that precursory procedural knowledge enabled our participants to benefit from the approximation task that was also administered without further guidance. This indicates that even an abstract arithmetic principle like commutativity, for which children possess informal precursory knowledge from everyday life, can be induced without any verbal explanation. Our Experiment 3 also supports this suggestion. The results showed that an explicit verbal explanation about commutativity alone did not elicit a larger procedural benefit than the approximation induction and, more importantly, this explicit instruction did not foster conceptual knowledge either. Several findings already provided evidence for the existence of precursory knowledge in mathematics, and suggested that children understand basic arithmetic (i.e., addition or subtraction) as long as the tasks are carried out approximately (Barth et al., 2005; Burr et al., 2010; Cantlon et al., 2009; Gilmore et al., 2007, 2010; Gilmore & Spelke, 2008; Knops et al., 2009; Piazza, 2010; Xu & Spelke, 2000).

In addition, in our Studies 1 and 2, we confirmed existing evidence that these approximate competencies even predict later math performance in school (Mazzocco et al., 2011; de Smedt, Verschaffel, & Ghesquière, 2009). A few other studies also indicated that activating precursory arithmetic concepts by inductions can facilitate children's exact symbolic calculation performance (e.g. Hyde et al., 2014, see also Chapter 3.2). Our current study extended these results by providing evidence that an approximation induction not only positively affects calculation and number processing, but also the spontaneous exploitation of a *specific abstract arithmetic principle* in a symbolic exact representational format even when no hint about its existence was provided beforehand. Another important point regarding procedural knowledge is that, without the symbolic approximation induction, none of the conditions tested in the current studies displayed a benefit from commutative problems, at least not to a degree that led to significant differences between commutative and noncommutative problems. Even when comparing children who only received noncommutative problems (control group) with the computation-first condition, the benefit of commutative problems was not significant. Thus, this finding supports evidence that children up to third grade do not consistently spot and use the commutativity shortcut when receiving no hint about the existence of commutative problems (Gaschler et al., 2013; Hannula & Lehtinen, 2005; Hannula, Lepola, & Lehtinen, 2010).

Altogether, the current results show that symbolic approximation problems can not only help to enhance general number processing or the execution of simple arithmetic (Hyde et al., 2009; Obersteiner et al., 2013) but also the use of a quite abstract and specific arithmetic principle. It strongly suggests that children already possess precursory knowledge about the principle of commutativity when entering school. They can rely on strategies derived from this knowledge, but seem to need external triggering to activate them, for instance with symbolic approximate calculation (e.g. Gilmore et al., 2007). Abstract conceptual knowledge in terms of an integration of conceptual and procedural knowledge about the commutativity principle seems to develop later and probably independent of such precursory knowledge (e.g., Verschaffel et al., 1994). This might be one reason why it seems so difficult to enhance this conceptual knowledge.

15 General Discussion

The theoretical frame of this thesis emerged from a common confounding of the two factors 'demanded accuracy' and 'format' in current research. We addressed the superordinate question if in numerical competencies of children just starting formal education, one of these two factors is dominating in determining their performance, interrelations between different skills, and susceptibility to socioeconomic influences. Thus, our goals were to investigate the potential of symbolic approximate arithmetic to predict and to foster elementary school children´s formal math performance, as well as to test this format for its robustness against socioeconomic influences. For nonsymbolic approximate arithmetic, both has been frequently found. We assumed the symbolic approximate arithmetic task to be of similar value in these regards due to an important proposition of the Triple Code Model (Dehaene, 1992). According to the TCM, approximate processing of quantities is carried out in the fundamental analogue magnitude code, the ANS, no matter if presented nonsymbolically or as Arabic numerals. Exact processing in both formats, however, has to be performed in one of the two symbolic codes proposed in the model. Despite frequent reference to the TCM, studies that investigate approximation and compare it to exact calculation mostly do only realize approximation in a nonsymbolic, and exact processing in a symbolic format. A systematic realization of all 2x2 combinations (approximate symbolic, approximate nonsymbolic, exact symbolic, exact nonsymbolic) in one sample was missing, at least for the age group of elementary school children and for approximate arithmetic.

We started to fill this academic gap by conducting two longitudinal studies (Study 1 and 2) in which we accompanied children´s first year of school. At the beginning of grade 1, we assessed their performance in exact and approximate arithmetic each in both a symbolic and nonsymbolic format, and at the end of term we administered a standardized math test. This enabled us to investigate correlations between exact and approximate arithmetic skills of both formats, as well as predictive relationships between the ANS-based abilities and formal math performance. We furthermore assessed children´s SES and thus were able to look into which of the skills are prone to detrimental socioeconomic influences and which are not.

In a second, experimental series (Experiments 3 and 4 a – c), we tested if an induction of symbolic approximate arithmetic containing problems that were connected via a specific arithmetic principle would facilitate the use and understanding of that principle in a more formal, symbolic exact context as typically encountered in school.

In the following, the hypotheses and research questions are discussed together with our results. Although the TCM provides a general theoretical framework, the literature concerning each of the research questions and hypotheses does only in part overlap. So we will fit our results in the respective literature at first, before at the end of this chapter we return to the TCM.

In the longitudinal studies, we investigated (1a) the pattern of intercorrelations between the performances in the 2x2 tasks as well as (1b) two competing hypotheses concerning the influence that children´s SES would have on their achievements in the different problem combinations. We furthermore tested (2a) the hypothesis that symbolic approximate arithmetic should be at least as good a predictor for formal math performance at the end of first grade as nonsymbolic approximate arithmetic, and (2b) that this predictive relationship should not be fully explainable by a mediation through numerical ordering ability.

In the experimental series, based on the results of the longitudinal studies, we explored (3) the suitability of symbolic approximate arithmetic to induce the mathematical principle of commutativity for a subsequent more formal context. We compared its effect to that of an according nonsymbolic induction and to an explicit instruction of the principle. Furthermore, we investigated if procedural, conceptual, or both kinds of knowledge might profit from one of the tasks or the instruction. After establishing the specific efficiency of the symbolic approximate arithmetic task in Experiment 3, we replicated the experiment in three age groups in Experiments 4 a – c, ranging from the very beginning of first grade to third grade.

Our first two questions were about the interrelations between first grader´s arithmetic performances, depending on the demanded accuracy (approximate vs. exact) and the format these were presented in (nonsymbolic vs. symbolic); as well as which of these skills would be negatively affected by a low SES. For the latter, in two competing hypotheses we investigated two potential mechanisms that could explain how SES might gain its impact on children´s numerical abilities. If differences in home environment lead to poorer symbolic number knowledge (Gilmore et al., 2010, Jordan et al., 1992), one should find that mostly those tasks presented as Arabic numerals would vary with children´s SES. If on the other hand differences in math performance according to SES emerge because of a mediating role of linguistic skills, which have been found to correlate with both SES and exact math performance measures, we would expect specifically the performance in the exact tasks to differ with SES, no matter if presented symbolically or nonsymbolically. After investigating these aspects, one could tentatively applicate the results onto the superordinate question if early numerical abilities are distinguishable (regarding their difficulty and their vulnerability to socioeconomic influence) according to one specific factor. And if so, is this factor the format a problem is presented in (as silently suggested by for example Gilmore et al., 2010), or might it rather be the demanded accuracy of the operation (as implicated in the TCM, Dehaene, 1992; see also Mejias & Schiltz, 2013)?

In the first study, rather clear lines emerged between approximate and exact performances, with only the latter varying with SES. Furthermore, besides a marginal significant correlation between both approximate tasks and the exact nonsymbolic measure, significant intercorrelations only existed within the approximate tasks, as well as within the exact tasks. Study 2, however, demonstrated that the two questions of intercorrelations and the impact of SES must be considered in combination. Here, all performances were intercorrelated, with the strongest connections displayed whenever the exact nonsymbolic measure was part of the correlation. Partialling out this dominant common influence left only a marginally significant correlation between both symbolic arithmetic tasks. At first this looked contradictive to the correlations obtained in the first study. Besides an important integrative function of the ability to exactly manipulate quantities without the need to understand Arabic numerals (exact nonsymbolic arithmetic) that had only been adumbrated in Study 1, it here seems that beyond that, symbolic number knowledge connects children´s numerical abilities. In the first study, it had rather looked like demanded accuracy would be the more crucial factor in children´s performances. As not only our SES measure but also the distribution of low and high SES in the respective sample differed between the studies, a plausible reason for his divergence might be that the SES composition in Study 1 was not diverse enough. There had been an overproportion of high-SES participants. Therefore, we conducted partial correlation analyses for the low- and the high-SES participants of Study 2, and indeed the pattern in the high-SES subsample resembled the one found in Study 1. Their performances in both approximate tasks correlated significantly with each other, as well as the performances in both exact tasks. In this subsample, there was no correlation between exact symbolic arithmetic and the approximation tasks anymore. However, we found large positive correlations between nonsymbolic exact processing and the approximate tasks. In the low-SES subgroup, on the other hand, the strongest connections were again between nonsymbolic exact processing and all other measures, but also high correlations between symbolic exact arithmetic and all (also the approximate) tasks were found. Between the approximate measures however, there was no significant relationship.

Because we had found the distinct correlations of exact nonsymbolic processing to approximate measures in both subsamples, we again computed partial correlations without the common variance of this measure. Without its integrative influence, in the low-SES subsample only the correlation between both symbolic measures remained significant, whereas it was the correlation between both approximate measures in the high-SES subsample.

Something else had changed in the results of Study 2 compared to Study 1. Besides the exact measures, in Study 2, also the scores in the approximate nonsymbolic task were worse in low- than in high-SES children. We think that this results from greater practice and progress of the high-SES participants in this problem format. The pattern of performances in the 2x2 tasks in this subsample supports this notion. They differed significantly in their accuracy in both approximate tasks, as well as in both exact tasks. But there was no main effect, that is, no general advantage for nonsymbolic (or symbolic) problems. Instead, it was always the task we would describe as the more 'common in everyday life' (which is also mirrored in the confounding that was the starting point for our research): In approximation, this is the nonsymbolic, and in exact processing it is the symbolic task in which high-SES children showed better performance. This indicates that experience with the task formats plays an important role in our results. While first graders typically strive to become accomplished in exact calculation, estimating results of problems incorporating Arabic numerals might seem like a setback to them. Approximating nonsymbolic quantities probably does not induce the demand of exact calculation to an equal degree as approximating Arabic numerals. From this it follows that high-SES first graders, who have improved all numerical abilities tested in the other three main tasks, have the least training in symbolic approximation. In this task, their performance is on the level of their low-SES peers. But why did we not find the correlation between approximate nonsymbolic arithmetic and SES in the first study? We attribute this missing finding to the smaller sample size in Study 1, but also to the greater diversity of SES in our sample of Study 2. The skills measured in the various tasks might be more or less sensitive to SES. Exact processing might be so sensitive to socio-economic environment that an according effect shows also within the comparably homogeneous and small sample of Study 1, while in the approximate nonsymbolic task, it needs bigger differences in children´s home environment to produce a measurable difference (see also Chen & Li, 2014; or Vogel et al., 2016). This was the case in our second sample.

Our results support the claim of the TCM that approximate processes are carried out in the same code, while exact numerical processing 'happens' in a different one. In Study 1 this was visible in intercorrelations mainly existing between tasks that demanded the same degree of accuracy. However, our second study matches the result that Mejias and Schiltz (2013) had reported for their 5-6 year old sample. They had realized the 2x2 design for the case of number representation and found that in this age group, all tasks were intercorrelated. Our results are in line with Mejias and Schiltz´s conclusion. They suggest that after the ANS being relatively robust against external influences at first, the integration of approximate and exact number representation emerges late in development, that is, shortly before or when starting formal education. This is mirrored in our second study by the intercorrelations between all main tasks. That the integration of approximate and exact skills is driven by increasingly exact abilities calibrating one´s ANS, as Kolkman et al. (2013) suggested, becomes also visible in the superordinate role of exact nonsymbolic arithmetic demonstrated in our Study 2 (and adumbrated in Study 1). This task was the one responsible for the intercorrelations between exact and approximate processing; indicating its integrative role. Although this task showed strong intercorrelations to the other performances of both low- and high-SES participants, it was itself subject to an influence of children´s SES, with low-SES participants performing significantly worse. This contradicts current accounts that only take poor symbolic number knowledge into account as the mediating source between the socioeconomic situation and low-SES children´s worse numerical performances (Jordan et al., 2008). It rather seems that also language skills, which should affect exact processing of any kind, play a role in how SES gains impact on numerical abilities (see Noble, Norman, & Farah, 2005). The reported data are evidence in favor of the above mentioned and often ignored proposition of the TCM. The results argue against the acceptability of confounding demanded accuracy and format; and against accounts that treat symbolic and nonsymbolic approximation as two distinct abilities (see for example Ebersbach et al., 2013; Crollen et al., 2011).

A second focus in the longitudinal studies were the predictive relationships of our approximate measures to math performance at the end of term. As hypothesized, symbolic approximate arithmetic predicted a significant proportion of variance in the DEMAT 1+ at the end of first grade, also after controlling for the effects of age and working memory. Contrary to Study 1, in Study 2 also the nonsymbolic approximate arithmetic performance predicted math performance. However, after controlling for age and working memory of the participants, and in line with the results of Study 1, the symbolic approximate measure explained a notably larger proportion of variance than the nonsymbolic one. These results confirm and extend cumulative evidence in current research that ANS-based abilities can predict formal math performance almost a year later. While there is a large body of research stating a predictive relationship between nonsymbolic comparison measures and exact math performance for different age ranges (preschoolers: Libertus et al., 2013; vanMarle et al., 2014; elementary school: Pinheiro-Chagas et al., 2014; adults: Lyons & Beilock, 2011), our findings arose from the administration of a different ANS measure, the approximate arithmetic task, both in a symbolic and a nonsymbolic format. Our results are in accordance with those of Pinheiro-Chagas et al. (2014). They administered different ANS measures and had already found approximate nonsymbolic arithmetic to predict math performance in grades 1 to 6 and to actually mediate the predictive capacity of the comparison measure. Park and Brannon (2013) also found approximate nonsymbolic addition to be a better training for subsequent exact arithmetic in adults than an intervention based on ordinal knowledge.

In our Study 2, numerical ordering ability also contributed to the explained variance of formal math achievement at the end of term, with its predictive capacity ranging between those of symbolic and nonsymbolic approximation. To test the frequently proposed mediation of the connection between ANS and formal math through numerical ordering, in a second regression model we explored to which extent the approximation measures were predictive beyond numerical ordering ability. Both measures kept significant predictive capacity, but that of the symbolic measure was markedly reduced when being put into the regression model after numerical ordering ability. Mediation analyses confirmed that the prediction of formal math performance through nonsymbolic approximation was not mediated by numerical ordering. On the contrary, the according relationship between symbolic approximation and math achievement was significantly mediated by numerical ordering. However, this mediation was only partial, in line with the results of Lyons et al. (2014) who found that from the start of formal education, ordinal knowledge only starts to gain importance for children´s numerical skills. This supports our idea that not ordinal knowledge per se is what mediates the relationship between ANS and math achievement, but rather the ability to transfer the understanding of the employed regularities of an ANS task onto a formal, symbolic context. Earlier studies that had reported the mediation (Lyons & Beilock, 2011) had only used comparison tasks as their ANS measure. As mentioned above, depending on the respective instructions, these tasks can be viewed as mostly requiring an ordinal judgment. This makes it less surprising that the same ability brought into the representational format of the math achievement measure (Arabic symbols) works as a mediator. We on the other hand found no mediation through numerical ordering between nonsymbolic approximate arithmetic and formal math performance, indicating that it depends on the kind of ANS measure which mediating relationships are detected. This contradicts Lyons and Beilock´s (2011) assumption that ordinal skills are the general interface between ANS and later exact mathematics. They had suggested that the ability to represent a quantity relative to other quantities might be the key to transform the coarse representations of the ANS to increasingly exact ones. However, as we observed only a partial mediation between children´s approximate symbolic arithmetic skills and later formal math performance through numerical ordering, and none at all when nonsymbolic approximation was the predictor, we suggest that other factors connecting ANS tasks and formal math assessment have to be considered, too. The fact that a mediation through numerical ordering was found between symbolic approximation and the DEMAT 1+ scores might be explained by the common format of the approximation and the NO task (Arabic numerals), which naturally is also important for formal math performance. This notion receives further support by the finding that in the hierarchical multiple regression model in which numerical ordering was put into the equation before the approximation measures, both approximation tasks predicted later math performance to an equal degree. Still, both approximation measures kept significant predictive capacity beyond numerical ordering. This might hint to the importance of the ability to not only represent, but actually manipulate quantities on the coarse representational level of the ANS, which might be transferred onto exact contexts.

The results obtained in the longitudinal studies have brought forth the potential of symbolic approximate arithmetic: Amongst all 2x2 realizations of demanded accuracy and presentation format, this was the only task that consistently was not influenced by children´s SES, and furthermore the one that in both studies reliably predicted participant´s math performance at the end of term. These findings, taken together with existing research demonstrating the possibility to foster exact arithmetic by means of nonsymbolic approximation (e.g., Hyde et al., 2014; Park & Brannon, 2013), qualify this task as a promising intervention or training means which might induce understanding and exploitation of arithmetic principles (see for example Sherman & Bisanz, 2009). While the existing research has mainly provided evidence that simple addition and subtraction performance can be fostered by approximation, to our knowledge there are no attempts to actually boost the use and understanding of more abstract principles with approximation. Besides the ecological value of knowing and applying principle-based shortcuts, extending the training research to arithmetic principles allows us to specify the kind of knowledge that might benefit from the approximation induction: Can we only foster procedural knowledge (that is, strategy-use) or also (metacognitive) conceptual knowledge because of an intuitive understanding of regularities on the fundamental ANS level? The approximate symbolic task was promising as an induction because due to our results, all children should profit to an equal degree from this induction task. On the contrary, in a nonsymbolic format, children with higher SES might be more practiced already at the beginning of formal education. We thus supposed that in mixed-SES classes, symbolic approximate arithmetic should be more effective in inducing an arithmetic principle than nonsymbolic approximate arithmetic.

This hypothesis, tested in the Experiments 3 and 4 a – c with the additive law of commutativity as our test case, was confirmed. We found that in second graders, a symbolic approximate arithmetic induction containing commutativity problems could foster the subsequent procedural exploitation of commutativity-based shortcuts in (exact and symbolic) formal problems to an equal degree as an explicit instruction about the principle. An analogue approximate *nonsymbolic* induction, however, had no comparable beneficial influence. In series 4 a – c, we replicated the effect of the approximate symbolic induction in first graders who had just started school and had never been taught the principle, in first graders at the end of first grade who had learned about commutativity shortly before; and finally in third graders who should have substantial practice in applying principles like commutativity. All these age groups profited procedurally from the approximate symbolic induction. In each but the youngest sample, we also administered a measure for conceptual understanding of commutativity. In this task, all conditions within each of the four experiments showed comparable performance. Thus, none of the inductions, but – even more notably – neither the direct and explicit instruction about the additive law of commutativity could benefit conceptual understanding (at least not in the short term as was tested in our experiment). SDT measures indicated that participants already had some conceptual understanding of commutativity, but that this understanding could just not be strengthened any further by our inductions and instruction. This is in line with the results obtained by Sherman and Bisanz (2009) who did not manage to induce conceptual transfer with their (nonsymbolic) induction either. Also the integration of both kinds of knowledge could not be fostered by our induction. It also has to be noted that the beneficial effect on procedural knowledge was strongest in the youngest sample (Experiment 4 a) and became weaker with increasing age of participants. This might on the one hand be based on the logarithmic-to-linear shift proposed for the first years of elementary school (Booth & Siegler, 2006). With longer time in school and increasing experience in formal mathematical processing, referring to the ANS might lose its efficacy and the beneficial effect of the inductions would decrease. Note, however, that Park and Brannon (2013) found positive effects of nonsymbolic approximate arithmetic even in adults. Our results also point to another explanation. On a descriptive level our findings show a training effect (that is, more correct solutions in the noncommutative subset) in the computation-first condition only in the youngest sample. This might hint to an increasing spontaneous exploitation of commutativity in older children also without our induction, even though this exploitation does not produce a statistically secured commutativity effect.

As already mentioned, so far, mostly nonsymbolic approximation inductions have been found to benefit exact arithmetic (Hyde et al., 2014; Park & Brannon, 2013). We, on the other hand, found symbolic approximation to be the better means to induce an arithmetic principle. In line with our results in the Studies 1 and 2, we suppose that symbolic approximate arithmetic is a kind of task from which *all* children within a SES-mixed class can profit to an equal degree, hence producing the most pronounced benefit in subsequent problems.

All in all, our results support the propositions of the TCM. For the tested age group at the beginning of formal education, we found indication of stronger connections between the mastery of approximation (respectively exact arithmetic) in a symbolic and in a non-symbolic format, than between performances in tasks with the common factor of demanding the understanding of Arabic symbols (respectively not demanding it). The process that Lyons et al. (2012) termed 'symbolic estrangement' cannot be seen in this group. In their adult participants, they had found that comparison judgments across different formats took more time than when two nonsymbolic arrays or two Arabic numerals had to be compared. They suggested a development of symbolic numerical skills distinct from the fundamental ANS. However, while we cannot rule out with our results that with increasing age and experience in formal mathematics, number symbols might not or might only more ponderously activate the according analogue magnitude, our results show that this is not yet the case at the beginning of school. Mejias and Schiltz (2013) found all 2x2 tasks (in the domain of number representation) being intercorrelated right before starting school, while one year before it had only been tasks that presented numerical content nonsymbolically. We, on the other hand, found that when starting school, abilities of low-SES first graders rather were integrated by symbolic number knowledge. However, their high-SES peers – who we assume are ahead in their development (Jordan et al., 2008, Jordan & Levine, 2009) – again show close connections between performances in same operations regardless of their format. Hence, although the suggested estrangement might happen at a later time, we assume that both kinds of representations develop closely interrelated. Taking our results into account together with the findings of Mejias and Schiltz suggests an iterative development in which the newly acquired representational format of Arabic numerals is determining and integrative for children´s performances especially at the beginning of its usage. After its mastery is established, however, it seems that abilities become reintegrated and that at after that the kind of process is more determining for performance than the mere format.

A remarkable result of our research is the prominent role of exact nonsymbolic arithmetic processing. Not only was it the most difficult task for our participants, but it also showed correlations with all other performances, especially in low-SES participants. We assume that this task tackles a superordinate ability that drives early numerical development and the transition from nonsymbolic to symbolic expertise. This is in line with another prominent account of the development of numerical understanding, coming from Resnick (1992, see also Geary, 2006). She suggested that children understand arithmetic principles on increasingly abstract levels. The first level is termed mathematics of *protoquantities*, in which children reason about concrete objects mostly qualitatively and without any reference to specific numerosities. On an approximate level, however, quantitative reasoning is possible, for example by predicting that the action of adding items to a collection increases its size. On the second level, the level of *quantities*, children can already process number symbols in an exact manner as long as these have concrete object referents. The third level (mathematics of *numbers*) implicates that children do not need object referents anymore to reason and draw conclusions about numbers and less basic arithmetic principles. On the last level of Resnick´s model, the level of *operators*, children finally construct general arithmetic principles by treating not only numbers but also operations themselves as conceptual entities that can be reasoned about abstractly. The transition from the second (quantities) to the third stage (numbers) has been found to happen somewhere around grades 1 to 3 (Baroody et al., 1983) and thus matches our finding that the ability to exactly manipulate quantities presented with concrete object referents (and not only as Arabic numbers) is especially important at this developmental stage. It is one of several steps to increasing abstraction and more comprehensive concepts of arithmetic operations and principles.

That conceptual knowledge at this point is still rather limited is further demonstrated by our attempts to foster conceptual understanding of commutativity. With our inductions that referred to precursory knowledge of that principle by demanding only approximate solutions, we did not succeed to benefit subsequent metacognitive knowledge (conceptual understanding) in a more formal context. In accordance with the results of Sherman and Bisanz (2009) who did not manage to foster conceptual knowledge of the arithmetic principle of equivalence with their exact and concrete induction, we suppose that more experience with this principle might be needed to show also in conceptual gains (see also McNeil & Alibali, 2000; or Fyfe et al., 2012). However, our approximate symbolic arithmetic task was as suitable an induction for the procedural exploitation of the law of commutativity as an explicit instruction, and even exceeded an analogue nonsymbolic task in its efficacy. This fits with the TCM idea of a semantic transcoding route between the symbolic and analogue magnitude code (Crollen et al., 2011; Dehaene, 1992; Dehaene et al., 2003). Activating the analogue representation (ANS) when approximating symbolic problems can facilitate retranslating of the processes executed in the ANS and the understanding of the associated principles back to the symbolic representation and the formal context afterwards. According to other research (Hyde et al., 2014; Park & Brannon, 2013) and in line with our finding that both symbolic and nonsymbolic approximate arithmetic predict later math performance (see also Gilmore et al., 2010), nonsymbolic approximation should also be suited as an intervention to boost exact arithmetic. Although our results do not exclude this possibility (in Experiment 3 – the only one in which we administered the nonsymbolic approximate induction – we only compared its effect to that of the symbolic induction and an explicit instruction, not to a baseline of procedural knowledge like in Experiments 4 a – c), undoubtedly the symbolic induction produced a greater benefit. Concluding from our results obtained in the longitudinal studies, we suppose that this goes back to the task´s robustness against socioeconomic influences. Thus, in a mixed sample, symbolic approximation seems to reach all learners to an equal degree, leading to a marked procedural benefit in a subsequent formal mathematical context.

To sum it up, our two series of longitudinal studies (Study 1 and 2) and intervention experiments (Experiments 3 and 4 a – c) obtained the following insights. (1) First grade is a sensible period to study numerical cognition. While the influence of socioeconomic factors have already spread from recently acquired formal skills to informal abilities with an important integrative function of the capacity to process quantities exactly, within groups of similar SES the abilities are interrelated differentially. Low-SES children show a stronger connection between performances that demand symbolic number knowledge than their high-SES peers, while these in turn show a better integration of both formats (symbolic and nonsymbolic) when these represent the same kind of operation (approximate vs. exact arithmetic). (2) While Gilmore et al. (2010) found nonsymbolic approximate arithmetic skills being unaffected by SES in preschoolers, our results show that in first grade this does not hold anymore. In our second, more balanced sample, approximate nonsymbolic arithmetic performance was worse in low- than in high-SES children. The opposing finding that symbolic approximate arithmetic does *not* vary with SES can be explained by this being a combination of format and demanded accuracy that children had encountered least often and – due to the motivation to become 'proficient mathematicians' – that they had not practiced it in their everyday life. From that it would follow that this kind of processing on the representational level of the ANS might be the least refined by the newly acquired exact skills. This is in line with on the one hand the demonstrated robustness against socioeconomic influences, and on the other hand its independence from working memory. (3) Approximate arithmetic at the beginning of first grade is suited to predict formal math performance at the end of term. The relationship is only partly mediated by numerical ordering ability for approximate symbolic arithmetic and not at all for approximate nonsymbolic arithmetic. This argues against the notion of ordinal knowledge being the genuine interface between the ANS and exact math skills, as proposed by Lyons and Beilock (2011). In line with current research (DeSmedt et al., 2013), approximate nonsymbolic arithmetic was a more unstable predictor than the symbolic measure. This was mirrored in a nonsignificant regression in Study 1 and the smaller proportion of explained variance in Study 2. (4) Approximate symbolic arithmetic is well suited to induce an arithmetic principle that in difficulty ranges beyond the most basic regularities. An according induction that provided the opportunity to discover the additive law of commutativity resulted in stronger procedural exploitation than achieved by an analogue nonsymbolic induction task. In second graders, the fostering effect of symbolic approximation was of equal size than the one obtained after a direct and explicit instruction of the principle. The beneficial effect of the symbolic approximate induction task was replicated in third graders and two groups of first graders, one of which just had started school and never heard about the principle before. In line with current literature, conceptual knowledge could not be enhanced in any of the age groups.

15.1 Limitations

The studies and experiments reported in this thesis are subject to some limitations that should be kept in mind when interpreting the results and their implications. As already mentioned, due to alterations of tasks during the course of testing and the resulting exclusion of two participating classes, analyses of our first study were conducted with a rather unbalanced sample regarding children´s SES. We tried to incorporate this fact in our analyses of the second study´s data and were able to explain and secure changes in the results by this differing sample compositions. Furthermore, all our studies and experiments were conducted in a group setting. Naturally, this produces more noise in the data than laboratory and computerized research. Nevertheless, it can also be seen as a specific value of our work that we obtained and replicated many effects despite this setting and its possible interfering factors. One specific issue concerns our question for the reason of socioeconomic impact on numerical skills. From the correlations of both exact main tasks with children´s SES, we concluded that not only symbolic number knowledge, but also language skills probably play a role in the mechanism by which socioeconomic factors gain the according influence. However, note that in our Study 2, also approximate nonsymbolic processing was influenced by SES. We assume that already at this early stage of numerical development, at least high-SES children have trained the more common problem formats, to which also nonsymbolic approximation can be counted. This indicates that the two competing hypotheses we formulated regarding SES influence can probably be answered more clearly in younger samples in the future. It thus needs more systematic longitudinal studies that compare the 2x2 performances in samples of different SES and different age ranges. To secure our interpretations, in future studies not only the numerical skills tested here should be incorporated, but also an explicit measure of language skills.

In the experimental series, we have to keep in mind that our approximate inductions there were designed to serve as an intervention. To reliably secure that the influence from approximate arithmetic to formal math is unidirectional, future research should replicate these experiments with an approximation measure that is equally sensitive as the computation task. Another question coming up with these experiments might concern our assessment of conceptual and procedural knowledge. We tried to unobtrusively measure both kinds of knowledge separately, but we cannot rule out that our computation task also taps conceptual knowledge of commutativity. We suppose, however, that this should only be the case to a minor degree. Otherwise, we would expect performances in the computation and the judgment task to correlate, which was not the case.

15.2 Implications

All in all, our results are further support for a strong connection between early, approximate numerical abilities and the more formal exact competencies as they are required in school. At the same time in which the studies and experiments for this thesis were conducted, research interest in the question how symbolic and nonsymbolic approximation might differ in this regard, and a growing awareness of the necessity to compare these kinds of approximation have developed. For example, a recently published meta-analysis (Schneider et al., 2016) confirmed the stronger relationship between symbolic comparison and formal math assessment (compared to the nonsymbolic comparison measure). However, our results extend these findings in demonstrating that even when the ANS measure testing this relationship is not comparison but approximate arithmetic, the relationship to formal math performance is stronger for the symbolic measure than for the nonsymbolic one. While there is diverging evidence concerning the role of ordinal abilities for the relationship between comparison tasks and math performance (Vogel et al., 2015, found no mediation of NO between symbolic comparison and exact math performance; Lyons & Beilock, 2011 found a full mediation for the according relationship using a nonsymbolic comparison measure); our results contribute converging evidence that approximate *arithmetic* tasks have their own predictive capacity beyond ordinal knowledge and should get more attention in future research. At the developmental stage of beginning formal education, they even might be the more meaningful measure (Pinheiro-Chagas et al., 2014, Gilmore et al., 2014) and additionally provide the beneficial possibility to foster the understanding of arithmetic principles.

Our results show that activating children´s early knowledge of commutativity by a symbolic approximation task positively influenced the strategy use in formal arithmetic. This was the case either before or *after* children had received an according instruction in school. Therefore, we assume that children do not automatically activate their precursory mathematical knowledge in order to support their understanding of formal mathematical principles taught in primary school. However, our findings indicate that teachers can help children by explicitly referring to their informal knowledge and ANS-based competencies. Although there is certainly a need to look for more methods and ways to improve the conceptual understanding of commutativity, it seems promising to include nonsymbolic, as well as symbolic approximate tasks as economical and practicable means in mathematical instruction. Our findings confirm the TCM proposition that also Arabic numerals (symbolic number representation) can be used in approximation tasks (Dehaene, 1992; Dehaene et al., 2003, Gilmore et al., 2007). According to the results we obtained in Studies 1 and 2, using an Arabic representation in approximate arithmetic tasks for first graders holds the advantage that this format seems to be least trained, and proficiency in solving such problems does not differ with children´s socioeconomic status. Thus, in classes where there is a distribution of very different SES levels, symbolic approximate arithmetic is the setting that all children can refer to in equal ease. Furthermore, to induce mathematical principles nonsymbolically might seem like a setback to older children. Yet, in our study, symbolic approximation could successfully foster the exploitation of commutativity even in third graders. Thus, our findings suggest that symbolic approximation tasks can help children to spot and apply more efficient strategies in elementary school. We suggest that inducing a principle in a symbolic approximate representation is suited to link informal understanding to the understanding and exploitation of the numerical version of the principle in a formal context. This might be an important premise for paving the way to an understanding of the principle in its truly abstract conception on the long run, enabling the learners not only to spontaneously use it, but also to integrate it in a broader context of more advanced mathematics. This is especially promising because in the future it might open the possibility to teach mathematical principles in elementary school independently of language barriers (as some principles obviously are not better understood after explicit – verbal – instruction) and disadvantages because of varying socioeconomic backgrounds in class.

16 Literature

Anderson, S. W., Damasio, A. R., & Damasio, H. (1990). Troubled letters but not numbers. *Brain and Cognition*, *113*, 749–766.

Anderson, U. S., & Cordes, S. (2013). 1 < 2 and 2 < 3: non-linguistic appreciations of numerical order. *Frontiers in Psychology*, *4*, 695.

Attout, L., Noël, M.-P., & Majerus, S. (2014). The relationship between working memory for serial order and numerical development: A longitudinal study. *Developmental Psychology*, *50*(6), 1667–1679.

Baroody, A. (1987). The Development of Counting Strategies for Single-Digit Addition. *Journal for Research in Mathematics Education*, *18*(2), 141–157.

Baroody, A. (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). Lawrence Erlbaum Associates.

Baroody, A., & Gannon, K. E. (1984). The Development of the Commutativity Principle and Economical Addition Strategies. *Cognition and Instruction*, *1*(3), 321‑339.

Baroody, A. J., & Ginsburg, H. (1983). The Effects of Instruction on Children's Understanding of the "Equals" Sign. *The Elementary School Journal*, *84*(2), 198–212.

Baroody, A., Ginsburg, H., & Waxman, B. (1983). Children´s Use of Mathematical Structure. *Journal for Research in Mathematics Education*, *14*(3), 156–168.

Baroody, A., Wilkins, J. L., & Tiilikainen, S. H. (2003). The Development of Children's Understanding of Additive Commutativity: From Protoquantitative Concept to General Concept? In A. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills. Constructing Adaptive Expertise* (1st ed., pp. 127–160). Lawrence Erlbaum Associates.

Barth, H., Baron, A., Spelke, E., & Carey, S. (2009). Children’s multiplicative transformations of discrete and continuous quantities. *Journal of Experimental Child Psychology*, *103*(4), 441–454.

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, (39), 14117–14121.

Berends, I. E., & van Lieshout, E. C. (2009). The effect of illustrations in arithmetic problem-solving: Effects of increased cognitive load. *Learning And Instruction*, *19*(4), 345‑353.

Bermejo, V., & Rodriguez P. (1993). Children's understanding of the commutative law of addition. *Learning And Instruction*, *3*(1), 55–72.

Berteletti, I., Lucangeli, D., & Zorzi, M. (2012). Representation of numerical and non-numerical order in children. *Cognition*, *124*(3), 304–313.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, *29*(4), 711–721.

Blevins-Knabe, B., & Musun-Miller, L. (1996). Number Use at Home by Children and Their Parents and Its Relationship to Early Mathematical Performance. *Early Development & Parenting,*, *5*(1), 35–45.

Bönig, Dagmar (2001): „Das Ungefähre in der richtigen Antwort". Zur Bedeutung des Schätzens beim Umgang mit Größen. *Die Grundschulzeitschrift*, *15*(141), 43–45.

Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology*, *114*(3), 375–388.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *42*(1), 189‑201.

Bos, W., Wendt, H., Köller, O., & Selter, C. (2012). Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich: TIMSS 2011. Münster: Waxmann.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240.

Brannon, E. M., & van de Walle, G. A. (2001). The Development of Ordinal Numerical Competence in Young Children. *Cognitive Psychology*, *43*(1), 53–81.

Brownell, W. A. (1935). Psychological considerations in learning and teaching arithmetic. In W.D. Reeve (Ed.), *The Teaching of Arithmetic: Tenth Yearbook of the National Council of Teachers of Mathematics* (pp. 1-31). New York: Teachers College, Columbia University.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision*, *10*(6), 1‑10.

Calvo, A., & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, *130*(3), 278–288.

Canobi, K. H. (2005). Children's profiles of addition and subtraction understanding. *Journal of Experimental Child Psychology, 92,* 220–246.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (1998). The role of conceptual understanding in children's addition problem solving. *Developmental Psychology*, *34*(5), 882–891.

Canobi, K. H., Reeve, R. A., & Pattison, P. E. (2002). Young children's understanding of addition concepts. *Educational Psychology*, *22*(5), 513–532.

Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences*, *13*(2), 83–91.

Charras, P., Brod, G., & Lupiáñez, J. (2012). Is 26 + 26 smaller than 24 + 28? Estimating the approximate magnitude of repeated versus different numbers. *Attention, Perception, & Psychophysics*, *74*(1), 163‑173.

Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica*, *148*, 163–172.

Chesney, D. L., & Haladjian, H. H. (2011). Evidence for a shared mechanism used in multiple-object tracking and subitizing. *Attention, Perception, & Psychophysics*, *73*(8), 2457–2480.

Cohen, L., Verstichel, P., & Dehaene, S. (1997). Neologistic Jargon Sparing Numbers: A Category-specific Phonological Impairment. *Cognitive Neuropsychology*, *14*(7), 1029–1061.

Cowan, R. (2003). Does it all add up? Changes in children's knowledge of addition combinations, strategies, and principles. In A. Baroody & A. Dowker (Eds.), *The Development of Arithmetic Concepts and Skills. Constructing Adaptive Expertise* (1st ed., pp. 35–74). Lawrence Erlbaum Associates.

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*(4), 407–420.

Crollen, V., Castronovo, J., & Seron, X. (2011). Under- and over-estimation: a bi-directional mapping process between symbolic and non-symbolic representations of number? *Experimental Psychology (formerly Zeitschrift für Experimentelle Psychologie)*, *58*(1), 39–49.

de Smedt, B., Noël, M.-P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education*, *2*(2), 48–55.

de Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 469– 479.

Dean Jr., D., & Kuhn, D. (2007). Direct instruction vs. discovery: The long view. *Science Education*, *91*(3), 384–397.

DeCaro, M. S., & Rittle-Johnson, B. (2012). Exploring mathematics problems prepares children to learn from instruction. *Journal of Experimental Child Psychology*, *113*(4), 552–568.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*, 1–42.

Dehaene, S. (1999). Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science*, *284*(5416), 970–974.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*(3-6), 487–506.

Dowker, A. (2009). Use of derived fact strategies by children with mathematical difficulties. *Cognitive Development*, *24*(4), 401–410.

Dowker, A. (2012). Derived fact strategies. In *Individual Differences in Arithmetic. Implications for Psychology, Neuroscience and Education*, pp. 123–147. New York: Psychology Press.

Dowker, A. (2012). Individual Differences in Arithmetic: Implications for Psychology, Neuroscience and Education. New York: Psychology Press.

Dowker, A. (2014). Young children's use of derived fact strategies for addition and subtraction. *Frontiers in Human Neuroscience, 7.*

Ebersbach, M., & Erz, P. (2014). Symbolic versus non-symbolic magnitude estimations among children and adults. *Journal of Experimental Child Psychology*, *128*, 52–68.

Ebersbach, M., Luwel, K., & Verschaffel, L. (2013). Comparing apples and pears in studies on magnitude estimations. *Frontiers in Psychology*, *4.*

Erikson, R., Goldthorpe, J. H., & Portocarrerp, L. (1979). Intergenerational class mobility in three Western European societies: England, France and Sweden. *The British Journal of Sociology*, *30*(4), 415–441.

Farrington-Flint, L., Canobi, K. H., Wood, C., & Faulkner, D. (2007). The role of relational reasoning in children's addition concepts. *British Journal of Developmental Psychology*, *25*(2), 227–246.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, Instruments, & Computers*, *39*(2), 175–191.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. Trends in Cognitive Sciences, 8(7), 307–314.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), The nature of intelligence (pp. 231 - 235). Hillsdale, NJ: Lawrence Erlbaum Associates.

Fyfe, E. R., Rittle-Johnson, B., & DeCaro, M. S. (2012). The effects of feedback during exploratory mathematics problem solving: Prior knowledge matters. *Journal of Educational Psychology*, *104*(4), 1094–1108.

Gallistel, C., & Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*, *4*(2), 59–65.

Gallistel, R. (1990). *The organization of learning*. Cambridge, MA: Bradford Books/MIT Press.

Gallistel, R., & Gelman, R. (1992). Preverbal and verbal counting and computation. Cognition, 44, 43 – 74.

Ganor-Stern, D. (2015). When you don't have to be exact: Investigating computational estimation skills with a comparison task. *Acta Psychologica*, *154*, 1–9.

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.

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.

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

Gilmore, C. K., & Spelke, E. S. (2008). Children's understanding of the relationship between addition and subtraction. *Cognition*, *107*(3), 932–945.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*(7144), 589–591.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*(3), 394–406.

Gilmore, C., Attridge, N., de Smedt, B., & Inglis, M. (2014). Measuring the approximate number system in children: Exploring the relationships among different tasks. *Learning and Individual Differences*, *29*, 50–58.

Godau, C., Haider, H., Hansen, S., Schubert, T., Frensch, P. A. & Gaschler, R. (2014). Spontaneously spotting and applying shortcuts in arithmetic - a primary school perspective on expertise. *Frontiers in Psychology, 5*, 1-11.

Hackman, D. A., Betancourt, L. M., Gallop, R., Romer, D., Brodsky, N. L., Hurt, H., & Farah, M. J. (2014). Mapping the Trajectory of Socioeconomic Disparity in Working Memory: Parental and Neighborhood Factors. *Child Development*, *85*(4), 1433–1445.

Haider, H., Eichler, A., Hansen, S., Vaterrodt, B., Gaschler, R. & Frensch, P. A. (2014). How we use what we learn in Math: An integrative account of the development of commutativity. *Frontline Learning Research, 1,* 1-21.

Halberda, J., & Feigenson, L. (2008). Set representations required for the acquisition of the “natural number” concept. *Behavioral and Brain Sciences*, *31*(06), 655.

Hannula, M. M. & Lehtinen, E. (2005). *Spontaneous focusing on numerosity and mathematical skills of young children.*

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*(4), 394–406.

Hansen, S. M., Haider, H., Eichler, A., Godau, C., Frensch, P. A., & Gaschler, R. (2015). Fostering formal commutativity knowledge with approximate arithmetic. *PloS ONE*, DOI:10.1371/journal.pone. 014255.

Hatano, G. (1988). Social and motivational bases for mathematical understanding. *New Directions For Child Development :55-70, 41*, 55–70.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Methodology in the social sciences. New York, NY: Guilford Press.

Hiebert, J., & Lefevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis, 1-27, xiii, 309.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction, 14*((3), 251-283.

Hoffmann, D., Hornung, C., Martin, R., & Schiltz, C. (2013). Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds. *Journal of Experimental Child Psychology*, *116*(4), 775–791.

Huang, Y. T., Spelke, E., & Snedeker, J. (2010). When Is Four Far More Than Three?: Children's Generalization of Newly Acquired Number Words. *Psychological Science*, *21*(4), 600–606.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology*, *61*(4), 343–365.

Hyde, D. C., Khanum, S., & Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition, 131*(1), 92–107.

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106*, 1221–1247.

Jordan, N. C., Huttenlocher, J., & Cohen Levine, S. (1992). Differential Calculation Abilities in Young Children From Middle- and Low-Income Families. *Developmental Psychology*, *28*(4), 644–653.

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice, 22*(1), 36–46.

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

Kalaman, D. A., & LeFevre J. (2007). Working memory demands of exact and approximate addition. *European Journal Of Cognitive Psychology,19*(2), 187–212.

Klein, E., Moeller, K., & Willmes, K. (2013). A neural disconnection hypothesis on impaired numerical processing. *Frontiers in Human Neuroscience*, *7*.

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*(2), 105–115.

Knops, A., Viarouge, A., & Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: Evidence from the operational momentum effect. *Attention, Perception, & Psychophysics*, *71*(4), 803–821.

Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning And Instruction*, *25*, 95–103.

Koshmider, J. W., & Ashcraft, M. H. (1991). The development of children's mental multiplication skills. *Journal of Experimental Child Psychology*, *51*(1), 53–89.

Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition*, *105*(2), 395–438.

LeFevre, J.-A., Greenham, S. L., & Waheed, N. (1993). The Development of Procedural and Conceptual Knowledge in Computational Estimation. *Cognition and Instruction*, *11*(2), 95–132.

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*(4), 285–303.

Lemaire, P., Lecacheur M., & Farioli F. (2000). Children's strategy use in computational estimation. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *54*(2), 141–148.

Lenhard, W. & Lenhard, A. (2006). *ELFE-Training: Förderung des Leseverständnisses für Schüler der 1. bis 6. Klasse*. Göttingen: Hogrefe.

Levine, S., & Jordan, N. C. H. J. (1992). Development of Calculation Abilities in Young Children. *Journal of Experimental Child Psychology, 53*, 72 - 103.

Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences*, *25*, 126–133.

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subject designs. *Psychonomic Bulletin & Review*, *1*(4), 476–490.

Lonnemann, J., Linkersdorfer, J., Hasselhorn, M., & Lindberg, S. (2011). Symbolic and non-symbolic distance effects in children and their connection with arithmetic skills. *Journal of Neurolinguistics*, *24*(5), 583‑591.

Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition*, *121*(2), 256–261.

Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General*, *141*(4), 635–641.

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, *17*(5), 714–726.

Mandler, G., & Shebo, B. J. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General*, *111*(1), 1–22.

Martinez Perez, T., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, *111*(4), 708–723.

Matthews, P., & Rittle-Johnson, B. (2009). In pursuit of knowledge: Comparing self-explanations, concepts, and procedures as pedagogical tools. *Journal of Experimental Child Psychology*, *104*(1), 1–21.

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the Approximate Number System predicts later school mathematics performance. *PLoS ONE*, *6*(9), ArtID: e23749.

McCloskey, M., Aliminosa, D., & Sokol, S. M. (1992). Facts, rules and procedures in normal calculation: Evidence from multiple single-patient studies of impaired arithmetic fact retrieval. *Brain and Cognition*, *17*(2), 154–203.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology*, *92*(4), 734–744.

McNeil, N. M., Fuhs, M. W., Keultjes, M. C., & Gibson, M. H. (2011). Influences of problem format and SES on preschoolers' understanding of approximate addition. *Cognitive Development*, *26*(1), 57–71.

Meck, W. H., & Church, R. M. (1980). A Mode Control Model of Counting and Timing Processes. *Journal Of Experimental Psychology: Animal Behavior Processes*, *9*(3), 320–334.

Mejias, S., & Schiltz, C. (2013). Estimation abilities of large numerosities in Kindergartners. *Frontiers in Psychology*, *4*.

Moeller, K., Klein, E., Nuerk, H.-C., & Cohen Kadosh, R. (2012). A Unitary or Multiple Representations of Numerical Magnitude? – the Case of Structure in Symbolic and Non-Symbolic Quantities. *Frontiers in Psychology*, *3*.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of Numerical Inequality. *Nature*, *215*(5109), 1519–1520.

Müller, W., Wirth, H., Bauer, G., Pollak, R., & Weiss, F. (2007). Entwicklung einer Europäischen Sozioökonomischen Klassifikation. *Wirtschaft und Statistik*, *5*, 527–530. Retrieved from https://www.destatis.de/DE/Publikationen/WirtschaftStatistik/Monatsausgaben/WistaMai07.pdf?__blob=publicationFile.

Mussolin, C., Mejias, S., & Noël, M.-P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition*, *115*(1), 10–25.

Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, *8*(1), 74–87.

Noël, M.-P., & Rousselle, L. (2011). Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. *Frontiers In Human Neuroscience*, *5*.

Notebaert, K., Nelis, S., & Reynvoet B. (2011). The magnitude representation of small and large symbolic numbers in the left and right hemisphere: An event-related fMRI study. *Journal of Cognitive Neuroscience*, *23*(3), 622–630.

Obersteiner, A., Reiss, K., & Ufer, S. (2013). How training on exact or approximate mental representations of number can enhance first-grade students' basic number processing and arithmetic skills. *Learning And Instruction*, *23*, 125–135.

Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology*, *49*(6), 1103–1112.

Park, J., & Brannon, E. M. (2013). Training the Approximate Number System Improves Math Proficiency. *Psychological Science*, *24*(10), 2013–2019.

Patro, K., & Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *Journal of Experimental Child Psychology*, *111*(3), 534–542.

Petermann, F., & Petermann, U. (2007). *Hamburg-Wechsler-Intelligenztest für Kinder – Vierte Auflage. Manual [Hamburg-Wechsler-Intelligence Test for children. Fourth edition. Manual]*. Bern: Huber.

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition*, *116*(1), 33–41.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A Magnitude Code Common to Numerosities and Number Symbols in Human Intraparietal Cortex. *Neuron*, *53*(2), 293–305.

Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I.,… (2014). In How Many Ways is the Approximate Number System Associated with Exact Calculation? *PLoS ONE*, *9*(11), 1–13.

Resnick, L. B. (1989). Developing mathematical knowledge. *American Psychologist.44*, 162–169.

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.

Riley, M., Greeno, J., & Heller, J. (1983). Development of children´s problem-solving ability in arithmetic. In H. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–196). New York: Academic Press.

Rilling, M., & McDiarmid, C. (1965). Signal Detection in Fixed-Ratio Schedules. *Science*, *148*(3669), 526–527.

Rips, L. J., Bloomfield, A., & Asmuth, J. (2008). From numerical concepts to concepts of number. *Behavioral and Brain Sciences*, *31*(6), 623–642.

Rittle Johnson, B. (2006). Promoting Transfer: Effects of Self-Explanation and Direct Instruction. *Child Development*, *77*(1), 1–15.

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*(2), 346–362.

Schneider, M. &. S. E. (2010). The Developmental Relations Between Conceptual and Procedural Knowledge: A Multimethod Approach. *Developmental Psychology*, *46*(1), 178–192.

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology*, *47*(6), 1525–1538.

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & Smedt, B. de. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, n/a.

Sherman, J., & Bisanz, J. (2009). Equivalence in symbolic and nonsymbolic contexts: Benefits of solving problems with manipulatives. *Journal of Educational Psychology*, *101*(1), 88–100.

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

Siegler, R. S. (2000). Unconcious Insights. *Current Directions in Psychological Science*, *9*(3), 79–83.

Siegler, R. S., & Booth, J. L. (2005). Development of Numerical Estimation: A Review. In Jamie I. D. Campbell (Ed.), *Handbook of Mathematical Cognition* (pp. 197‑212). New York: Psychology Press.

Siegler, R. S., & Opfer, J. E. (2003). The Development of Numerical Estimation: Evidence for Multiple Representation of Numerical Quantity. *Psychological Science*, *14*(3), 237–243.

Siegler, R. S., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General*, *127*(4), 377–397.

Sophian, C., Harley, H., & Manos Martin, C. S. (1995). Relational and Representational Aspects of Early Number Development. *Cognition and Instruction*, *13*(2), 253–268.

Spelke, E. S. (2011). Natural Number and Natural Geometry. In S. Dehaene & E. Brannon (Eds.), *Space, Time and Number in the Brain. Searching for the Foundations of Mathematical Thought (Attention and Performance)* (pp. 287–317). Amsterdam: Elsevier.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137–149.

Starkey, G. S., & McCandliss, B. D. (2014). The emergence of "groupitizing" in children´s numerical cognition. *Journal of Experimental Child Psychology*, *126*, 120‐137.

Starkey, P., & Klein, A. (2008). Sociocultural influences on young children´s mathematical knowledge. In O. N. Saracho & B. Spodek (Eds.), *Contemporary perspectives in early childhood education. Contemporary perspectives on mathematics in early childhood education* (pp. 253–276). Charlotte, NC: IAP/Information Age Pub.

vanMarle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science*, *17*(4), 492–505.

Verschaffel, L., de Corte, E., & Lasure, S. (1994). Realistic considerations in mathematical modeling of school arithmetic word problems. *Learning And Instruction*, *4*(4), 273–294.

Vogel, S. E., Remark, A., & Ansari, D. (2015). Differential processing of symbolic numerical magnitude and order in first-grade children. *Journal of Experimental Child Psychology*, *129*, 26–39.

Wagner, J. B., & Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition*, *119*(1), 10–22.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488.

Wilson, A. J., Dehaene, S., Dubois, O., & Fayol, M. (2007). Effects of an Adaptive Game Intervention on Accessing Number Sense in Low-Socioeconomic-Status Kindergarten Children. *Mind, Brain, and Education, 3*(4), 224–234.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, (358), 749–750.

Xenidou-Dervou, I., de Smedt, B., van der Schoot, M., & van Lieshout, E. C. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and Individual Differences*, *28*, 119–129.

Xenidou-Dervou, I., van Lieshout, E. C. D. M., & van der Schoot, M. (2014). Working Memory in Nonsymbolic Approximate Arithmetic Processing: A Dual-Task Study With Preschoolers. *Cognitive Science*, *38*(1), 101–127.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-monts-old infants. *Cognition*, *74*, B1-B11.

17 Appendix

Appendix A: Exact problems in Study 1

Trial	Problem	Result
1	5 + 4	9
2	3 + 6	9
3	4 + 3	7
4	5 + 6	11
5	9 – 5	4
6	10 – 7	3
7	8 – 5	3
8	7 – 3	4

Appendix B: Approximation problems in Study 1

	Tim		Lisa		
Trial	Operation	Result/Reference Quantity	Operation	Result/Reference Quantity	Ratio
1	18 + 17	35		21	3:5
2		24	20 + 16	36	2:3
3		65	15 + 24	39	3:5
4	13 + 17	30		45	2:3
5		48	14 + 22	36	3:4
6	34 - 14	20		12	3:5
7		30	31 – 13	18	3:5
8	27 – 15	12		18	2:3
9	40 – 22	18		27	2:3
10		20	42 – 27	15	3:4

Appendix C: Exact problems in Study 2

Trial	Problem	Result
1	5 + 3	8
2	4 + 6	10
3	8 + 4	12
4	6 + 7	13
5	4 + 5	9
6	5 + 9	14
7	7 + 5	12
8	4 + 7	11
9	8 + 6	14
10	7 + 8	15

Appendix D: Approximation problems in Study 2

	Tim		Lisa		
Trial	Operation	Result/Reference Quantity	Operation	Result/Reference Quantity	Ratio
1	18 + 17	35		21	3:5
2		24	20 + 16	36	2:3
3		65	15 + 24	39	3:5
4	13 + 17	30		45	2:3
5	25 + 35	60		45	3:4
6		48	14 + 22	36	3:4
7		28	25 + 17	42	2:3
8	26 + 19	55		33	3:5
9		18	11 + 13	24	3:5
10	12 + 15	27		36	3:4

Appendix E: Trials of the Numerical Ordering Task in Study 2

Page			**Sequence**			
	Position 1	Position 2	Position 3	Position 4	Correct symbol to mark	
1	3	4	7	9	☺	
	2	4	6	3		☹
	9	2	8	4		☹
	5	6	7	9	☺	
	4	5	6	8	☺	
	4	6	8	9	☺	
	7	9	2	6		☹
	1	5	7	8	☺	
	4	6	5	3		☹
	7	6	3	8		☹
2	16	15	13			☹
	17	19	12			☹
	14	16	18		☺	
	13	14	17		☺	
	17	16	13			☹
	15	16	17		☺	
	12	18	14			☹
	11	15	17		☺	
	14	16	13			☹
	14	15	16		☺	
3	12	48	24			☹
	14	21	28		☺	
	13	17	30		☺	
	14	18	21		☺	
	16	13	8			☹
	11	54	77		☺	
	17	22	5			☹
	15	18	21		☺	
	17	15	7			☹
	14	21	11			☹

Note that in children´s booklets, both faces were printed after each sequence, here we depicted only the correct one for each trial

Appendix F: Problems of the approximate induction tasks in Experiments 3 and 4

	Tim		Lisa		
Trial	Operation	Result/Reference Quantity	Operation	Result/Reference Quantity	Ratio
1	49 + 26	75	19 + 31	50	3:2
2	28 + 29	57	39 + 37	76	2:4
3	42 + 45	87	45 + 42	87	com
4	35 + 31	66	31 + 35	66	com
5	38 + 28	66	42 + 57	99	2:3
6	49 + 29	78	29 + 49	78	com
7	25 + 29	54	19 + 17	36	3:2
8	32 + 19	51	31 + 54	85	3:5
9	49 + 47	96	47 + 49	96	com
10	13 + 37	50	37 + 13	50	com
11	57 + 36	93	36 + 57	93	com
12	51 + 53	104	41 + 37	78	4:3
13	31 + 32	63	32 + 31	63	com
14	45 + 60	105	97 + 43	140	3:4

Appendix G: Problems of the Computation Task in Experiments 3 and 4

Comutative Subset

Trial	Problem	Result
1	3+5+4	12
2	**4+9+8**	21
3	**4+8+9**	21
4	6+2+5	13
5	**9+7+2**	18
6	**2+9+7**	18
7	**6+3+2**	11
8	**6+2+3**	11
9	8+9+6	23
10	**7+2+6**	15
11	**6+7+2**	15
12	7+4+8	19
13	5+2+3	10
14	**9+6+2**	17
15	**2+9+6**	17
16	**9+8+5**	22
17	**9+5+8**	22
18	3+2+4	9
19	9+7+8	24
20	**7+9+4**	20
21	**7+4+9**	20
22	9+2+5	16
23	**2+9+3**	14
24	**2+3+9**	14
25	**9+6+5**	20
26	**5+6+9**	20
27	4+2+7	13
28	**3+4+9**	16
29	**9+3+4**	16
30	5+6+7	18

Commutative problems are printed in bold.

Noncommutative Subset

Trial	Problem	Result
1	5+3+4	12
2	8+9+4	21
3	6+7+8	21
4	5+2+6	13
5	2+7+9	18
6	9+4+5	18
7	2+6+3	11
8	2+4+5	11
9	8+6+9	23
10	4+2+9	15
11	7+6+2	15
12	8+7+4	19
13	2+5+3	10
14	9+3+5	17
15	6+2+9	17
16	5+8+9	22
17	9+7+6	22
18	4+3+2	10
19	7+8+9	24
20	7+5+8	20
21	9+7+4	20
22	2+9+5	16
23	3+2+9	14
24	3+6+5	14
25	3+9+8	20
26	5+9+6	20
27	4+7+2	13
28	5+7+4	16
29	4+9+3	16
30	7+6+5	18

Appendix H: Problems of the Judgment Task in Experiments 3 and 4

Trial	Problem	Result
1	2+7+9	18
2	9+5+4	18
3	**2+6+5**	13
4	**6+5+2**	13
5	8+7+5	20
6	**3+5+6**	14
7	**6+5+3**	14
8	2+9+5	16
9	**6+7+9**	22
10	**9+6+7**	22
11	3+4+8	15
12	**8+9+4**	21
13	**4+9+8**	21
14	**3+5+8**	16
15	**5+8+3**	16
16	6+8+9	23
17	5+4+8	17
18	2+6+9	17
19	**4+7+9**	20
20	**7+4+9**	20

Commutative problems are printed in bold. Note that only the test problems (that is, without the 10 training problems of the first page) are depicted. Children did not see the results, but instead a circle that was to mark when they thought a problem could be solved without computation.