CHAPTER 2

Working Memory in Nonsymbolic Approximate Arithmetic Processing: A Dual-Task Study With Pre-schoolers

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*, 101–27. doi:10.1111/cogs.12053

ABSTRACT

Preschool children have been proven to possess nonsymbolic approximate arithmetic skills before learning how to manipulate symbolic math and thus before any formal math instruction. It has been assumed that nonsymbolic approximate math tasks necessitate the allocation of Working Memory resources (WM). WM has been consistently shown to be an important predictor of children's math development and achievement. The aim of our study was to uncover the specific role of WM in nonsymbolic approximate math. For this purpose, we conducted a dual-task study with preschoolers with active phonological, visual, spatial and central executive interference during the completion of a nonsymbolic approximate addition dot-task. With regard to the role of WM, we found a clear performance breakdown in the central executive interference condition. Our findings provide insight into the underlying cognitive processes involved in storing and manipulating nonsymbolic approximate numerosities during early arithmetic.

INTRODUCTION

"Number is the ruler of forms and ideas,
and the cause of gods and daemons"
(Pythagoras as quoted by Iamblichus of Chalcis in Thomas Taylor, 1986)

Nowadays these metaphoric words are illustrated empirically in numeracy being necessary for later life achievement (Duncan et al., 2007; Finnie & Meng, 2001; Reyna & Brainerd, 2011). Given the significance of the domain, research has been flourishing around the question of how children's early ability to learn mathematics develops; specifically around the math-specific and non-specific cognitive precursors of mathematical achievement (Bull, Espy & Wiebe, 2008; De Smedt, Verschaffel, & Ghesquière, 2009; Geary et al., 2009; Gullick, Sprute & Temple, 2010; Holloway & Ansari, 2009; Krajewski & Schneider, 2009; Mundy & Gilmore, 2009) and their interrelations (e.g. Logie & Baddeley, 1987; Noël, 2009; Rasmussen & Bisanz, 2005).

Working memory (WM) emerges as a well-established domain-general cognitive predictor of math performance (Raghubar, Barnes, & Hecht, 2010), playing an important role in early mental arithmetic (DeStefano & LeFevre, 2004). At the same time, studies have highlighted a math-specific precursor of mathematical achievement: the very early ability to conduct basic arithmetic operations with large nonsymbolic numerosities, known as nonsymbolic approximate arithmetic skills (Barth et al., 2006; Gilmore, McCarthy & Spelke, 2010). Some assume that these skills comprise the architectural foundation upon which symbolic math skills are built (e.g. Mundy & Gilmore, 2009), underscoring the importance of a better understanding of their underlying cognitive structure. In this study we sought to test the assumption that nonsymbolic approximate processing necessitates WM. More specifically, we focused on uncovering which specific WM component is involved in this kind of processing in preschool age.

Nonsymbolic approximate arithmetic

A variety of empirical evidence suggests the existence of a cognitive ability that runs across species and improves with development, an inherent precursor of mathematical skills, known as nonsymbolic approximate arithmetic. Animals (Flombaum, Junge, & Hauser, 2005), human infants (McCrink & Wynn, 2004; Xu & Spelke, 2000) and adults who have received no formal instruction or schooling (Pica, Lemer, Izard, & Dehaene, 2004) can conduct basic mathematic operations with approximate numerical magnitudes. Preschool children, who have not yet received any formal math instruction can compare, add and subtract nonsymbolic approximate numerical magnitudes (Barth, et al., 2006; Gilmore, et al., 2010; Gilmore & Spelke, 2008) even when the elements presented are different in format and modality (Barth, Beckmann, & Spelke, 2008; Barth, La Mont, Lipton & Spelke, 2005). Often these skills are referred to as reflecting an innate "Approximate Number System" (ANS). In the literature, evidence for the ANS comes from several different types of tasks, which, however, have been recently found to be uncorrelated (Gilmore, Attridge & Inglis, 2011). It should be noted that in this study we place focus only on the nonsymbolic approximate skills of the ANS, skills that involve the addition and comparison of large numerosities.

It is theorized that exact symbolic verbal mathematic skills – i.e., as taught in school – develop on top of and are fostered by approximate nonsymbolic arithmetic skills (Mundy & Gilmore, 2009; see Noël & Rouselle, 2011 for an alternative view). For example, Gilmore and colleagues (2010) showed that preschoolers' nonsymbolic approximate addition skills were associated with their formal symbolic mathematical performance, even when controlling for intelligence and literacy skills. Thus, it is imperative to understand and uncover the cognitive processes underlying children's nonsymbolic approximate arithmetic skills.

The most common task for assessing nonsymbolic approximate arithmetic skills is a computer-animated task (Barth, et al., 2005; Barth et al., 2006; Gilmore et al., 2010), which we will refer to from here on as the dot-task. A trial of the addition dot-task consists of the following steps: an initial blue dot array appears on the screen and is then covered by a rectangular box, then an additional array of blue dots hides within the box and lastly a set of red dots appears next to it. At the end of each trial, children have to estimate whether they saw more red dots or more blue dots. In the dot-task, a ratio effect on performance arises from the distance of the summed blue set and the red set. As the numerical difference or distance between the two sets becomes smaller, their ratio approaches 1 and performance declines (e.g. Barth et al., 2006). For example, if the two blue dot-sets add up to forty, it is easier to estimate the correct response when they are compared to a set of seventy red dots than to a set of forty. The large numerical distance makes their comparison much easier. It is postulated that this occurs because the mental representations of two numerical magnitudes, which are close to each other, overlap and are therefore harder to compare (Izard & Dehaene, 2008). This ratio effect is also presumed to be reflected in the participants' mean reaction response times (Noël, Rouselle, & Mussolini, 2005 as cited in De Smedt & Gilmore, 2011). This assumption, however, has not been tested in the nonsymbolic approximate arithmetic domain because tasks used so far did not allow reaction time (RT) registration. We developed a dot-task that permitted the recording of RT data and thus the acquisition of a more fine-grained illustration of children's nonsymbolic approximate arithmetic cognitive processes. Previous research with tasks assessing comparison of small nonsymbolic numerosities has demonstrated the ratio effect in RTs by showing children's performance being slower in harder to compare trials (Holloway & Ansari, 2009; Rouselle & Noël, 2007; Soltész, Szücs & Szücs, 2010).

So, what underlies the process of nonsymbolic approximate addition? In the dot-task, participants must mentally retain and add the two blue dot- sets, remember the summed numerosity and then compare it to the red dot-set. This procedure appears to involve working memory. Barth and colleagues (2006) already assumed working memory load involvement in their nonsymbolic approximate addition and subtraction tasks' implementation. However, to our knowledge, no previous study has examined in detail the role of WM in nonsymbolic approximate arithmetic processing.

Working memory and arithmetic

The most prominent theoretical account of Working Memory (WM) is the tripartite WM model, originally conceptualized by Baddeley and Hitch in 1974. According to this model, WM is a multicomponent cognitive architectural system that is responsible for the short-term storage and manipulation of a limited amount of elements during the execution of cognitive activities (Baddeley, 1986, 2002, 2003). It entails a master system, *the central executive (CE),* and two slave subsystems, the *phonological loop (PL)* and the *visuo-spatial sketchpad (VSSP).* The central executive component has a supervising role; it is an executive system which regulates and controls cognitive processes run by the two slave subsystems. The phonological loop is responsible for retaining verbal information, whereas visuo-spatial information is maintained within the visuo-spatial sketchpad. Since its original conceptualization, empirical accounts have lead to the development and extension of this multicomponent model (see Baddeley, 1996a; 2000; 2001; 2003). The role of the central executive was for a long time unclear. Based on accumulating findings, Repovš and Baddeley (2006; pp. 14) proposed that "*in the realm of working memory tasks, executive processes seem to be involved whenever information within the stores needs to be manipulated*". In other words, the slave subsystems are free of executive processes only when they involve simple representation and maintenance.

The literature distinguishes two kinds of methodological designs utilized for assessing the role of these WM components (Raghubar et al., 2009): experimental dual-task studies and correlational designs. The dual-task methodology is considered as the most reliable experimental design since it uncovers the on-line underlying WM resources allocated in complex task processing. However, it has been predominantly used in studies with adults (e.g. Fürst & Hitch, 2000; Lee & Kang, 2002; Trbovich & LeFevre, 2003). The dual task design involves the execution of a primary task (e.g. arithmetic task) while simultaneously performing a secondary task which loads – and therefore interferes with – a specific WM component. It is based on the principle that, if a specific WM component is necessary for the cognitive processing of the primary task, one will identify a performance breakdown or reaction time increase on either the primary or the secondary task in the corresponding interference condition compared to the conditions where these tasks were performed in a stand-alone form (baseline).

Dual-task studies with children are very limited. In their review, Raghubar and colleagues (2009) identify only two with primary school-aged children (Imbo & Vandierendonck, 2007a; McKenzie, Bull & Gray, 2003). McKenzie and colleagues (2003) examined the developmental changes in the use of the slave WM components in exact verbal symbolic arithmetic (i.e. with Arabic numbers in the form of a + b = c). Two age-groups of children were used: one with mean age 6.91 years and the other 8.94 years. Phonological and visuo-spatial interference occurred with the concurrent presentation of secondary tasks. In the respective interference conditions, children either heard irrelevant speech or looked at dynamic visual noise without needing to react to these secondary tasks. This type of interference is characterized as passive. In an active interference condition participants are asked to also respond to the secondary task (e.g. Imbo & Vandierendonck, 2007a). This way the interactive effect of the interference is indexed and thus performance breakdowns due to the load are reflected on either the primary or the secondary task. As highlighted by McKenzie et al. (2003), one reason for them to choose passive secondary tasks was because it was uncertain whether, especially the younger children, could perform active concurrent secondary tasks. Their results showed that younger children relied solely on visuo-spatial strategies when solving verbally presented exact symbolic arithmetic problems, whereas older children used also phonological strategies. Our study takes WM research in mathematical cognition a step further. We introduced for the first time active WM interference to children as young as preschoolers showing both the feasibility and the effectiveness of such an experimental design in this age group.

We know most about the early stages of learning arithmetic and the role of WM from studies using correlational designs. It has been argued that preschoolers' performance in arithmetic is in fact restricted due to their limited WM capacity (Klein & Bisanz, 2000). Specifically, Rasmussen and Bisanz (2005) demonstrated the developmentally differentiated relationship of WM components with distinct arithmetic problem formats. They tested 5- and 6-year old children's PL, VSSP and CE skills and their performance in two different arithmetic problem formats: verbal (using story problems) and nonverbal (using chips). Preschool children's performance on the nonverbal simple addition task was found to be related to their VSSP WM capacity, contrary to older children who relied on their PL capacity. The authors argued that preschool children make use of a mental model to represent objects and conduct arithmetic manipulations, contrary to older children who make use of phonological coding strategies. Their nonverbal task was nonsymbolic in nature. In their study, however, approximate arithmetic was not examined since exact responses were required for the arithmetic problems, with operand set sizes ranging from one to seven.

Essentially, Rasmussen and Bisanz's (2005) study revealed the importance of the VSSP WM component for early nonsymbolic arithmetic, but was limited by the fact that only one task was utilized to assess it. WM literature has shown evidence for the fractionation of this component into a visual and a spatial subcomponent (Baddeley, 2003; Darling, Della Sala, Logie, & Cantagallo, 2006; Logie, 1986). Notably, Hegarty and Kozhevnikov's (1999) results highlight the importance of this fractionation, since spatial and visual representations were shown to be differentially related to mathematical success. For this reason we designed both a visual and a spatial interference condition in order to test whether they play different roles in the process of mentally representing nonsymbolic approximate arithmetic information.

On the other hand, Noël's (2009) research on preschoolers' simple addition skills, emphasized the role of the CE component. Children were presented with drawings of objects, such as cows, with which they were asked to conduct basic additions. Contrary to the previously mentioned studies, here children were free to solve the problems in any way they preferred and could even use their fingers or tokens in the process. Noël's arithmetic problems were presented in a combined visual and verbal manner and in the given presentation format both symbolic and nonsymbolic information was involved. It was shown that in a free situation the predictive power of the CE appears stronger and more significant than that of the other components in preschoolers' simple addition.

Nevertheless, nonsymbolic approximate processing was not examined. To our knowledge, our study is the first to study the underlying WM processing in nonsymbolic approximate arithmetic.

The present study

Our main aim was to examine the relationship between nonsymbolic approximate arithmetic and WM as conceptualized by Baddeley's multicomponent model. Thus, a dual-task study was conducted with active phonological, visual, spatial and central executive interference during the completion of a nonsymbolic approximate addition task, i.e. the dot-task. Based on Rasmussen and Bisanz's (2005) findings, we hypothesized that its' underlying processing will depend on VSSP WM and not the PL in preschoolers. As indicated earlier, the CE appears to play an important role in children's arithmetic processing (see also Raghubar et al., 2009). During the implementation of the dot-task, the mental representation of the first appearing blue quantity set must be updated after the second one is presented in order to form the basis against which the red set can be compared. Based on this updating process, we also hypothesized the CE WM component being involved in nonsymbolic approximate arithmetic processing (Morris & Jones, 1990).

Our secondary aim was to replicate existing findings on preschoolers' ability to successfully conduct addition with large nonsymbolic approximate quantities (Barth et al., 2006). Moreover, with our dot-task we scoped for the use of RT data as an additional source of information, which will facilitate the acquisition of a more coherent picture of the processes underlying children's nonsymbolic approximate arithmetic skills.

METHOD

Participants

Sixty-two children (25 boys, 37 girls) aged from 5.23 years to 6.44 years (mean age 5.95 years) were recruited from five kindergartens (second grade according to the Dutch educational system) in two cities in the Netherlands. We included children from the whole ability range. None of these children had any learning problems, diagnosed developmental or retardation disorders and all of them completed testing. Written consent forms were acquired from their legal guardians.

Design

A dual-task interference design was used to examine the differential contribution of the WM components on nonsymbolic approximate arithmetic performance assessed with the dot-task. Children solved the primary task in five sessions: one without interference (dot-task alone) and the others together with the implementation of secondary tasks for phonological, visual, spatial and central executive interference.

Children's intelligence was assessed two months earlier in a group-wise manner (4 to 7 children) with the Raven's Colored Progressive Matrices (Raven, Raven & Court, 1998). Based on their mean and standard deviation scores, the children were divided into three intelligence groups (low, average, high). In order to control for any effects related to the order of presentation of tasks, task order was counterbalanced by using two presentation conditions: (a) visual stand-alone, PL stand-alone, spatial stand-alone, CE stand-alone, dual-visual, dual-PL, dual-spatial, dual-CE and dot-task stand-alone or (b) the exact opposite order. Half of the children of each intelligence group were assigned to the first order and the other half to the second. Independent sample t-test analyses indicated no order of presentation effect for all the dot-task trials ($p = .12$) nor for any of the secondary tasks, namely the visual stand-alone ($p = .21$) and dual ($p = .18$), the spatial stand-alone ($p = .51$) and dual ($p = .46$), the PL stand-alone ($p = .24$) and dual ($p = .28$) and lastly the CE stand-alone ($p = .76$) and dual ($p = .81$).

Material

All tasks were computerized and developed with the experimental software E-Prime, version 1.2 (Psychological Software Tools, Pittsburgh, PA, USA).

Primary task

On the basis of Barth et al.'s (2005, 2006, 2008) and Gilmore et al.'s (2010) studies, we developed a task for assessing children's nonsymbolic approximate addition skills. Our dot-task differed from other versions mainly on a response and trial construction level.

The task started with a welcoming cartoon figure and the experimenter introduced it as a computer game with dots. In line with Barth et al. (2008), 6 practice trials followed, in order for the children to fully understand the events of the task. Initially the experimenter asked the child to identify the blue and the red dots on the screen in order to check for cases of color blindness. Figure 1 illustrates the events of a dot-task trial and the instructions the experimenters narrated during practice. The duration of each animated

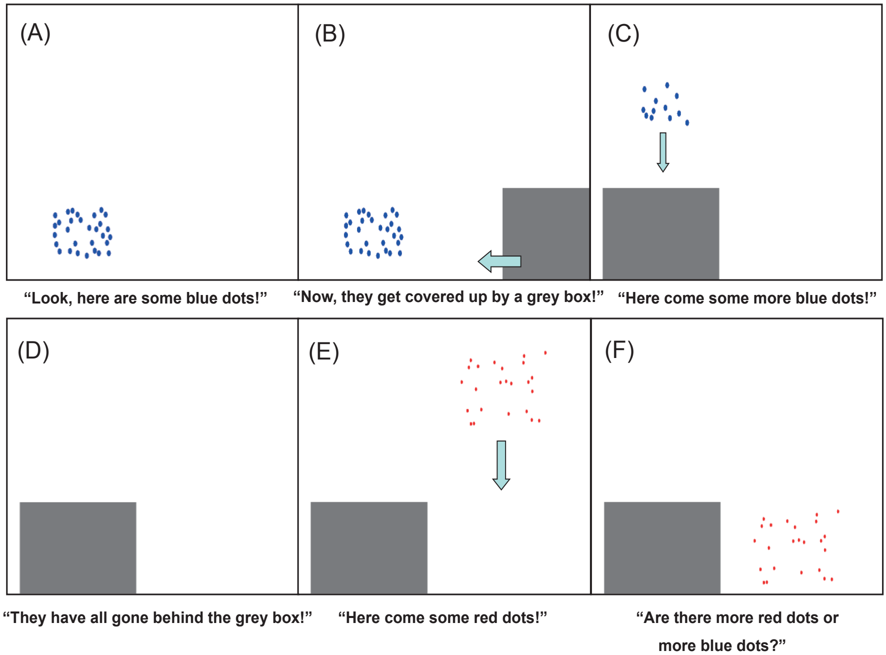

Figure 1. The dot-task; illustration of a trial and the narrated instructions during practice.

event within a trial (e.g. blue array falling down) was 1300ms and between each event there was a 1200ms wait interval. Consequently the sequence of presentation of the stimuli was too fast to allow the children to count the relatively large sets of dots.

In our study, children were instructed to decide as accurately and as fast as possible whether more blue dots or more red dots appeared on the screen and press the corresponding response button. Two response boxes were situated in front of them. They were instructed to press the left one with the blue sticker, if they thought the blue dots were more, or the right one with the red sticker, if they thought the red were more. Response registration (accuracy and reaction time) was initiated from the moment the complete red dot-array appeared on the right upper side of the screen. From that point on, participants had a maximum of 7000 ms to respond; thereafter the next trial would be initiated. A complete trial lasted approximately 15 seconds. There was a 300ms interval between the end of a trial and the initiation of the next one. Before the experimenter initiated the testing block, children were told that they would no longer receive feedback.

The actual game consisted of 24 testing trials (see Appendix, Table A1), which were presented in a random sequence. To control for responses being reliant on continuous quantity variables when comparing the summed blue set with the red dots, in half of the trials, dot size, total dot surface area, total dot contour length and density were positively correlated with numerosity while array size was negatively correlated with numerosity. The opposite relations occurred for the other half of the trials (Gilmore et al., 2010). Dot size was constant within the two blue arrays and variable across the summed blue and red arrays: 3 or 4.5 mm diameter (Barth et al., 2006). Our dot-stimuli were developed with MATLAB 7.5 R2007b.

Within the trials, the sum of blue dot arrays and the comparison red dot array differed by three ratios: 4:7 (easy), 4:6 (middle), 4:5 (difficult) and our numerosities ranged from 6 to 70 dots across all trials. In half of the trials within each ratio the comparison numerosity was larger and in the other half it was smaller. Similar to previous studies, we controlled for the use of any non-addition strategies, in order to be able to assess, for example, whether children resorted to a strategy such as always only pressing the red button. In previous studies (e.g. Barth et al., 2006; Barth et al., 2008) the criteria for ascribing a trial as one that predicts a specific non-addition strategy or not, were based on numerosity distances; e.g. a distance of 1, 2 or 5 was characterized as small. We believe, however, that a criterion based solely on the difference in dots can be characterized as arbitrary and that such a judgment should be also relative to the size of the other dot addends. Thus, on the trial construction level, we constructed our nonaddition strategy controls on ratio-based criteria. For example, to judge whether the distance "x" between the larger of the two blue addends and the red addend was far, medium or near (see strategy Near/Far in the Appendix) we calculated the ratio of their distance and ascribed it on one of three levels on the scale of 0 to 1. If $x \leq 0.33$ then the distance was ascribed as far, if $x \leq 0.66$ as medium and if $x \leq 0.99$ as near.

Secondary Tasks

Secondary tasks were developed to tap and interfere with the four different WM components during primary task completion. They were also performed in stand-alone conditions with a 15 sec delay replacing the primary task. When possible, we made use of several difficulty levels in order to cover the full span of performance variability in the corresponding WM skills. With this we aimed to examine whether there were different interference effects for the various difficulty spans in the secondary tasks. Furthermore, for the dual-task conditions, which had varying difficulty spans, the dot-task trials were counterbalanced across the spans based on ratio and in turn these trials were randomized within each span. This was done to prevent any outcome interference effects in the secondary tasks being related to ratio effects in the dot-task.

Visual WM. We developed a variation of the Abstract Patterns task which has been shown to successfully tap the Visual WM component (Logie & Pearson, 1997). As illustrated in Figure 2, a matrix pattern with half of its boxes white and the other half black, was displayed on the screen before and after each dot-task trial. In the stand-alone condition, instead of conducting a dot-task trial, children were instructed to look at a fixation cross for 15 secs. After the dot-task trial or the delay, a second abstract pattern (target) appeared. In half of the trials (12) the target was the same as the original whereas in the other half (12) it was different. The child was prompted to say aloud whether the second pattern was the same or different and the experimenter would register the response. There was no time-limit for this reaction. As in previous studies, pattern size began at 2 x 2 and increased by two squares in each difficulty span. Previous research has shown that children aged five to six are able to perform successfully on this task at least until span 4 (Logie & Pearson, 1997). The rationale of the dual-task design is that one's performance will break down on the hardest conditions. Therefore, our task's level of difficulty ranged from span 2 to 5 with six trials in each span. In all spans there were targets that differed only in one box; in span 3, 4 and 5 there were also targets that had two different boxes and in span 4 and 5 there was one case of a target that had 3 and one of 4 different boxes compared to the original pattern. An instructional slide would prompt the experimenter to inform the child about the initiation of a new span.

Spatial WM. The Corsi Blocks task (Ang & Lee, 2008) was adapted for our Spatial WM interference condition. In each trial a sequence of crosses appeared in nine randomly positioned blocks. A cross was displayed for 500 ms, then disappeared and subsequently one more cross appeared in a different block based on the corresponding span (see Figure 3 for an example trial). Children aged five to six have been shown to perform successfully on this task up to span 3, i.e. a spatial pattern made by three crosses. In order to cover the complete range of levels of difficulty variation our spatial task included four spans. After the dot-task trial (dual condition) or the delay (stand-alone condition), a target corsi blocks pattern was displayed. In half of the trials (12) the target was the same as the original and in the other half (12) it was different. Half of the target patterns (6) differed in location of appearance of the cross but not in sequence whereas the other half (6) in sequence but not location. Children were instructed to recall both sequence and location of appearance of the crosses. Their response (target same or different) was vocal and registered by the experimenter. Again, there was no time limit for these responses. Each span included six trial sequences.

Phonological WM. A Letter Span task was developed to interfere with the PL WM component. In this task participants heard a series of recorded consonants (one sec

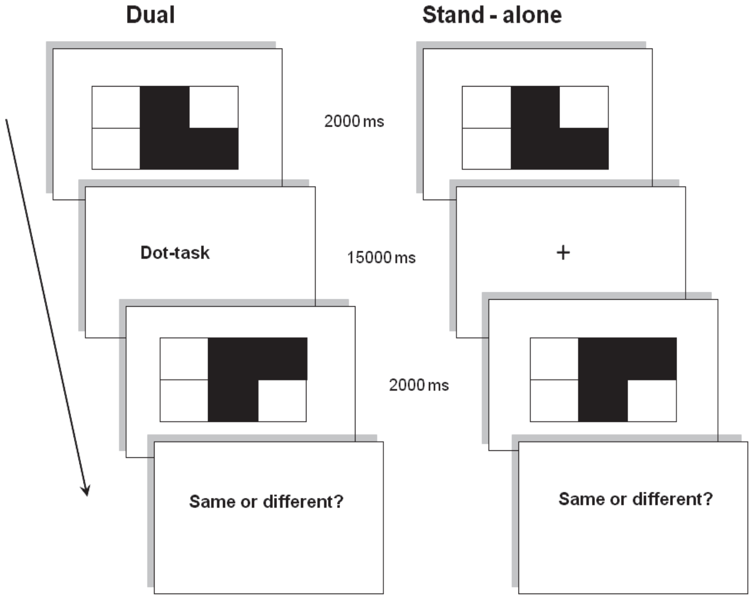

Figure 2. Illustration of the sequence of events in a span 3 trial of the Abstract Patterns task in the dual and stand-alone condition.

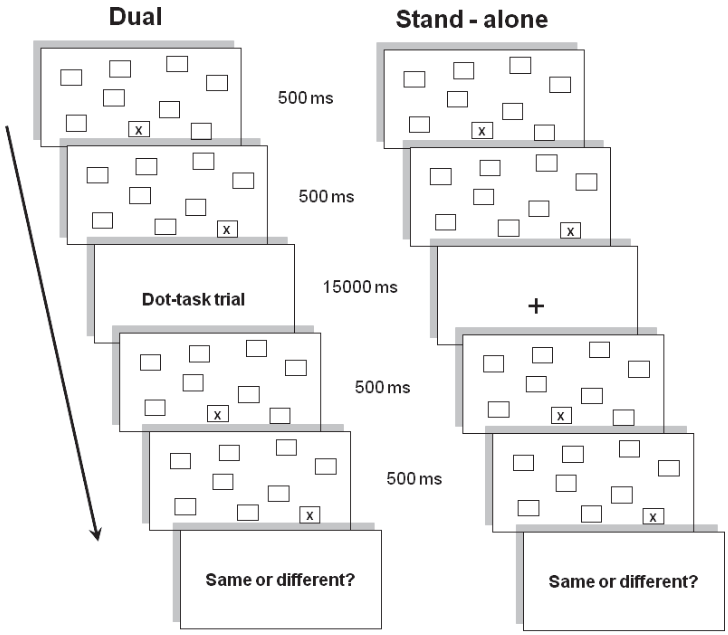

Figure 3. Illustration of the sequence of events in a span 2 trial of the Corsi Blocks task in the dual and stand-alone condition.

each) and were prompted to reproduce them after implementing a dot-task trial (dual condition) or the delay (stand-alone condition). The experimenter registered whether the reproduced response was correct or incorrect. Difficulty level was based on the number of consonants to be retained. Children aged five perform successfully up to sequences of three consonants (Chuah & Maybery, 1999). As in the cases of the visual and the spatial task, we included an additional span making the phonological task range from two to four consonant spans, with eight sequences in each span.

Consonant sequences were grouped according to Dutch pronunciation (Szmalec, Vandierendonck & Kemps, 2005). To avoid confusion the consonants B, P, M and N were excluded. We used the following groups of consonants: (D, T), (C), (F,S), (G), (H, K), (J), (L), (Q), (R), (V, W), (X) and (Z). Strings were made by random selection of a letter with no replacement of a consonant group within a string, no alphabetic pairs and no common combinations (Chuah & Maybery, 1999).

Central Executive WM. The Continuous Choice Reaction Time Task–Random (CRT-R; Imbo & Vandierendonck, 2007a) was used to load and interfere with the CE WM component. Due to the design of our study, the response procedure was different in our task compared to the original. Children heard high and low tones and they had to say aloud “A” or “O” correspondingly. Pilot testing showed that five year-old children found the sequence of the tones to be too fast, thus in our study the interval between two consecutive tones was longer (3000 and 3500 ms). Responses were recorded with voice-recorders and subsequently scored by the experimenter. In the dual-task condition, this task was performed concurrently with the dot-task, whereas in the stand-alone condition they performed the CRT-R for 5 mins (same as the duration of the dot-task).

Procedure

Children completed the five stand-alone and four dual-tasks within five sessions; four sessions of 30 minutes approximate duration (two tasks per session) and one of 6 minutes (dot-task alone). These sessions took place on six different days within a period of approximately two weeks. With the permission of the teacher, the experimenter took each child to a quiet room within the school setting, where they conducted the tasks. All tasks, apart from the secondary CE task, were performed on HP Compaq 6710b laptops with a 15-inch screen; the children were seated approximately 60cm away from the screen. The CE secondary task (CRT-R) was played on a Samsung NC20 12.1-inch notebook and for the PL condition headphones were used.

The experimenter introduced the tasks as games and the children were told that after each task they would receive a sticker. Firstly the experimenter explained the instructions to the child and then initiated the practice trials of the given task. The secondary tasks consisted of four practice trials whereas the primary task of six trials, consistent with Barth and colleagues' (2006) procedure. Therefore, in dual-task conditions children received these ten practice trials followed by four dual-task practice trials. The exception was the CE condition, where the task was practiced for half a minute in the stand-alone condition and together with the primary tasks' practice trials in the dual condition. During practice, children received computerized feedback with the display of a happy or mildly sad cartoon face at the end of each trial. Before initiating the testing block, the experimenter made sure the child had understood what he or she would have to do. During testing no feedback was provided. Throughout the games children were encouraged to stay focused and were reminded what they had to do if necessary. Every game ended with a very happy cartoon face indicating that they did a "good job" to reinforce them positively and sustain their interest and motivation.

Statistical analyses

In the literature one notices that it is still very common to apply ANOVA analyses on proportional data even though the nature of this data will very often violate the assumption of equal variances (Jaeger, 2008). Part of our data consisted of dichotomous data (e.g. a child's response was assessed as correct or incorrect). Trials with dichotomous responses were aggregated into frequency variables with a lower (zero) and upper limit (the maximum score). The distribution of these variables, therefore, was essentially binomial. Consequently, the variance close to the extremes of the scale was lower than the variance close to the midpoint of the scale (like in the case of proportions with 0 and 1 as extremes and .5 as midpoint), which can result in unequal variances in the comparison of two (experimental) conditions. Data transformations like the arcsine-square-root transformation are often not sufficient to mitigate this violation of the assumption of equal variances (Jaeger, 2008). For this reason, for our binomial accuracy data, we made use of a relatively new extension of logistic regression analyses for repeated measures; the so-called Generalized Estimating Equations or GEE analysis (Jaeger, 2008). GEE analysis can be seen as the repeated measures version of the Generalized Linear Model (not to confuse with the General Linear Model). To circumvent the inherent problems of proportions, GEE provides a link function of the predictors with the logit that is the natural log of the odds, i.e. p / (1 - p), where "p" stands for proportion. The logit (or 'log-odds') is not constrained by a lower and upper limit of the range of scores, but varies symmetrically from minus infinity to plus infinity around zero as midpoint of the scale. Hypothesis testing with GEE is based on maximum likelihood estimation. Therefore, the Wald χ^2 statistic is used instead of *F,* as is the case of regular ANOVA. The odds ratio (OR) constitutes a measure of effect size (Ferguson, 2009). OR gives the ratio between the odds of one (experimental) condition compared to the odds of another.

For our continuous outcomes, such as RT data and also the CRT-R scores, which had no theoretical upper limit and no practical lower limit, corresponding ANOVA tests were used.

RESULTS

We acquired accuracy scores for all stand-alone and dual-tasks; we also accumulated reaction time (RT) data for the dot-task on baseline and under the interference conditions for all trials on which children responded correctly. We first examined children's performance on the primary task (baseline). Subsequently, their baseline performance was compared to each dot-task interference condition in order to identify in which ones children's performance broke down. Lastly, WM demands were also examined with respect to secondary task performance.

Dot-task

Children succeeded in our stand-alone dot-task. They estimated significantly more frequently which of the sets was larger (60.48 %, Wald $\chi^2(1, 62) = 50.54$, $p < .001$, $OR = 1.5$) than they would at chance level (50 %). Specifically, they performed above chance on all three ratios: the easy ($p < .001$), the middle ($p < .001$) and the difficult ratio ($p < .05$). Their responses were not based on any physical features of the dots (continuous quantity variables) and they did not resort to any guessing non-addition strategies (see Appendix, Table A.2). Figure 4 shows the characteristic ratio effect for accuracy, Wald $\chi^2(2, 62) = 43.79$, $p < .001$. These results indicated a steady decrease in performance from the largest to the smallest ratio (closest to 1). A similar effect, however, was not acquired for their RTs, F (2, 122) = .74, $p = .929$.

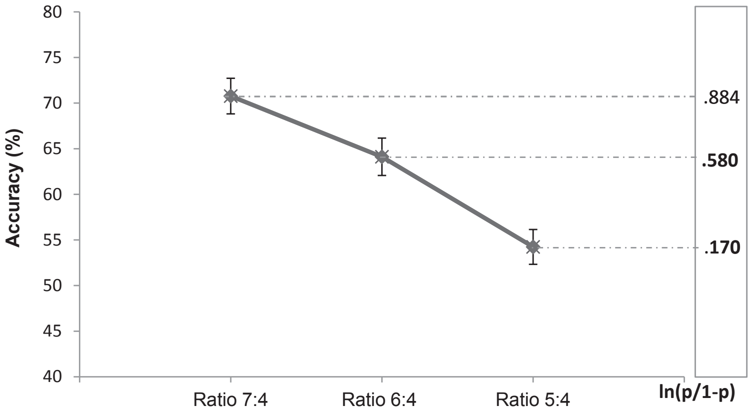

Figure 4. The characteristic ratio effect in children's accuracy on the primary task: performance (% correct) declines as the quantities' ratio approaches one, i.e. from the easy (7:4), to the middle (6:4) and the difficult (5:4) ratio. Error bars reflect standard error values. Logit values of the corresponding transformed original values are presented in the table on the right; note that "p" stands for "proportions"- estimated marginal means of the original scores.

Primary task: Baseline vs. interference conditions

To examine the effect of the interference conditions on nonsymbolic approximate addition, a 3 x 5 factorial GEE analysis was conducted with the 3 Ratios (easy, middle, difficult) and 5 Interference conditions (baseline, visual, spatial, PL, CE). Results showed the expected main interference effect, Wald $\chi^2(4, 62) = 45.69$, $p < .001$, a main ratio effect, Wald $\chi^2(2, 62) = 111.07$, $p < .001$ and no interaction effect. The main ratio effect followed the anticipated direction; performance decreased from easy to middle and difficult ratio across all conditions. With regard to the interference effect, compared to baseline children's performance broke down only in the CE interference condition ($p < .001$) and not in the visual ($p = .715$), the spatial ($p = .936$), or the PL condition ($p = .508$). Figure 5 illustrates the pattern of performance across the five conditions and the corresponding logit values. Performance in the CE interference condition was significantly above chance level (>50%), 54.82 %, Wald $\chi^2(1, 62) = 13.47$, $p < .001$, $OR = 1.2$.

For the RT data, a 3 x 5 repeated measures ANOVA was conducted with the 3 Ratios (easy, middle, difficult) and 5 Interference conditions (baseline, visual, spatial, PL, CE). For mean RTs, Mauchly's test indicated that the assumption of sphericity had been violated for Interference, $\chi^2(9, 61) = 49.17$, $p <.001$, Ratio, $\chi^2(2, 61) = 7.28$, $p = .026$, and their interaction, $\chi^2(35, 61) = 123.7$, $p <.001$. Therefore, degrees of freedom were corrected using

Greenhouse-Geisser estimates. As expected, a main interference effect, $F(2.8, 170.94) = 10.25$, $p < .001$, was found (Figure 6). No main ratio or interference by ratio interaction effect was shown. Consistent with the accuracy score results, children's mean RT was significantly higher than that of baseline only in the interference condition where their central executive (CE) was loaded ($p < .001$). The spatial ($p = 1.000$), the visual ($p = .617$) and the PL ($p = .764$) interference conditions did not reach significant difference to that of baseline.

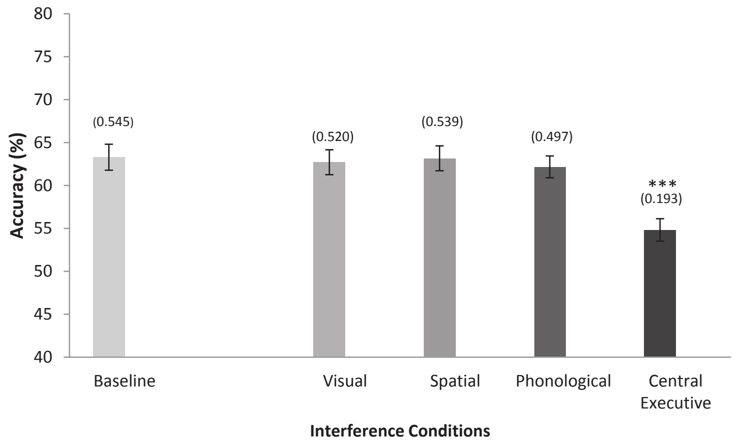

Figure 5. Children's performance (percentage accuracy) across the interference conditions. Parentheses include the corresponding logit values, ln(p/1-p), where "p" stands for proportion. Error bars reflect corresponding standard error values. *** $p < .001$.

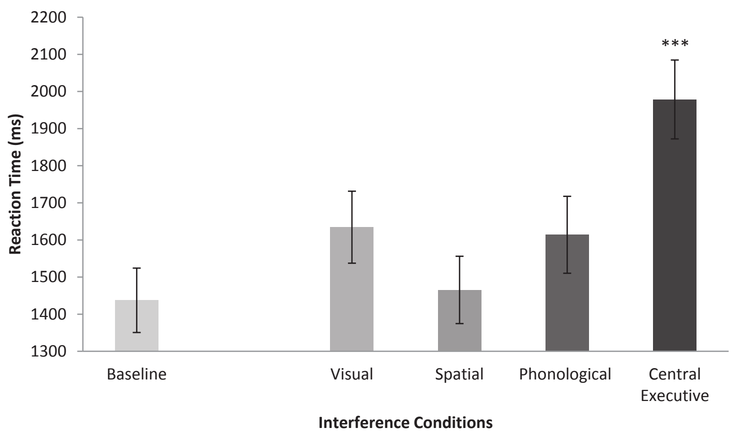

Figure 6. Mean RTs across the interference conditions. Only the central executive WM interference condition was significantly higher ($p < .001$) than that of baseline, i.e. dot-task alone. Error bars reflect corresponding standard error values.

Secondary tasks

WM loading demands can be indexed by performance breakdowns on either the primary or the secondary tasks. A paired samples t-test of the difference in performance on the CE secondary task (CRT-R) between the stand-alone and the dual condition indicated a significant performance breakdown in the second, $t\ (61) = 8.11$; $p < .001$. For our visual task (Abstract Patterns), a 2 x 4 factorial GEE analysis was conducted between the two conditions (stand-alone and dual) and the four difficulty spans. Results indicated no condition or condition by span effect. Only a significant main span effect was found, Wald $\chi^2(3, 62) = 21.37, p < .001$. For the spatial task (Corsi Blocks), a similar 2 x 4 factorial GEE analysis resulted in both a main span, Wald $\chi^2(3, 62) = 45.43$, $p < .001$, and a condition by span interaction effect, Wald $\chi^2(3, 62) = 12.66$, $p = .005$, indicating an involvement of the spatial WM component. To further elaborate on this interaction effect, we conducted corresponding GEE analyses for each Span of difficulty of the Corsi Blocks task (Figure 7). The expected condition effect was found only for the easiest span (Span 1), Wald $\chi^2(1, 62) = 9.88$, $p = .002$. Parameter estimate results showed that children's performance significantly dropped in the dual condition compared to the stand-alone one in this span, $OR = 0.6$. No corresponding condition effect was found for the rest of the difficulty spans.

Similarly, for the PL task (Letter Span), a 2 x 3 GEE analysis was conducted over the two conditions (dual and stand-alone) and the three difficulty spans. A main span, Wald $\chi^2(2, 62) = 337.35$, $p < .001$, and, surprisingly, a main condition, Wald $\chi^2\ (1, 62) = 5.42$, $p = .020$, effect were found, but no interaction effect. Parameter estimates indicated that children's performance in the dual PL task significantly dropped compared to the stand-alone version, Wald $\chi^2(1, 62) = 4.78$, $p = .029$, $OR = 0.7$).

DISCUSSION

Our aim was to uncover the relationship between nonsymbolic approximate arithmetic and WM. For this purpose we conducted for the first time a dual-task study with preschool children in which we actively interfered with the phonological, visual, spatial and central executive WM components while implementing a nonsymbolic approximate addition task, i.e. the dot-task. At baseline, our results replicated previous findings that show the ability of preschool children to perform above chance level in the dot-task. The characteristic ratio effect in accuracy was also replicated. With RT data, however, we did

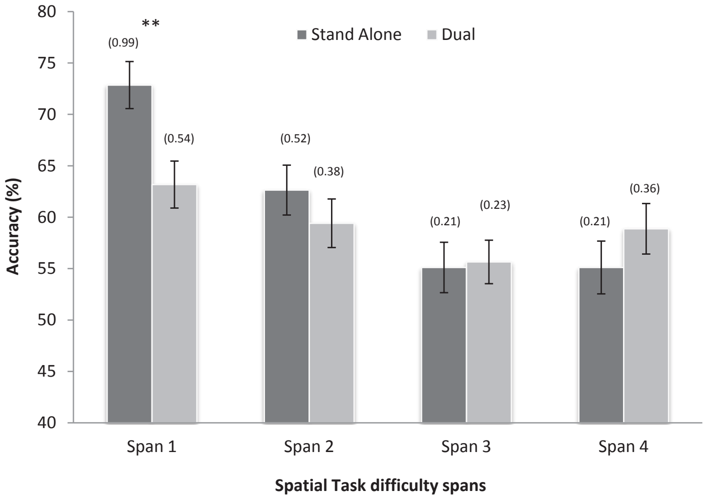

Figure 7. Children's performance (% accuracy) across the four difficulty spans of the Spatial Task (Corsi Blocks). Their performance significantly dropped in the first difficulty span. Parentheses include the corresponding logit values, ln(p/1-p), where "p" stands for proportion and error bars reflect standard error values. ** $p < .01$

not find a similar effect. Regarding the role of WM in nonsymbolic approximate addition processing, results confirmed our hypothesis showing a predominant clear-cut effect of the Central Executive (CE) component. Visual and spatial WM were not confirmed as important factors for nonsymbolic approximate processing. Surprisingly though, we found a Phonological Loop (PL) effect on the secondary task performance; we argue that this effect reflected the role of this component in action control.

A precondition for examining the role of WM in nonsymbolic approximate addition was to show that we indeed provoked approximate addition with our task. Nonsymbolic approximate addition results replicated previous findings that prove preschool children to be able to successfully add large nonsymbolic quantities prior to having received any formal arithmetic instruction in school (Barth et al., 2005, 2006, 2008; Gilmore et al., 2010). Likewise, they did so without resorting to any strategies other than addition per se – e.g. by using systematic response preferences, such as choosing only the red quantity as being larger – or by basing their responses on perceptual characteristics other than the numerosity of the nonsymbolic stimuli. Children performed above chance level and the characteristic ratio-effect was shown for accuracy, which supports the assumption of the existence of a mental number line system underlying approximate quantity estimation abilities (Izard & Dehaene, 2008).

Our study is the first to examine the ratio effect with RTs in nonsymbolic approximate addition. Iuculano, Moro and Butterworth (2011) used RT data on a similar nonsymbolic approximate addition task but did not examine the corresponding ratio effect. To our surprise the ratio effect was not evident with the RT data. This effect has been consistently demonstrated in previous research (e.g. Holloway & Ansari, 2009; Rouselle & Noël, 2007; Soltész et al., 2010), which, however, made use of nonsymbolic numerical magnitude comparison tasks. These tasks differ from our nonsymbolic approximate addition task in three main aspects; they (a) do not entail the element of addition, (b) deal with much smaller numerosities, ranging from one to nine and, (c) call for a response to simultaneously presented stimuli. This last element suggests that perhaps with the current animated dot-task design it was not possible to collect accurate RTs. The RT interference effect, which was consistent with the corresponding accuracy result, however, contradicts this interpretation. Perhaps our RT measurement was reliably sensitive to the interference effect but not sensitive enough to capture the ratio effect due to the inherent design of this task. On the other hand, the remaining two elements of differentiation between the dot-task and the previous comparison tasks, suggest possible differences in the skills that these tasks actually attempt to measure. Our dot-task is a far more complex cognitive task, where children were asked to add large quantities that ranged from 6 to 70.

For long, many assumed that the skills assessed with the nonsymbolic magnitude comparison and the nonsymbolic approximate addition tasks could be placed under the same 'theoretical umbrella; that of the so-called "Approximate Number System (ANS)". This is mainly because of the consistent and common underlying signature effects such as that of the ratio and distance effect (Gilmore, Attridge & Inglis, 2011). Gilmore, Attridge & Inglis (2011), however, provided evidence for the lack of correlation between participants' performances in these tasks. As one of their explanations they suggest the possibility that these tasks may draw on different domain-general abilities, such as WM. In accordance to that argument, we postulate that the nonsymbolic comparison tasks and nonsymbolic approximate addition tasks, such as our dot-task, may call upon different underlying cognitive processes. Of course, the lack of an RT ratio effect must be replicated and further research is needed for its elaborate explanation. Moreover, future research should determine the different mechanisms underlying nonsymbolic magnitude comparison and nonsymbolic approximate addition.

With regard to the role of WM**,** our findings confirmed our main expectation. WM underlies nonsymbolic approximate addition processing. For interference during the primary task (dot task), our results on both the accuracy and RT data revealed a clear-cut interference effect. Specifically, as expected, preschoolers' performance was hindered in the CE interference condition. There is of course also the matter of the strategic tradeoff between the primary and the secondary tasks. Comparisons of performance between the secondary tasks conducted in the stand-alone and in the dual-task condition indexed once again a breakdown on the CE secondary task. Our findings, therefore, demonstrated a coherent picture for the necessity of CE WM demands. This result is consistent with previous research demonstrating the importance of executive resources in children's mathematical cognition (Noël, 2009; Raghubar et al., 2009). The exact role of the CE demands further elucidation. The CE task that we used, namely the CRT-R, is a task widely utilized to tap the CE (Imbo & Vandierendonck, 2007a; Tronsky, McManus & Anderson, 2008) as a homunculus subcomponent of WM. The functions of the CE, however, can be further fractionated (Repovš & Baddeley, 2006). During the dot-task, a participant must mentally update the mental representation of the first blue array with the second in order to form a summed set, which can then be compared with the red array. Future research should determine whether it is specifically the executive process of updating that is required during nonsymbolic approximate addition.

On the basis of Rasmussen and Bisanz's (2005) findings on preschoolers' nonsymbolic arithmetic, we had initially also hypothesized a predominant role for the visuospatial component of WM in nonsymbolic approximate arithmetic. We, therefore, explored the effect of the Visual and Spatial WM subcomponents. Surprisingly, our results revealed no significant effect for the Visual and hardly any for the Spatial WM component since the only effect found for the latter was limited to its easiest span. In a dual-task design, if a WM interference effect is to be assumed, it must be evidenced at least in the hardest ratios. Taking a closer look at Figure7, which depicts children's performance in the different span levels of the spatial secondary task in the dual and the stand-alone conditions, one notices that performance drops close to chance (50%) after the easiest span in all conditions. It appears there was a floor effect. We believe that this task was too hard for our children, resulting in a limited variability of performance, which in turn did not allow for any interference effects to be visible. We advise future studies to make use of easier visuospatial interference tasks in order to illuminate the role of the visuospatial sketchpad in nonsymbolic approximate arithmetic. An alternative explanation for this surprising result may be derived from the early studies examining the cognitive processing of expert versus novice chess players (Baddeley, 1996b, 2002; Gobet, 1997). Expert chess playing has been found to not be a result of higher visuospatial WM processing but rather due to the advanced pattern recognition level of the player. Similarly, it is possible that in this assumed innate skill of nonsymbolic approximate arithmetic, some sort of visuo-spatial mental operation does take place, which is not, however, adjunct to visuospatial WM.

However, apart from the preceding arguments, why did Rasmussen and Bisanz (2005) find the VSSP playing a predominant role in nonsymbolic processing and we did not? In their nonsymbolic addition task an experimenter would show a number of chips to the child, cover them up with a box and then he or she would add more under this box. Children were asked to replicate the amount of chips they saw with their own collection of chips. Operands in this task ranged from one to five, answers from three to eight and it necessitated an exact response. Thus, our differentiated findings imply also differences in the underlying cognitive processing between the two tasks. To our knowledge, research, thus far, has examined the differentiation between exact and approximate symbolic arithmetic processing (Kucian, von Aster, Loenneker, Dietrich, & Martin, 2008) but not between exact and approximate nonsymbolic arithmetic processing. What arises from our pattern of findings is that in nonsymbolic approximate processing, preschool children ultimately rely on their CE for successful implementation. In our dot-task, children could not represent each object/dot separately, as in the case of Rasmussen and Bisanz's (2005) task and thus, due to the large amount of dots, the CE component takes over and compensates by processing condensed whole arrays of dots and updating them within WM. It would be interesting for future studies to examine this assumption by specifically examining the differences in cognitive resources allocated for the processing of nonsymbolic exact and approximate arithmetic.

Unexpectedly, secondary task performance results also identified PL involvement. According to Krajewski and Schneider's (2009) theoretical model, children from a very young age start utilizing quantity discrimination words such as "much" or "more". It may be assumed that children made use of such a strategy to solve the nonsymbolic approximate arithmetic problems, i.e. by applying phonological tags on the arrays presented. Such an explanation, even though interesting, is also unsafe. Other studies have shown children to start utilizing phonological WM and corresponding strategies at a later age (McKenzie, Bull & Gray, 2003; Rasmussen & Bisanz, 2005). We believe, therefore, that this unexpected result of PL involvement was shown due to the unavoidable instructions that were given to the children during dot-task implementation. During testing we observed these young children being easily distracted while conducting the given complex tasks. For this reason, it was necessary in some occasions to give them instructions in order to sustain their attention during dot-task implementation such as "look at the dots", "pay attention". It is very plausible that this may be the practical explanation of the PL interference effect evident on the corresponding secondary task when performed under the dual-task condition. Children heard the verbal instructions and at the same time had to remember the series of letters. This explanation is in line with the findings that regard the PL also as playing a role in the control of one's behavior (Baddeley, 2003; Baddeley, Chincotta, & Adlam, 2001). At the same time, this observation constitutes a limitation of our study. Future dual-task research should focus in developing a context where no such verbal instructions are needed, so that clear WM PL storing implications can be concluded.

This study is also limited by the fact that different WM task-designs were utilized to load and interfere with the corresponding WM abilities. The CE interference task was a continuous one, whereas the rest of the secondary tasks took place "before" and "after" each primary-task trial and entailed discrete levels of difficulty. Nevertheless, this is common practice within the dual-task literature (e.g. Imbo & Vandierendonck, 2007b) due to the practical restrictions of wanting to load and interfere purely on a specific WM component without also interfering with the actual skill (not related to WM) that the primary task is tapping. We argue, however, that our results cannot be interpreted based on the differences between the designs of the tasks. Performance in each interference condition was only compared with that of baseline. In other words, interference conditions were not compared amongst each other. Also, the effects of disruption of the CE cannot be attributed to a higher task difficulty, as all other WM tasks included harder levels of difficulty than usual performance in this age and analyses were conducted on the span-level. Future research in this field could pursue the design of similar WM interference tasks that would allow also the examination of the difference of effects between each condition. For example, future studies may alternatively use articulatory suppression, where in the dual condition participants repeat an irrelevant word such as "the", as an alternative to our PL interference condition (see Baddeley, 2001). Furthermore, innovative VSSP interference conditions could be developed, such as those used by Lanfranchi and colleagues (2012), that interfere with the primary task during its completion. The important issue in developing and using these secondary tasks for a dual-task study is that they tap the different WM components as purely as possible. Our secondary tasks were developed in that manner. The tasks we used for tapping the slave subsystems of WM were free of executive resources since they necessitated sole representation and maintenance of the corresponding information (Repovš & Baddeley, 2006). On the other hand, the CE secondary task necessitated manipulation within the store. Actually, it called for manipulation within the PL store, since the task had verbal characteristics. Future research should indicate if this CE interference result would also be evident in a condition where the corresponding task needed manipulation within the VSSP store.

The findings of this study generate methodological as well as cognitive, developmental and applied educational psychology implications. We demonstrated that effective dual-task studies with active WM interference can be conducted with children as young as preschoolers. Nonsymbolic approximate representations have been characterized as being central to human knowledge of mathematics (Gilmore & Spelke, 2008). It is even assumed that nonsymbolic approximate arithmetic comprises the building blocks on top of which symbolic exact arithmetic skills are developed and enhanced (Mundy & Gilmore, 2009). We showed that preschoolers' nonsymbolic approximate addition skills necessitate central executive resources. Now, the question is raised of whether it is actually nonsymbolic approximate skills that play a role in later math development or do these skills mediate the effect of WM processing on mathematical achievement? Our findings constitute a stepping-stone in the path for uncovering and understanding the underlying cognitive architecture of early arithmetical skills.

REFERENCES

Ang, S. Y., & Lee, K. (2008). Central executive involvement in children's spatial memory. *Memory, 16*, 918-933. doi: 10.1080/09658210802365347

Baddeley, A. D. (1986). *Working memory*. Oxford, UK: Oxford University Press.

Baddeley, A. D. (1996a). *Exploring the central executive*. *Quarterly Journal of Experimental Psychology A, 49*, 5–28. doi:10.1080/713755608

Baddeley, A. D. (1996b). Working memory in chess. *Memory and Cognition, 24*, 83-93.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417–423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. (2002). Is working memory still working? *European Psychologist, 7,* 85–97. doi:10.1027//1016-9040.7.2.85

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorder, 36*, 189-208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A., Chincotta, D., & Adlam, A. (2001). Working memory and the control of action: evidence from task switching. *Journal of Experimental Psychology: General, 130*, 641-657. doi:10.1037//0096-3445.130.4.641

Baddeley, A., & Hitch, G. (1974). *Working Memory*. In G. Bower (Ed.), The psychology of learning and motivation (pp. 47-90). New York: Academic Press.

Barth, H., Beckmann, L., & Spelke, E. S. (2008). Nonsymbolic, approximate arithmetic in children: abstract addition prior to instruction. *Developmental Psychology, 44*, 1466-1477. doi:10.1037/a0013046

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Nonsymbolic arithmetic in adults and young children. *Cognition, 98*, 199-222. doi:10.1016/j.cognition.2004.09.011

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 of the United States of America, 102*, 14116-14121. doi:10.1073/pnas.0505512102

Bull, R., Espy, K. A., & Wiebe, S. A. (2008). Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205-228. doi:10.1080/87565640801982312

Chuah, Y. M., & Maybery, M. T. (1999). Verbal and spatial short-term memory: common sources of developmental change? *Journal of Experimental Child Psychology, 73*, 7-44. doi:10.1006/jecp.1999.2493

Darling, S., Della Sala., Logie, R. H., & Cantagallo, A. (2006). Neuropsychological evidence for separating components of visuo-spatial working memory. *Journal of Neurology, 253*, 176-180. doi:10.1007/s00415-005-0944-3

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. doi:10.1016/j.jecp.2009.01.010

DeStefano, D. & LeFevre, J. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353-386. doi:10.1080/09541440244000328

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. (2007). School readiness and later achievement. *Developmental Psychology, 43,* 1428-1446. doi:10.1037/0012-1649.43.6.1428

Ferguson, C. J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice, 40*, 532-538. doi:10.1037/a0015808

Finnie, R., & Meng, R. (2001). Cognitive skills and the youth labour market. *Applied Economics Letters, 8*, 675-679. doi:10.1080/13504850110037877

Flombaum, J. I., Junge, J. A., & Hauser, M. D. (2005). Rhesus monkeys (Macaca mulatta) spontaneously compute addition operations over large numbers. *Cognition, 97*, 315-325. doi:10.1016/j.cognition.2004.09.004

Fürst, A. J., & Hitch, G. J. (2000). Separate roles for executive and phonological components of working memory in mental arithmetic. *Memory and Cognition, 28*, 774-782

Geary, D. C., Bailey, D. H., Littlefield, A., Wood, P., Hoard, M. K., & Nugent, L. (2009). First-grade predictors of mathematical learning disabilities: A latent-class trajectory analysis. *Cognitive Development, 24*, 411-429. doi:10.1016/j.cogdev.2009.10.001

Gilmore, C. K., Attridge N., & Inglis, M. (2011). Measuring the approximate number system. *The Quarterly Journal of Experimental Psychology, 64*, 2099-2109. doi:10.1080/17470218.2011.574710

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Nonsymbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition, 115*, 394-406. doi:10.1016/j.cognition.2010.02.002

Gilmore, C. K., & Spelke, E. S. (2008). Children's understanding of the relationship between addition and subtraction. *Cognition, 107*, 932-945. doi:10.1016/j.cognition.2007.12.007

Gobet, F. (1997). A pattern-recognition theory of search in expert problem solving. *Thinking and reasoning, 3*, 291-313. doi:10.1080/135467897394301

Gullick, M. M., Sprute L. A., & Temple, E. (in press). Individual differences in working memory, nonverbal IQ, and mathematics achievement and brain mechanisms associated with symbolic and nonsymbolic number processing. *Learning and Individual Differences*, 21, 644–654. doi:10.1016/j.lindif.2010.10.003

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103*, 17-29. doi:10.1016/j.jecp.2008.04.001

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*, 684-689. doi:10.1037//0022-0663.91.4.684

Imbo, I., & Vandierendonck, A. (2007a). The development of strategy use in elementary school children: working memory and individual differences. *Journal of Experimental Child Psychology, 96*, 284-309. doi:10.1016/j.jecp.2006.09.001

Imbo, I., & Vandierendonck, A. (2007b). The role of phonological and executive working memory resources in simple arithmetic strategies. *European Journal of Cognitive Psychology, 19*, 910-933. doi:10.1080/09541440601051571

Iuculano, T. Moro, R., & Butterworth, B. (2011). Updating working memory and arithmetical attainment in school. *Learning and Individual Differences, 21,* 655-661. doi:10.1016/j.lindif.2010.12.002

Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106,* 1221–1247. doi:10.1016/j.cognition.2007.06.004

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434-446. doi:10.1016/j.jml.2007.11.007

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology, 54*, 105-115.

Krajewski, K., & Schneider, W. (2009). Exploring the impact of phonological awareness, visual–spatial working memory, and preschool quantity–number competencies on mathematics achievement in elementary school: Findings from a 3-year longitudinal study. *Journal of Experimental Child Psychology, 103,* 516-531. doi:10.1016/j.jecp.2009.03.009

Kucian, K., von Aster, M., Loenneker, T., Dietrich, T., & Martin, E. (2008). Development of neural networks for exact and approximate calculation: a fMRI study. *Developmental Neuropsychology, 33*, 447-473. doi:10.1080/87565640802101474

Lanfranchi, S., Baddeley, A., Gathercole, S. & Vlanello, R. (2012). Working memory in down syndrome: is there a dual task deficit? *Journal of Intellectual Disability Research, 56,* 157-166. doi:10.1111/j.1365-2788.2011.01444.x

Lee, K. M., & Kang, S. Y. (2002). Arithmetic operation and working memory: differential suppression in dual tasks. *Cognition, 83*, B63-68.

Logie, R. H. (1986). Visuo-spatial processing in working memory. *The Quarterly Journal of Experimental Psychology, 38*, 229-247.

Logie, R. H., & Baddeley, A. D. (1987). Cognitive processes in counting. *Journal of Experimental Psychology, Learning. Memory & Cognition, 13*, 310-326.

Logie, R. H., & Pearson, D. G. (1997). The inner eye and the inner scribe of visuo-spatial working memory: Evidence from developmental fractionation. *European Journal of Cognitive Psychology, 9*, 241-257.

McCrink, K., & Wynn, K. (2004). Large number addition and substraction by 9-month-old infants. *Psychological Science, 15*, 776-781. doi:10.1111/j.0956-7976.2004.00755.x

McKenzie, B., Bull, R., & Gray, C. (2003). The effects of phonological and visual-spatial interference on children's arithmetical performance. *Educational and Child Psychology, 20*, 93-108.

Morris, N., & Jones, D. M. (1990) Memory updating in working memory: The role of the central executive. *British Journal of Psychology, 81*, 111-121.

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Pychology, 103*, 490-502.

Noël, M. P. (2009). Counting on working memory when learning to count and add: a preschool study. *Developmental Psychology, 45*, 1630-1643.

Noël, M. P., & Rouselle, 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*: 165.

Noël, M. P., Rousselle, L., & Mussolin, C. (2005). Magnitude representation in children: Its development and dysfunction. In J. I. D. Campbell (Ed.), Handbook of mathematical cognition (pp. 179–195). Hove, UK: Psychology Press.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian Indigene Group. *Science, 306*, 499-503.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference and cognitive approaches. *Learning and Individual Differences, 20*, 110-122.

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology, 91*, 137-157.

Raven, J., Raven, J.C., & Court, J.H. (1998). *Manual for the Raven's Progressive Matrices and Vocabulary Scales*. Oxford, England: Oxford Psychologist Press.

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience, 139*, 5-21.

Reyna, V. F., & Brainerd C. J. (2011). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences, 17*, 147-159. doi:10.1016/j.lindif.2007.03.010

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, *102*, 361–95. doi:10.1016/j.cognition.2006.01.005

Soltész, F., Szücs, D., & Szücs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: A developmental study. *Behavioral and Brain Functions, 6*, 1-14.

Szmalec, A., Vandierendonck, A., & Kemps, E. (2005). Response selection involves executive control: evidence from the selective interference paradigm. *Memory and Cognition, 33*, 531-541.

Taylor, T. (1986). *Iamblichus' Life of Pythagoras*. Rochester, Vermont: Inner Traditions International.

Trbovich, P. L., & LeFevre, J. A. (2003). Phonological and visual working memory in mental addition. *Memory and Cognition, 31*, 738-745.

Tronsky, L. N., McManus, M., & Anderson, E. C. (2008). Strategy use in mental subtraction determines central executive load. *The American Journal of Psychology, 121*, 189-207.

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

APPENDIX

Nonaddition and continuous quantity strategy analyses

Children's performance on the nonsymbolic approximate addition task was further examined in order to assess whether they resorted to any strategies not related to addition (Barth et al., 2006; Gilmore et al., 2010). The 24 trials (Table A.1) were constructed in such a manner in order to allow for post-hoc examination of possible nonaddition strategies. Based on Barth and colleagues (2008), in Table A1 a trial that is listed as 1 indicates that the corresponding strategy predicts the correct response on this trial, -1 signifies that this strategy would lead to the incorrect response, whereas 0 means that this strategy does not provide a clear prediction.

Table A.2 summarizes our findings on the nonaddition strategy analyses. Firstly, we tested whether children demonstrated a response bias, i.e. a tendency to press only the blue button or the red button (strategies Blue and Red). If that was the case then children who chose this strategy would perform at 50% (since in half of the trials the correct response was the blue one and in the other half the red). This was not the case, they performed above chance level (50%). Secondly, we examined the case where children would base their response only on the distance level between the larger blue addend and the red addend (Near/far strategy). If there is a small distance (near), then the answer is likely to be in favor of the sum of the blue. On the other hand, if this distance is large (far) then the red response is most likely to be chosen. In our trials, distances between quantities were ratio-based; above chance level performance was tested on trials where this distance was medium. Again, results showed that this strategy was not used.

Thirdly, we tested whether children tended to compare only the first (B1vsR) or only the second blue array (B2vsR) with the red comparison set. If these strategies were used then children would perform below chance level on the trials that predicted the incorrect response.

Ratio based criteria were utilized to judge whether the given strategy was predictive or not of the correct or incorrect response in a given trial. Performance both on the B1vsR and B2vsR strategies was proven to be above chance level. Lastly, we examined whether children based their response only on the range of the size of the red array (RedRange), e,g. by choosing the red if it seemed large, or blue if the red array seemed quite small. Performance above chance level on this strategy was tested with trials where the size of the red array was in middle range. Results showed once more that this strategy had not been used.

Table A.1. *The dot-task testing trials.*

Trial	Ratio	1st blue array	2nd blue array	Sum of blue	Red array	Correct Response	Nonaddition strategies [a]						Continuous Quantity [c]
							Near/far	Blue	Red	B2 vs R	B1vsR	RedRange	
1	4:7	6	6	12	21	Red	Far (0.29)	-1 [b]	1	1	1	Middle	*B*
2	4:7	8	12	20	35	Red	Medium (0.34)	-1	1	1	1	Middle	*A*
3	4:7	15	13	28	49	Red	Far (0.31)	-1	1	1	1	Large	B
4	4:7	16	20	36	63	Red	Far (0.32)	-1	1	1	1	Large	A
5	7:4	20	8	28	16	Sum Blue	Near (0.80)	1	-1	-1	0	Small	B
6	7:4	30	12	42	24	Sum Blue	Near (0.80)	1	-1	-1	0	Middle	A
7	7:4	6	50	56	32	Sum Blue	Medium (0.64)	1	-1	1	-1	Middle	B
8	7:4	7	63	70	40	Sum Blue	Medium (0.63)	1	-1	1	-1	Middle	A
9	4:6	8	8	16	24	Red	Far (0.33)	-1	1	1	1	Middle	B
10	4:6	16	8	24	36	Red	Medium (0.44)	-1	1	1	1	Middle	A
11	4:6	12	20	32	48	Red	Medium (0.42)	-1	1	1	1	Large	B
12	4:6	20	20	40	60	Red	Far (0.33)	-1	1	1	1	Large	A
13	6:4	11	7	18	12	Sum Blue	Near (0.92)	1	-1	-1	0	Small	B
14	6:4	33	7	42	28	Sum Blue	Near (0.85)	1	-1	-1	0	Middle	A
15	6:4	25	35	60	40	Sum Blue	Near (0.88)	1	-1	0	-1	Middle	B
16	6:4	10	26	36	24	Sum Blue	Near (0.92)	1	-1	0	-1	Middle	A
17	4:5	7	9	16	20	Red	Medium (0.45)	-1	1	1	1	Small	B
18	4:5	12	12	24	30	Red	Medium (0.40)	-1	1	1	1	Middle	A
19	4:5	24	8	32	40	Red	Medium (0.60)	-1	1	1	1	Middle	B
20	4:5	34	6	40	50	Red	Near (0.68)	-1	1	1	1	Large	A
21	5:4	6	14	20	16	Sum Blue	Near (0.88)	1	-1	0	-1	Small	B
22	5:4	15	50	65	52	Sum Blue	Near (0.96)	1	-1	0	-1	Large	A
23	5:4	32	8	40	32	Sum Blue	Near (1.00)	1	-1	-1	0	Middle	B
24	5:4	40	10	50	40	Sum Blue	Near (1.00)	1	-1	-1	0	Middle	A

[a] These columns present information for the given trials with regard to the usage of possible nonaddition strategies: Near/far = response based on the ratio distance between the larger blue addend and the red; Blue = only the blue response is chosen; Red = only the red response is chosen; B2vsRstr = only the second blue addend is compared; B1vsRstr only the first blue addend is compared, RedRange = response is based on the relative size of the red array. [b] 1 = predicts correct answer for that trial, -1: does not predict correct answer for that trial, 0 = does not provide a clear prediction.
[c] Continuous quantity conditions: A = dot size, total dot surface area, total dot contour length and density positively correlated with number while array size negatively correlated with number; B = dot size, total dot surface area, total dot contour length and density negatively correlated with number while array size positively correlated with number.

Table A.2. *Nonaddition and continuous quantity strategy analyses.*

Nonaddition Strategy	Number of Trials	Mean Accuracy %	Wald χ2	df	*p* - value	Logit [a] values	> Chance (50%)
Blue	12	55.78	4.44	1	.035	0.23	yes
Rred	12	65.19	38.81	1	.000	0.63	yes
Near/Far	8	60.97	10.40	1	.001	0.29	yes
B1vsR	6	70.70	56.62	1	.000	0.88	yes
B2vsR	6	60.14	12.68	1	.000	0.39	yes
RedRange	14	59.90	30.92	1	.000	0.40	yes
Continuous Quantity							
A	12	47.04	0.81	1	.369	-0.12	no
B	12	73.92	108.96	1	.000	1.04	yes

[a] Computed with the formula: ln(p/1-p), where "p" stands for proportion

We also tested whether children based their response on features related to the presentation of the dots. We examined whether children responded based on combined variables of dot size, summed dot surface area, summed dot circumference, density and array total area. To control for these variables, our trials were presented in two conditions (see Table A1): in condition A, dot size, total dot surface area, total dot contour length and density were positively correlated with number while array size negatively correlated with number; condition B had the opposite relations. Table A.2 shows that strategy B was found to be significantly above chance level whereas strategy A was not. We cannot assume that children's performance can be accounted as relying on continuous quantity variables since they did not perform significantly below chance level (Gilmore et al., 2010).

ACKNOWLEDGEMENTS

The authors would like to thank all children, parents, teachers and schools that took part in this study. We would also like to thank: Petra Dekker for her great assistance in the data-acquisition procedure, Cor Stoof and Jarik den Hartog for their invaluable technical assistance and Dr. Ineke Imbo from Ghent University for her advice in the development of our dual task design. Furthermore, we would like to thank Prof. Dr. Alan Baddeley for his inspiring insights into our results and two anonymous reviewers. This research was conducted as part of a broader NWO (National Dutch Organization for Scientific Research; PROO 411 07 111) funded project.