

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/285322316>

Evidence for distinct magnitude systems for symbolic and non-symbolic number

ARTICLE *in* PSYCHOLOGICAL RESEARCH · DECEMBER 2015

Impact Factor: 2.47 · DOI: 10.1007/s00426-015-0734-1

READS

33

3 AUTHORS:



[Delphine Sasanguie](#)

University of Leuven

17 PUBLICATIONS 216 CITATIONS

[SEE PROFILE](#)



[Bert De Smedt](#)

University of Leuven

93 PUBLICATIONS 1,432 CITATIONS

[SEE PROFILE](#)



[Bert Reynvoet](#)

University of Leuven

72 PUBLICATIONS 2,084 CITATIONS

[SEE PROFILE](#)

Accepted in Psychological Research

Evidence for distinct magnitude systems for symbolic and non-symbolic number

Delphine Sasanguie^{1,2}, Bert De Smedt³ & Bert Reynvoet^{1,2*}

¹ Brain and Cognition, KU Leuven, 3000 Leuven, Belgium

² Faculty of Psychology and Educational Sciences @Kulak, 8500 Kortrijk, Belgium

³ Parenting and Special Education, KU Leuven, 3000 Leuven, Belgium

*** Corresponding author:**

Delphine Sasanguie, Ph.D.

Brain & Cognition, KU Leuven

Faculty of Psychology and Educational Sciences @Kulak

Etienne Sabbelaan 53; 8500 Kortrijk – Belgium

Tel.: +32 56/24.61.77

E-mail: Delphine.Sasanguie@kuleuven-kulak.be

Abstract

Cognitive models of magnitude representation are mostly based on the results of studies that use a magnitude comparison task. These studies show similar distance or ratio effects in symbolic (Arabic numerals) and non-symbolic (dot arrays) variants of the comparison task, suggesting a common abstract magnitude representation system for processing both symbolic and non-symbolic numerosities. Recently however, it has been questioned whether the comparison task really indexes a magnitude representation. Alternatively, it has been hypothesized that there might be different representations of magnitude: an exact representation for symbolic magnitudes and an approximate representation for non-symbolic numerosities. To address the question whether distinct magnitude systems exist, we used an audio-visual *matching* paradigm in two experiments to explore the relationship between symbolic and non-symbolic magnitude processing. In Experiment 1, participants had to match visually and auditory presented numerical stimuli in different formats (digits, number words, dot arrays, tone sequences). In Experiment 2, they were instructed only to match the stimuli after processing the magnitude first. The data of our experiments show different results for non-symbolic and symbolic number and are difficult to reconcile with the existence of one abstract magnitude representation. Rather, they suggest the existence of two different systems for processing magnitude, i.e. an exact symbolic system next to an approximate non-symbolic system.

Keywords: Magnitude processing, audio-visual matching paradigm, symbolic and non-symbolic magnitudes, exact symbolic magnitude system

Evidence for distinct magnitude systems for symbolic and non-symbolic number.

Cognitive models on magnitude representation are mostly based on the results of studies that use a magnitude comparison task. In a magnitude comparison task, participants need to indicate which of two presented numerical magnitudes is the larger. Typically, a numerical distance or ratio effect is observed: The larger the numerical distance or the ratio approaches to 1, the faster and more accurate participants respond (e.g., Barth, Kanwisher & Spelke, 2003; Defever, Reynvoet & Gebuis, 2013; Halberda & Feigenson, 2008; Sasanguie, Göbel, Moll, Smets & Reynvoet, 2012). Commonly, these effects are explained by assuming overlapping magnitude representations on a logarithmically compressed mental number line (Dehaene, 1997). On this line, magnitudes are represented by means of Gaussian curves, implying that a particular magnitude activates not only its corresponding representation, but also to a lesser extent the representations of numerically close magnitudes (Moyer & Landauer, 1967; Restle, 1970), making it more difficult to discriminate between such close magnitudes. Using this comparison task, previous behavioral studies (e.g. Holloway & Ansari, 2009; Mundy & Gilmore, 2009; Sasanguie, De Smedt, Defever & Reynvoet, 2012) revealed similar distance or ratio effects obtained with symbolic (e.g. digits) and non-symbolic stimuli (e.g. dot arrays). Moreover, neuroimaging studies have demonstrated that symbolic and non-symbolic comparison tasks both activate the same brain area (i.e. the Intraparietal Sulcus or IPS) and that both tasks display similar distance effects in the IPS (e.g. Eger, Sterzer, Russ, Giraud & Kleinschmidt, 2003; Holloway & Ansari, 2010; Notebaert, Pesenti & Reynvoet, 2010; Piazza, Pinel, Le Bihan & Dehaene, 2007). Together, these studies suggest one abstract magnitude representation system in the brain for processing both non- symbolic and symbolic magnitudes.

Alternatively, it has also been hypothesized that there might be notation-dependent representations of magnitude: one for processing symbolic magnitudes in an exact manner (i.e. only applying to a unique cardinality and thus no overlap between the representations of neighboring numbers) and one for approximate non-symbolic number processing (e.g. Carey, 2009; Le Corre & Carey, 2007; Noël & Rousselle, 2011; Sasanguie, Defever, Maertens & Reynvoet, 2014; Lyons, Ansari & Beilock, 2012). Indirect evidence for this idea has been provided by studies which observed a differential association between mathematics achievement and either non-symbolic or symbolic numerosity processing (for a review, see De Smedt, Noël, Gilmore & Ansari, 2013; for a meta-analysis, see Schneider et al., in press) and the absence of a correlation between distance effects in non-symbolic and symbolic comparison (Sasanguie et al., 2012, 2014). Moreover, there are neuroimaging data that support the idea of two different magnitude representation systems (e.g. Cohen Kadosh, Bahrami, Walsh, Butterworth, Popescu & Price, 2011; Cohen Kadosh, Cohen Kadosh, Kaas, Henik & Goebel, 2007). For instance, using Multivoxel Pattern Analysis (MVPA), Bulthé, De Smedt and Op de Beeck (2014) observed no neural overlap between digit and dot representations on three different spatial scales. However, direct behavioral evidence in favor of distinct magnitude systems for non-symbolic and symbolic number is still lacking.

The latter may be caused by the fact that many studies rely on magnitude comparison to investigate differences between non-symbolic and symbolic number. Recently however, it has been questioned whether the classic magnitude comparison task is a good task to index the representation of magnitude. First, it has been demonstrated that overlapping representations are not required to obtain a distance or ratio effect (e.g. Holloway & Ansari, 2008; Van Opstal, Gevers, De Moor & Verguts, 2008): Also in non-numerical comparison tasks, such as comparing navy ranks or academic positions, or comparing the brightness, length, area or duration of two stimuli, a distance effect has been observed (e.g. Chiao,

Bordeaux & Ambady, 2004; Cohen Kadosh, Brodsky, Levin & Henik, 2008; Holloway & Ansari, 2008; Van Opstal & Verguts, 2010). This suggests that distance and ratio effects obtained in comparison do not necessarily reflect magnitude representations. An alternative task used to address the magnitude representation is the numerical matching task. In this task, participants have to decide whether two magnitudes are numerically similar or not. Computational and behavioral data have shown that a numerical matching task is more appropriate to investigate the mental representations of magnitudes (Cohen Kadosh et al., 2008; Van Opstal & Verguts, 2011). Second, it has been suggested that the processing of non-symbolic and symbolic number is confounded by visual processing. In a non-symbolic discrimination task, participants have to discriminate two visual presented dot arrays. Several studies have shown that numerosity judgments measured by this task are sensitive to the continuous visual properties of the dot arrays, such as density (Gebuis & Reynvoet, 2012a, 2012b; Sophian & Chu, 2008; Tibber, Greenwood & Dakin, 2012), the size of the individual elements (Gebuis & Reynvoet, 2012a, 2012b; Hurewitz, Gelman & Schnitzer, 2006; Rousselle & Noël, 2008; Tokita & Ishiguchi, 2010), or patch size (Tokita & Ishiguchi, 2010). Also symbolic number discrimination may be confounded by physical – and not numerical - similarity. For instance, in a numerical matching task, Cohen (2009) found that participants' reaction times were a function of physical similarity of the visually presented digits (see also Defever, Sasanguie, Vandewaetere & Reynvoet, 2011). One alternative to circumvent visual similarity is to use audio-visual stimuli pairs (e.g. a visual presented Arabic numeral together with an auditory presented number word). This method has been successfully applied in previous numerical studies with children and adults by, for instance, Barth and colleagues (e.g., Barth, Kanwisher & Spelke, 2003; Barth, La Mont, Lipton & Spelke, 2005; 2006). However, in these studies a comparison instruction was used and we already explained that such an instruction does not guarantee that magnitude representations are activated. More

recently, Holloway, van Atteveldt, Blomert and Ansari (2013) used a numerical audio-visual matching paradigm, but only as a control condition for symbol processing because numerosity processing was not the main research focus of their study.

To circumvent these two problems that may have obscured differences between non-symbolic and symbolic number processing, we used a *cross-modal (audio-visual) matching paradigm* in the current study (see Figure 1). In Experiment 1, adult participants performed four numerical tasks in which they had to match a visually presented and an auditory presented stimulus. Visual stimuli were either Arabic numerals (e.g. 1) or dot arrays (e.g. ●●●); auditory stimuli were verbal number words (e.g. /five/) or tone sequences (e.g. /beep beep beep/). Moreover, they performed two control tasks: a color word-color matching task, as a control measure for their general matching ability, and a letter speech sound-letter matching task to check whether the observed findings were number-specific or could be generalized to other symbols than numbers. We hypothesized that, if symbolic and non-symbolic magnitudes are indeed processed by one abstract magnitude representation system, 1) all tasks should show a ratio effect because this effect is the signature of the activation of the magnitude representation system (Price & Ansari, 2013) and 2) performance on all tasks should be correlated with each other. Alternatively, if symbolic numerosities are represented exactly and different from non-symbolic numerosities, 1) there should be no ratio effect present in the symbolic tasks and 2) the performance on the symbolic and the non-symbolic tasks should not be correlated.

[INSERT FIGURE 1 ABOUT HERE]

Experiment 1

Method

Participants.

Thirty-seven university students from the KU Leuven Kulak participated for eight Euros. Three participants were removed from the analyses because they performed too slow or made too many errors ($>3SD$ above the group mean) in one of the experimental tasks. Consequently the final sample consisted of 34 adults ($M_{\text{age}} = 21.59$ years; $SD = 3.63$; 26 females).

Procedure, tasks and stimuli.

Each participant first conducted the word-color matching task (i.e. general matching ability measure) followed by the numerical and the letter (control) audio-visual matching tasks in a fully counterbalanced way. Numerosities between 1 and 9 were used. Because it remains debated (for an overview, see Hyde, 2011) whether numbers from within and above the so-called subitizing range (i.e. small number of items up to 4 which can be readily identified; Trick & Pylyshyn, 1994) are processed similarly (e.g. Cantlon & Brannon, 2006; Cordes, Gelman, Gallistel & Whalen, 2001) or differently (e.g., Feigenson, Dehaene & Spelke, 2004; Hyde & Spelke, 2009), ratios (i.e. relative distances) instead of absolute distances were used. This way, similar ratios could be analyzed for both number ranges: pairs 2-4 and 5-9 (small ratios 0.50 and 0.56) versus pairs 3-4 and 7-9 (large ratios 0.75 and 0.78).

Control audio-visual matching tasks

Word-color matching task. Stimulus presentation and recording of the data were controlled by E-prime 1.1 (Psychology Software Tools, <http://www.pstnet.com>). Participants heard a color word (e.g. /green/) – in Dutch. Afterwards, a colored (i.e. green, blue, red or white) square was presented in the center of the computer screen and participants were asked to indicate whether the color they saw on the screen matched or not with the color word they heard by pressing the corresponding keys ('a' for match and 'p' for non-match) on an AZERTY keyboard. The sounds were digitally recorded (sampling rate 44.1 kHz, 16-bit quantization) from a Dutch female speaker. Recordings were band-pass filtered (180 –

10.000 Hz), resampled at 22.05 kHz, and matched for loudness. The sounds were presented binaurally through loudspeakers at about 65 dB SPL. They were presented such that each recorded word lasted about 1000ms. Also the visual stimuli were presented for 1000 ms. Each trial was preceded by a fixation cross of 600 ms. Participants could respond during the visual stimulus presentation or during a blank screen that followed the stimulus presentation. The next trial started after an inter-trial interval of 1500 ms. Before the experiment started, 10 practice trials were given during which feedback was provided in order to make the participants familiar with the task demands. There were 60 experimental trials. Participants were seated at approximately 50 cm from the screen.

Sound-letter matching task. The procedure was identical as in the word-color matching task, except that the stimuli were auditory presented speech sounds of letter names (e.g. /a/) and visually presented letters from the alphabet (i.e. a, b, e or f, in white on a black background, Arial font 40). There was an equal number of trials per condition: four ‘match’ trials (a-a, b-b, e-e, f-f), four ‘non-match’ trials with a small distance of one (a-b, b-a, e-f and f-e) and four ‘non-match’ trials with a large distance of four letter in between in the alphabet (a-e, e-a, b-f and f-b), resulting in 12 trials which were randomly repeated five times.

Numerical audio-visual matching tasks

Number word-digit matching task. Stimuli were auditory presented number words (e.g., /one/) and visually presented digits (i.e. 2, 3, 4, 5, 7 and 9). Identical as in the word-color matching task, the sounds were presented such that each recorded word lasted about 1000ms. Two groups of ratios were used on a 0-1 scale: Small ratios (0.50 – 0.56) and large ratios (0.75 – 0.78), both with numbers within and outside the subitizing range (2-4 and 3-4 vs. 5-9 and 7-9). Each number from these pairs was presented auditory (as a number word) as well as visually (as an Arabic numeral). This way, a trial list of eight pairs was obtained, which was randomly presented five times (= 40 experimental trials). After adding 20 filler

trials containing two numbers that were numerically the same (e.g. /five/ - 5), the final trial list consisted of 60 trials, preceded by 10 practice trials.

Number word-dots matching task. Here, stimuli were auditory presented number words (e.g. /nine/ - duration set to about 1000ms) in combination with visually presented dot arrays containing one, two, three, four, five, seven or nine dots. Stimuli were generated with the Matlab script developed by Gebuis and Reynvoet (2011). The trial list again consisted of 60 trials: two pairs of a small ratio within the subitizing range (2-4/4-2) and two outside the subitizing range (5-9/9-5), in combination with two pairs of a large ratio within (3-4/4-3) and two outside the subitizing range (7-9/9-7) – all presented randomly five times with each time a different dot configuration – and an additional 20 filler trials in which the number word and the dot array had the same numerical value. Ten practice trials, randomly selected from the list, preceded the experiment.

Tones-digits matching task. Stimuli were auditory presented tone sequences containing one, two, three, four, five, seven or nine tones (e.g. /beep beep beep/). The tone sequences lasted for 1000 ms, independent of the number of tones included in the sequence. A hissing noise was presented for 1000 ms to draw the participants' attention to the whole length of the stimulus sequence. High and low pitches were randomly interspersed and the inter-tone intervals randomly varied. After the tone sequence, an Arabic numeral appeared on the screen and participants were asked whether the tone sequence and the digits numerically matched or not. Similar to the other numerical tasks, here the trial list also consisted of 10 practice trials and 60 experimental trials.

Tones-dots matching task. The auditory presented stimuli were identical as in the tones-digits matching task (i.e. tone sequences) and the visual presented stimuli were identical as in the number word-dots matching task (i.e. dot arrays containing 2, 3, 4, 5, 7 or 9 dots). For each number, several tone sequences and dot configurations were created. After 10 practice trials,

participants fulfilled 60 experimental trials: 20 of a small ratio, 20 of a large ratio and 20 filler trials (i.e. ratio 0).

Results

In some conditions of the numerical matching tasks, accuracy rates were moderate to low (see Table 1, e.g. performance on the tones-digit matching and the tones-dots matching task with large numerosities). We therefore used accuracy data rather than reaction times as our main measure to index performance. The reaction time data, however, showed exactly the same pattern of results and these analyses are included in Appendix 1.

[INSERT TABLE 1 ABOUT HERE]

Control audio-visual matching tasks.

Word-color matching task. Participants scored 98% correct ($SD = 3\%$) on this task.

Sound-letter matching task. Participants scored 95% ($SD = 6\%$), 98% ($SD = 3\%$) and 98% ($SD = 3\%$) correct on trials with the same letters ('match trials'), a distance of 1 and a distance of 4 letters in between in the alphabet ('non-match' trials), respectively.

Mean accuracy rates on the 'non-match' trials were submitted to a repeated measures analysis of variance (ANOVA) with distance (2 levels: 1 and 4) as within-subject variable. There was no main effect of distance, $F < 1$.

Numerical audio-visual matching tasks.

The mean accuracies of the numerical audio-visual matching tasks are shown per ratio and separately for the numbers within and outside the subitizing range in Table 1. The mean accuracy rates on the 'non-match' trials were submitted to a repeated measures analysis of variance (ANOVA) with task (4 levels: number word-digit matching, number word-dots matching, digits-tones matching and tones-dots matching), ratio (2 levels: small vs large) and number range (2 levels: within vs outside subitizing range) as within-subject variables.

All main effects were significant: There was a main effect of task, $F(3,31) = 63.064$, $p < .0001$, $\eta_p^2 = .859$, a main effect of ratio, $F(1,33) = 249.136$, $p < .0001$, $\eta_p^2 = .883$, and a main effect of number range, $F(1,33) = 87.150$, $p < .0001$, $\eta_p^2 = .725$. Moreover, there was a significant interaction between task and ratio, $F(3,31) = 65.211$, $p < .0001$, $\eta_p^2 = .863$, between task and number range, $F(3,31) = 15.118$, $p < .0001$, $\eta_p^2 = .594$, and between ratio and number range, $F(1,33) = 11.055$, $p = .002$, $\eta_p^2 = .251$, which were in turn embedded in a three-way interaction between task, ratio and number range, $F(3,31) = 4.295$, $p = .012$, $\eta_p^2 = .294$.

Because our scope of interest is on symbolic versus non-symbolic numbers, we disentangled these interactions by conducting separate analyses per task (and post-hoc analyses per number range when relevant).

Number word-digit matching task. The effect of ratio was not significant, $F < 1$, there was no effect of number range, $F(1,33) = 3.989$, $p = .054$, $\eta_p^2 = .108$, and no ratio \times number range interaction, $F < 1$.

Number word-dots matching task. There was a main effect of ratio, $F(1,33) = 18.774$, $p < .0001$, $\eta_p^2 = .363$, demonstrating higher accuracies on the small ratio's than on the large ratio's. Moreover, a main effect of number range was present, $F(1,33) = 19.781$, $p < .0001$, $\eta_p^2 = .375$, indicating more errors in the trials with the largest numbers (i.e. outside the subitizing range, 7-9). There was also an interaction between ratio and number range, $F(1,33) = 15.042$, $p < .0001$, $\eta_p^2 = .313$. Post-hoc paired t -tests showed that there was no ratio effect in the subitizing range, $t < 1$, and there was a significant ratio effect outside the subitizing range, $t(33) = 4.374$, $p < .0001$.

Tones-digit matching task. There was a main effect of ratio, $F(1,33) = 198.619$, $p < .0001$, $\eta_p^2 = .858$, demonstrating higher accuracies on the small ratio's than on the large ratio's. Moreover, a main effect of number range was observed, $F(1,33) = 20.86$, $p < .0001$,

$\eta_p^2 = .387$, indicating less errors on the trials with number from within the subitizing range.

There was no interaction between ratio and number range, $F < 1$.

Tones-dots matching task. There was a main effect of ratio, $F(1,33) = 139.591$, $p < .0001$, $\eta_p^2 = .809$, demonstrating higher accuracies on the small ratio's than on the large ratio's. There was also a main effect of number range, $F(1,33) = 60.670$, $p < .0001$, $\eta_p^2 = .648$, indicating more errors on the larger trials. Moreover, there was an interaction between ratio and number range, $F(1,33) = 5.074$, $p = .031$, $\eta_p^2 = .133$, demonstrating a slightly larger ratio effect in the larger number range. Post-hoc paired t -tests showed however significant ratio effects in both number ranges: $t(33) = 8.902$, $p < .0001$, and $t(33) = 8.238$, $p < .0001$, for within and outside the subitizing range respectively.

Numerical ratio effects.

To examine the ratio effect into more detail, we computed the size of the ratio effect for each participant by subtracting the average accuracy on trials with the most difficult ratios (i.e. 0.75 and 0.78) from the average accuracy on trials with the easiest ratios (i.e. 0.50 and 0.56) (see Holloway & Ansari, 2009; Sasanguie et al., 2013 for a similar method). We then compared the size of the ratio effect in the number word-digit matching task with the weakest ratio effect in one of the other three numerical tasks, to statistically test the absence of a ratio effect in the number word-digit task. A paired t -test showed that both ratio effects significantly differed from each other, $t(33) = 4.074$, $p < .0001$.

Correlation analysis.

Partial correlations, controlling for general matching ability measured by the performance on the word-color matching task, were conducted between the mean accuracies of the numerical audio-visual matching tasks and the control sound-letter matching task, in order to examine whether these tasks shared a common underlying mechanism.

No significant correlations were observed between the number word-digit matching task and one of the other numerical matching tasks, all $r(31) < .159$, $p > .356$. In contrast, a significant correlation was present between the sound-letter matching task and the number word-digit matching task, $r(31) = .435$, $p = .011$. All other numerical matching tasks (number word-dots, tones-digit and tones-dots) also correlated with each other (number word-dots and tones-dots, $r(31) = .353$, $p = .044$; tones-digit and tones-dots, $r(31) = .356$, $p = .042$ and number word-dots and tones-digit, $r(31) = .397$, $p = .022$).

Discussion

The results of Experiment 1 revealed a ratio effect in the three numerical audio-visual matching tasks containing at least one non-symbolic stimulus (i.e. the two mixed notation conditions and the pure non-symbolic condition). A ratio effect is assumed to be the result of overlapping Gaussian curves on the mental number line and therefore considered as a signature of the underlying magnitude representation (Dehaene, 1997; Price & Ansari, 2013). The ratio effect interacted with number range when visual dot arrays were presented (i.e. the number word-dots matching task and the tones-dots matching task). In these conditions, only in case of the larger numerosities, a clear ratio effect was present (or the ratio effect was larger than the ratio effect in the smaller number range). This might be so because the small number of dots have been ‘subitized’ by the participants (i.e. a process in which up to 4 items can readily be identified because they are represented exactly; Trick & Pylyshyn, 1994). The absence of a ratio effect in the pure symbolic condition (i.e. the number word-digit matching task) suggests that when processing digits, another (magnitude) representation is activated. The performance in the digit task was similar to the performance in the sound-

letter matching task, where also no ratio effect was present, indicating that similar processes are involved in processing both symbols (i.e. digits and letters)¹.

Correlations confirmed this dichotomy between conditions. The participants' performance on the number word-digit and the sound-letter matching task were correlated. Also their performance on the numerical tasks with non-symbolic stimuli (i.e. either dot arrays or tones) were associated with each other, but none of the performance on these tasks correlated with the performance on a pure symbolic task (i.e. the number word-digit and the sound-letter matching task). This analysis thus clearly revealed a distinction between the symbolic tasks on the one hand and the non-symbolic tasks on the other. Together, these results are difficult to reconcile with the idea of one abstract magnitude representation (e.g. Eger et al., 2003; Holloway & Ansari, 2009; 2010; Notebaert et al., 2010) and rather suggest that symbolic and non-symbolic numbers are processed differently.

Two possible explanations can be put forward for our observations: First, it is possible that the two pure symbolic tasks were solved using an *asemantic route* (i.e. without access to magnitude information). Indeed, the current results could be interpreted in terms of the Triple Code Model of Dehaene and Cohen (1995), to date still the most dominant cognitive model on numerical cognition. This theory postulates three main number representations – a visual identification code in which numbers are represented as symbols (e.g. '6'), a magnitude code in which numbers are represented as analogue magnitudes (e.g.

¹ Because Experiment 1 also contains two conditions/tasks in which participants have to match a non-symbolic and a symbolic number, the results regarding the ratio effects may also be informative for (a)symmetry of mapping between non-symbolic and symbolic number, an issue addressed previously by e.g. Mundy and Gilmore (2009). These authors showed smaller REs for non-symbolic-to-symbolic mapping than for symbolic-to-non-symbolic mapping in children. In contrast, in our study, accuracy results demonstrated a larger RE in the tones-digit matching task (i.e. the condition in which the non-symbolic stimulus was presented first) compared to the number word-dots matching task (symbolic stimulus presented first), $t(33)= 7.440$, $p < .0001$. However, our results are inconclusive regarding the issue of asymmetry in mapping because both conditions not only differ regarding the firstly presented number (symbolic versus non-symbolic), but also the modality of the non-symbolic number is different. In order to ensure that the differences between our two mixed conditions are due to the direction of mapping, it should be examined whether visual dots and auditory tones are represented equally precise by the analogue magnitude system. Future research should address this issue.

‘●●●●●●’); and a verbal code in which numbers are represented as number words (e.g. ‘six’). In terms of this model, the current results suggest that adults’ performance on the numerical audio-visual matching tasks with non-symbolic stimuli relies on the analogue magnitude code, whereas their performance on the number word-digit and the sound-letter matching task rather reflects the direct link between the visual identification code and the verbal code – without activating the analogue magnitude code.

Alternatively, symbolic numbers may activate different magnitude representations than non-symbolic numbers. This is in line with previous behavioral and neurocognitive studies showing differences between the processing of different numerical formats (e.g., Bulthé et al., 2014; Butterworth, 2010; Cohen Kadosh et al., 2011; Holloway & Ansari, 2009; Rousselle & Noël, 2011; Sasanguie et al., 2014). Possibly, using a cross-modal matching paradigm, the current experiment might therefore also point towards the existence of an *exact symbolic magnitude representation*, which was activated in these symbolic tasks and exists next to the approximate, analogue magnitude representation that is activated by non-symbolic numbers.

Based on Experiment 1, we are not able to fully dissociate between the two possible explanations for our results (i.e. ‘the asemantic route’ versus ‘the exact symbolic magnitude representation’), because we are not sure that the participants, when performing the symbolic audio-visual matching task, activated the numerical meaning - i.e. magnitude - of the symbols). In order to ascertain that the participants processed the magnitudes of the numbers, we conducted Experiment 2 in which we used the same symbolic number word-digit matching task (and a control non-symbolic tones-dots matching task) in a go/no-go experimental paradigm. Specifically, we additionally instructed the participants to perform the numerical matching only if both numbers were either smaller/larger than or equal to a specific magnitude. If this was not the case, they were instructed to withhold their response. This

way, we were sure that participants processed the magnitudes of the numbers prior to the matching and excluded an alternative asemantic solution strategy for symbolic numbers. If the dissociation between the performance in non-symbolic and symbolic conditions is replicated in Experiment 2, this would provide strong evidence for the existence of a separate exact symbolic magnitude representation.

Experiment 2

Method

Participants.

Twenty-four adult volunteers participated in this experiment. One participant was removed from the analyses because he/she performed below chance level (50%) in one of the experimental tasks. Consequently the final sample existed of 23 adults ($M_{\text{age}}= 25.48$ years; $SD= 2.56$; 17 females).

Procedure, tasks and stimuli.

The number word-digit matching task and the tones-dots matching task were identical as in Experiment 1, except for the number of trials that was doubled in two identical Blocks and the participants who were now additionally instructed only to match (= 'go') the auditory and the visual stimulus if both numbers were either smaller than or equal to 4 (Block 1 – 50% of the trials) or were larger than or equal to 5 (Block 2 – 50% of the trials). If else, they were instructed not to answer on the trial (= 'no go') and wait until the next trial appeared on the screen. In the latter case, the inter-trial interval was 2000 ms. The order of the tasks and the order of the Blocks within the tasks were counterbalanced over all participants.

Results

In the number word-digit matching task, 0.5% of the trials were false alarms (i.e. 'no go'- trials on which the subjects answered anyway) and 0.3% were misses (i.e. 'go'-trials on

which the subjects did not respond); in the tones-dots matching task, there were 0.3% false alarms and no misses.

For both tasks, only the performance on the ‘go trials’ was included in the analyses. In Table 2, the performance on the ‘go’- trials from the two tasks is presented per ratio and for the numbers within and outside the subitizing range.

[INSERT TABLE 2 ABOUT HERE]

Similar as in Experiment 1, the mean accuracy rates on the ‘non-match’- trials were submitted to a repeated measures analysis of variance (ANOVA) with task (2 levels: number word-digit matching and tones-dots matching), ratio (2 levels: small vs large) and number range (2 levels: within vs outside subitizing range) as within-subject variables. However, unlike Experiment 1, we here also immediately report the identical analyses conducted on the median reaction time data, because those do not show a similar pattern as the accuracy data and need some additional clarification. One should be cautious however when interpreting reaction time data of conditions in which the accuracies were close to chance (e.g. in the tones-dots matching task).

In the accuracy data, there was a main effect of task, $F(1,22) = 66.634, p < .0001, \eta_p^2 = .752$. There was also a main effect of ratio, $F(1,22) = 26.369, p < .0001, \eta_p^2 = .545$ and a main effect of number range, $F(1,22) = 22.224, p < .0001, \eta_p^2 = .503$. Furthermore, there was a task \times ratio interaction, $F(1,22) = 26.960, p < .0001, \eta_p^2 = .551$ and a task \times number range interaction, $F(1,22) = 32.280, p < .0001, \eta_p^2 = .595$.

In the reaction time data, there was a main effect of task, $F(1,22) = 103.973, p < .0001, \eta_p^2 = .825$ and a main effect of number range, $F(1,22) = 8.835, p = .007, \eta_p^2 = .287$. There was also an interaction between task and ratio, $F(1,22) = 9.555, p = .005, \eta_p^2 = .303$ and between ratio and number range, $F(1,22) = 24.710, p < .0001, \eta_p^2 = .529$, which were embedded in a task \times ratio \times number range three-way interaction, $F(1,22) = 9.363, p = .006$,

$\eta_p^2 = .299$. To disentangle these interactions, separate analyses per task were conducted (and post-hoc analyses per number range when relevant).

Number word-digit matching task.

Accuracies. There was no effect of ratio, $F(1,22) = 1.206, p = .284, \eta_p^2 = .052$, no effect of number range, $F < 1$, and no interaction between ratio and number range present, $F < 1$.

Reaction times. There was a significant main – though reversed – effect of ratio, $F(1,22) = 4.337, p = .049, \eta_p^2 = .165$, demonstrating faster responses on the large ratio's than on the small ratio's. Moreover, a main effect of number range was present, $F(1,22) = 8.347, p = .009, \eta_p^2 = .275$, indicating faster responses on the trials with the largest numbers. The interaction between ratio and number range was not significant, $F(1,22) = 3.304, p = .083, \eta_p^2 = .131$.

Tones-dots matching task.

Accuracies. There was a main effect of ratio, $F(1,22) = 30.008, p < .0001, \eta_p^2 = .577$, demonstrating higher accuracies on the small ratio's than on the large ratio's. There was also a main effect of number range, $F(1,22) = 30.826, p < .0001, \eta_p^2 = .584$, indicating more errors on the larger trials. The interaction between ratio and number range was not significant, $F(1,22) = 1.080, p = .310, \eta_p^2 = .047$.

Reaction times. There effect of ratio was not significant, $F(1,22) = 4.092, p = .055, \eta_p^2 = .157$. There was also no effect of number range, $F(1,22) = 3.765, p = .065, \eta_p^2 = .146$. There was a significant interaction between ratio and number range, $F(1,22) = 24.011, p < .0001, \eta_p^2 = .522$. Post-hoc paired t -tests showed that only in case of the smallest numbers, there was a significant ratio effect, $t(22) = 5.728, p .0001$. Most probably, the absence of a ratio effect in the largest numbers is because the 7-9 trials were too difficult and participants answered fast without being accurate (cf. only 53% correct).

Numerical ratio effects.

Similar as in Experiment 1, we computed the size of the ratio effect for each participant by subtracting the average accuracy (or reaction time) on trials with the most difficult ratios (i.e. 0.75 and 0.78) from the average accuracy (or reaction time) on trials with the easiest ratios (i.e. 0.50 and 0.56). In case of the reaction times, this difference score was also divided by the average reaction time on trials with the easiest ratios to correct for individual differences in reaction times (see Holloway & Ansari, 2009; Sasanguie et al., 2013 for a similar method). Paired *t*-tests showed that both ratio effects significantly differed from each other, $t(22) = 5.192$, $p < .0001$, for the accuracies and $t(22) = 2.902$, $p = .008$, for the reaction times.

Discussion

The absence of a ratio effect in the pure symbolic task of Experiment 1 could be explained by the direct link between the visual identification code and the verbal code – without activating the analogue magnitude code (i.e. an asemantic route). Alternatively, this result could be explained by a distinct, exact magnitude representation for symbolic numbers. In order to dissociate between these two alternative explanations, Experiment 2 was conducted. By means of a go/ no-go experimental design, the participants were obliged to process the numerical meaning (i.e. the magnitude) of the numbers before making an audio-visual matching decision. We reasoned that, if the participants would have processed the magnitude and there would still be no ratio effect present in the number word-digit matching task, this would provide us with clear evidence in favor of the existence of an exact symbolic magnitude representation.

In contrast to the non-symbolic task, the accuracy results of the symbolic matching task in Experiment 2 showed indeed no ratio effect. Together with the fact that the participants had processed the magnitude of the symbolic numerosities, this demonstrates

that human adults can process symbolic magnitudes without activating an abstract, analogue magnitude representation. This finding cannot be explained without assuming the existence of a distinct, exact symbolic magnitude representation, next to the approximate, analogue non-symbolic magnitude representation.

It must be noted, however, that unlike Experiment 1, the reaction time data of Experiment 2 did not entirely mirror the accuracy results. Also in the symbolic task there was an effect of ratio in the reaction time data. Closer inspection of these RT data, however, showed that – despite the absence of a significant ratio \times number range interaction – this ratio effect was only significant in case of the largest numbers (i.e. trials 5-9 and 7-9). Moreover, in contrast to a classic ratio effect (i.e. faster responses on the easy ratio), here the effect of ratio was reversed, $t(22) = 2.196, p = .039$: participants were slower on the 5-9 trials (easy ratio 0.56) than on the 7-9 trials (difficult ratio 0.78). Because the latter finding was only observed in Experiment 2 (– in Experiment 1, there was no ratio effect and no ratio by number range interaction in the RT data of this task, both $F_s < 1$), the only possible explanation is that this reversed ratio effect is evoked by the instructions within the larger number range. In that range, the participants had to decide whether both numbers were ‘larger than or equal to 5’ and were slowed down in case of the 5-9 trials, because the number 5 served as a reference, in contrast to the 7-9 trials, where their decision was not confounded by the reference 5. Consequently, in all probability, this might have resulted in a classic comparison *distance* effect (Moyer & Landauer, 1967), leading to faster responses on the larger distances (i.e. faster on 7-9 trials than on 5-9 trials because the former are further away from five). This reasoning is highly plausible because this problem was only observed in the larger number range. In the smaller number range, the instructions were ‘smaller than or equal to 4’, but both trials contained 4 as one of the to be matched numbers, i.e. 2-4 and 3-4. Here, both types of trials thus were ‘affected’ by the confound/slow down process of the reference. In the smaller number range, we did not observe a

ratio effect. Therefore, we put forward this only possible explanation for the presence of the reversed ratio effect in the large number range of the number word-digit matching task.

In sum, our results suggest different magnitude representations for symbolic and non-symbolic magnitudes and we can conclude that to perform the task in Experiment 2, our participants relied on a separate, exact symbolic magnitude representation.

Conclusion

In the current study, we examined whether human adults indeed possess a common abstract magnitude representation for processing symbolic and non-symbolic magnitudes (e.g. Holloway & Ansari, 2010; Notebaert et al., 2010; Piazza et al., 2007), using an audio-visual matching paradigm in two experiments. By means of this particular paradigm, we could circumvent two issues raised in the field of numerical cognition, based on previous studies: First, we used *matching* instead of comparison because this task is more appropriate to investigate magnitude representations (Cohen Kadosh et al., 2008; Van Opstal & Verguts, 2011). In contrast, in order to obtain a ratio effect in a comparison task, overlapping representations are not required. Second, by using *cross-modality*, we could avoid that non-symbolic and symbolic number discrimination is contaminated by physical (non-numerical) resemblance.

In Experiment 1, we could observe that all tasks with non-symbolic stimuli, either dot arrays or tone sequences, resulted in a ratio effect, indicating that an approximate, analogue magnitude representation was addressed in these tasks. In contrast, in the number word-digit or pure symbolic task, no ratio effect was present. Furthermore, the dichotomy between the symbolic and non-symbolic conditions was confirmed by the correlation analysis where the performance in the symbolic condition was related to the performance in letter matching, but not to the other non-symbolic numerical matching tasks. In Experiment 2, we found that,

even when the participants were obliged to process the numerical meaning (i.e. the magnitude) of the symbols, still no ratio effect was present in the pure symbolic task.

The present data provide evidence for *distinct magnitude systems* for symbolic and non-symbolic number. In addition, the absence of a ratio effect in the symbolic condition also suggests that, in contrast to the approximate representation for non-symbolic number, symbolic numbers are represented *exactly*. If stimuli are represented exactly (i.e. magnitude representations of numbers do not overlap with nearby magnitudes), no ratio effect is expected in a numerical matching task because nearby and far number pairs are both equally discriminable. Although many studies, almost all with comparison instructions, consistently reported numerical distance effects in symbolic processing, the absence of distance effects has been reported before in matching tasks. Both Cohen (2009) and Goldfarb, Henik, Rubinsten, Bloch-David & Gertner (2011) found no numerical distance effect when participants had to numerical match two visually presented digits. These findings are in line with our current results and the suggestion of an exact symbolic magnitude system, but apparently this pattern of results only emerges under matching instructions. We are aware that our assumption of an exact magnitude system for symbolic number is based on the absence of an effect and therefore are convinced that further research is needed to investigate the characteristics of the symbolic magnitude system. However, what our findings are clear on is that there are distinct magnitude representations for non-symbolic and symbolic numbers.

Acknowledgments

Delphine Sasanguie is a postdoctoral research fellow for the FWO (Research Foundation Flanders). This research was also supported by the Research Fund KU Leuven.

References

- Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*, 201–221.
- 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.
- Bulthé, J., De Smedt, B., & Op de Beeck, H.P. (2014). Format-dependent representations of symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel pattern analyses, *NeuroImage*, *87*, 311–322.
- Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences*, *14*, 534–541.
- Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*, 401–406.
- Carey, S. (2009). *The origin of concepts*. Oxford: Oxford University Press.
- Chiao, J.Y., Bordeaux, A.R., & Ambady, N. (2004). Mental representations of social status. *Cognition*, *93*, 49-57.
- Cohen, D. (2009). Integers do not automatically activate their magnitude representation. *Psychonomic Bulletin & Review*, *16*, 332-336.
- Cohen Kadosh, R., Bahrami, B., Walsh, V., Butterworth, B., Popescu, T., & Price, C. (2011). Specialization in the human brain: the case of numbers. *Frontiers in Human Neuroscience*, *5*, 1–9.
- Cohen Kadosh, R., Brodsky, W., Levin, M., & Henik, A. (2008). Mental representation: what can pitch tell us about the distance effect? *Cortex*, *44*, 470–477.

- Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and independent representations of numbers in the parietal lobes. *Neuron*, 53, 307–314.
- Cordes, S, Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review*, 8, 698–707.
- Dehaene, S. (1997). *The number sense*. Oxford: Oxford University Press.
- Dehaene, S, & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition*, 1, 83–120.
- Defever, E., Reynvoet, B., & Gebuis, T. (2013). Task and age dependent effects of visual stimulus properties on children’s explicit numerosity judgments. *Journal of Experimental Child Psychology*, 116, 216-233.
- Defever, E., Sasanguie, D., Vandewaetere, M., Reynvoet, B. (2012). What can the same-different task tell us about the development of magnitude representations?. *Acta Psychologica*, 140 (1), 35-42
- De Smedt, B., Noël, M.-P., Gilmore, C. & Ansari, D. (2013). The relationship between symbolic and non-symbolic numerical magnitude processing and the typical and atypical development of mathematics: a review of evidence from brain and behavior. *Trends in Neuroscience and Education*, 2, 48–55.
- Eger, E., Sterzer, P., Russ, M.O., Giraud, A.-L., Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron*, 37, 719–725.
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, 8, 307–314.
- Gebuis, T., & Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behavior Research Methods*, 43, 981–986.

- Gebuis, T., Reynvoet, B. (2012a). The interplay between visual cues and non-symbolic number and its continuous visual properties. *Journal of Experimental Psychology. General*, *141* (4), 642-648.
- Gebuis, T., Reynvoet, B. (2012b). The role of visual information in numerosity estimation. *PLoS One*, *7* (5), art.nr. e37426, (epub).
- Goldfarb, L., Henik, A., Rubinsten, O., Bloch-David, Y., & Gertner, L. (2011). The numerical distance effect is task dependent. *Memory & Cognition*, *39*, 1508-1517.
- Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the “Number Sense”: The Approximate Number System in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*, 1457–1465.
- Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children’s development of number comparison. *Developmental Science*, *11*, 644-649.
- 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.
- Holloway, I.D., & Ansari, D. (2010). Developmental specialization in the right intraparietal sulcus for the abstract representation of numerical magnitude. *Journal of Cognitive Neuroscience*, *22*, 2627–2637.
- Holloway, I.D., van Atteveldt, N., Blomert, L., & Ansari, D. (2013). Orthographic Dependency in the Neural Correlates of Reading: Evidence from Audiovisual Integration in English Readers. *Cerebral Cortex*, Dec. 17 [Epub ahead of print], doi: doi:10.1093/cercor/bht347.
- Hurewitz, F., Gelman, R., & Schnitzer, B. (2006). Sometimes area counts more than number. *Proceedings of the National Academy of Sciences of the United States of America*, *103*, 19599–19604.

Hyde, D.C. (2011). Two systems of non-symbolic numerical cognition. *Frontiers in Human Neuroscience*, 5:150, doi: 10.3389/fnhum.2011.00150.

Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: an electrophysiological investigation of small and large number representations. *Journal of Cognitive Neuroscience*, 21, 1039–1053.

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, 395–438.

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, 635-641.

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

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

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*, 5:165, 1-4.

Notebaert, K., Pesenti, M., & Reynvoet, B. (2010). The neural origin of the priming distance effect: distance-dependent recovery of parietal activation using symbolic magnitudes. *Human Brain Mapping*, 31, 669–677.

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, 293-305.

Price, G., & Ansari, D. (2013). Dyscalculia: Characteristics, Causes, and Treatments.

Numeracy, 6, 2, doi: <http://dx.doi.org/10.5038/1936-4660.6.1.2>

- Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, 91, 191-205.
- Rousselle, L., & Noël, M.-P. (2008). The development of automatic numerosity processing in preschoolers: evidence for numerosity-perceptual interference. *Developmental Psychology*, 44, 544–560.
- Sasanguie, D. De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, 30, 344-357.
- Sasanguie, D., Defever, E., Maertens, B., Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergartners. *The Quarterly Journal of Experimental Psychology*, 67, 271-280.
- Sasanguie, D., Göbel, S., Moll, K., Smets, K., & Reynvoet, B. (2013). Acuity of the approximate number sense, symbolic number comparison or mapping numbers onto space: what underlies mathematics achievement? *Journal of Experimental Child Psychology*, 114, 418-431.
- Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S.S., Stricker, J., & De Smedt, B. (in press). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*.
- Sophian, C., & Chu, Y. (2008). How do people apprehend large numerosities? *Cognition*, 107, 460–478.
- Tibber, M. S., Greenwood, J. A., & Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *Journal of Vision*, 12, 8.

- Tokita, M., & Ishiguchi, A. (2010). How might the discrepancy in the effects of perceptual variables on numerosity judgment be reconciled? *Attention, Perception & Psychophysics*, *72*, 1839–1853.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, *101*, 80–102.
- Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: priming and comparison distance effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*, 419-425.
- Van Opstal, F., & Verguts, T. (2011). The origins of the numerical distance effect: The same– different task. *Journal of Cognitive Psychology*, *23*, 112–120.

Figure 1

Illustration of the audio-visual matching paradigm: numerical and control tasks.

In all tasks, participants were first presented with an auditory stimulus (number word, tone sequence, color word or speech sound) and afterwards with a visual stimulus (a digit, a dot array, a color or a letter). They had to indicate whether what they saw matched with what they heard.

Numerical audio-visual matching tasks



Control audio-visual matching tasks



Table 1

Mean accuracies and corresponding standard deviations of the numerical audio-visual matching tasks used in Experiment 1, displayed per ratio and per number range.

Task	ACCURACIES (% CORRECT)				
	Number range				
	<i>Same numerical value</i>	Within subitizing range		Outside subitizing range	
		<i>Ratio</i>		<i>Ratio</i>	
	<i>Small: 0.50 (trials 2-4)</i>	<i>Large: 0.75 (trials 3-4)</i>	<i>Small: 0.56 (trials 5-9)</i>	<i>Large: 0.78 (trials 7-9)</i>	
Number word-digit matching	94 (5)	99 (4)	99 (3)	97 (5)	97 (5)
Number word-dots matching	86 (10)	97 (6)	96 (7)	97 (5)	79 (22)
Tones-digit matching	73 (16)	96 (7)	66 (22)	87 (15)	52 (18)
Tones-dots matching	72 (13)	97 (5)	72 (18)	85 (15)	48 (22)
Task	REACTION TIMES (MS)				
Number word-digit matching	443 (78)	437 (65)	435 (82)	459 (79)	456 (90)
Number word-dots matching	697 (185)	533 (121)	599 (89)	648 (37)	1043 (501)
Tones-digit matching	756 (200)	742 (237)	868 (317)	741 (182)	832 (380)
Tones-dots matching	846 (272)	652 (57)	831 (255)	732 (177)	866 (375)

Table 2

Mean accuracies, median reaction times and the corresponding standard deviations of the two tasks conducted in Experiment 2, per ratio and per number range.

Task	ACCURACIES (% CORRECT)					
	Number range					
	<i>Same numerical value</i>	Within subitizing range		Outside subitizing range		
		<i>Ratio</i>		<i>Ratio</i>		
	<i>Small: 0.50 (trials 2-4)</i>	<i>Large: 0.75 (trials 3-4)</i>	<i>Small: 0.56 (trials 5-9)</i>	<i>Large: 0.78 (trials 7-9)</i>		
Number word-digit matching	95 (6)	95 (7)	96 (7)	95 (6)	97 (6)	
Tones-dots matching	70 (14)	94 (7)	77 (21)	78 (20)	53 (23)	
Task	REACTION TIMES (MS)					
	Number word-digit matching	516 (138)	531 (139)	527 (125)	652 (220)	587 (192)
	Tones-dots matching	868 (212)	793 (149)	979 (250)	975 (243)	927 (280)

Appendix 1

Experiment 1: Reaction time results

Control audio-visual matching tasks.

Word-color matching task. Median reaction time (RT) on the correct responses of this task was 443.38 ms ($SD = 79.28$ ms).

Sound-letter matching task. Median RTs on the ‘non-match’ trials were submitted to a repeated measures analysis of variance (ANOVA) with distance (2 levels: 1 and 4) as within-subject variable. There was no main effect of distance, $F < 1$.

Numerical audio-visual matching tasks.

The median RTs of the numerical audio-visual matching tasks are shown per ratio and separately for the numbers within and outside the subitizing range in Table 1. The median RTs on the ‘non-match’ trials were submitted to a repeated measures analysis of variance (ANOVA) with task (4 levels: number word-digit matching, number word-dots matching, digits-tones matching and tones-dots matching), ratio (2 levels: small vs large) and number range (2 levels: within vs outside subitizing range) as within-subject variables².

All main effects were significant: There was a main effect of task, $F(3,30) = 51.902$, $p < .0001$, $\eta_p^2 = .838$, a main effect of ratio, $F(1,32) = 29.038$, $p < .0001$, $\eta_p^2 = .476$, and a main effect of number range, $F(1,32) = 14.377$, $p < .0001$, $\eta_p^2 = .310$. Moreover, there was a significant interaction between task and ratio, $F(3,30) = 18.827$, $p < .0001$, $\eta_p^2 = .653$, and between task and number range, $F(3,30) = 16.103$, $p < .0001$, $\eta_p^2 = .617$, which were in turn embedded in a three-way interaction between task, ratio and number range, $F(3,30) = 6.146$,

² In the analyses including the RT data of the tones-dots matching task, 1 participant dropped out because (s)he performed erroneously on all trials of certain conditions. Overall however, this participant was not an outlier.

$p = .002$, $\eta_p^2 = .381$. To disentangle these interactions, separate analyses per task were conducted (and post-hoc analyses per number range if necessary).

Number word-digit matching task. There was no main effect of ratio, $F < 1$. There was a main effect of number range, $F(1,33) = 8.671$, $p < .001$, $\eta_p^2 = .208$, showing faster reaction times on the trials with numbers from within the subitizing range. There was no interaction between ratio and number range, $F < 1$.

Number word-dots matching task. There was a main effect of ratio, $F(1,33) = 46.033$, $p < .0001$, $\eta_p^2 = .582$, demonstrating faster reaction times on the small ratio's than on the large ratio's. There was also a main effect of number range, $F(1,33) = 43.639$, $p < .0001$, $\eta_p^2 = .569$, demonstrating much faster reaction times on the trials with the smallest numbers (i.e. within the subitizing range). There was also an interaction between ratio and number range, $F(1,33) = 19.737$, $p < .0001$, $\eta_p^2 = .374$. Post-hoc paired t -tests showed significant ratio-effects in both number ranges: $t(33) = 5.231$, $p < .0001$, and $t(33) = 5.649$, $p < .0001$, for within and outside the subitizing range respectively, although visual inspection of the data suggested a larger ratio effect in the larger number range.

Tones-digit matching task. There was a main effect of ratio, $F(1,33) = 9.535$, $p < .001$, $\eta_p^2 = .224$, demonstrating faster reaction times on the small ratio's than on the large ratio's. There was no main effect of number range and no ratio \times number range interaction.

Tones-dots matching task. There was a main effect of ratio, $F(1,32) = 19.986$, $p < .0001$, $\eta_p^2 = .384$, demonstrating faster reaction times on the small ratio's than on the large ratio's. There was no effect of number range and no ratio \times number range interaction.

Numerical ratio effects.

The size of the ratio effect was computed for each participant by subtracting the ~~average reaction time on trials~~ average reaction time on trials with the most difficult ratios (i.e. 0.75 and 0.78) from the average reaction time on trials with the easiest ratios (i.e. 0.50 and 0.56). This difference was

then divided by the average reaction time on trials with the easiest ratios to correct for individual differences in reaction times (see Holloway & Ansari, 2009; Sasanguie et al., 2013 for a similar method). We compared the size of the ratio effect in the number word-digit matching task with the weakest ratio effect in one of the other three numerical tasks, to statistically test the absence of a ratio effect in the number word-digit task. Paired *t*-tests showed that both ratio effects significantly differed from each other, $t(33) = 3.991$, $p < .0001$.