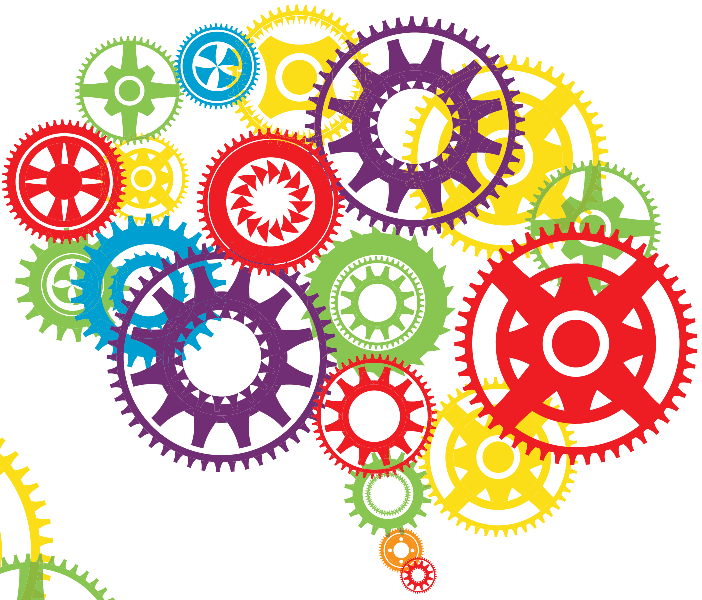
# THE FUNCTIONAL AND NEURAL MECHANISMS OF NUMEROSITY PROCESSING: FROM PERCEPTION TO COGNITION

EDITED BY: Michele Fornaciai, Joonkoo Park and Roberto Arrighi
PUBLISHED IN: Frontiers in Human Neuroscience



frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-046-6
DOI 10.3389/978-2-88976-046-6

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

# THE FUNCTIONAL AND NEURAL MECHANISMS OF NUMEROSITY PROCESSING: FROM PERCEPTION TO COGNITION

Topic Editors:
**Michele Fornaciai,** International School for Advanced Studies (SISSA), Italy
**Joonkoo Park,** University of Massachusetts Amherst, United States
**Roberto Arrighi,** University of Florence, Italy

**Citation:** Fornaciai, M., Park, J., Arrighi, R., eds. (2022). The Functional and Neural Mechanisms of Numerosity Processing: From Perception to Cognition. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88976-046-6

# Table of Contents

frontiers | Frontiers in Human Neuroscience

EDITORIAL
published: 07 April 2022
doi: 10.3389/fnhum.2022.880645

# Editorial: The Functional and Neural Mechanisms of Numerosity Processing: From Perception to Cognition

*Michele Fornaciai*[1]*, *Joonkoo Park*[2,3] *and Roberto Arrighi*[4]

*[1] Department of Cognitive Neuroscience, International School for Advanced Studies (SISSA), Trieste, Italy, [2] Department of Psychological and Brain Sciences, University of Massachusetts Amherst, Amherst, MA, United States, [3] Commonwealth Honors College, University of Massachusetts Amherst, Amherst, MA, United States, [4] Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy*

**Keywords: numerosity perception, numerical cognition, subitizing, groupitizing, mathematical abilities, perceptual adaptation, spatial-numerical association, connectedness**

**Editorial on the Research Topic**

**The Functional and Neural Mechanisms of Numerosity Processing: From Perception to Cognition**

Perceiving numerosity—the number of discrete items in a set—represents a fundamental step to understand the surrounding environment. Its ubiquitous nature across animal species suggests that it provides important advantages for survival. It is also thought to serve as an important basis for advanced mathematical thinking in humans. The diverse nature of the perceptual and cognitive functions linked to numerosity has attracted a large number of researchers with different perspectives, methodologies, and levels of investigation. In line with this, we present here a series of ten articles capturing the multifaceted nature of numerosity perception and numerical cognition.

First, some of the contributions aimed to achieve a deeper understanding of the brain mechanisms tuned to numerosity by leveraging on perceptual illusions and/or contextual effects.

For example, Li et al. investigated numerosity perception in the periphery of the visual field, an area of particular interest due to its lower spatial "resolution" and the tendency to pool information across larger spatial extents. The results indicate that numerosity estimates for a given target area are robustly distorted by irrelevant contextual information in the surrounding areas, with the relative weight of these two sources of information depending on the position participants deployed their attention to.

Numerosity is related to space not only when it comes to central vs. peripheral vision, but also in terms of how quantities are mapped to spatial location (i.e., along a "mental number line") and to spatial extent (i.e., the coupling of numerical and spatial magnitude). Viarouge and de Hevia addressed the interaction between these two types of mapping, showing that they may arise from a single representational system.

Another powerful tool to study numerosity perception is adaptation. Here, Togoli and Arrighi leveraged on this technique to show that adaptation generalizes not only across vision and audition but also touch, despite this latter modality exploits a completely different reference frame (hand/body centered). Haptic numerosity adaptation is indeed able to strongly distort perceived numerosity presented visually and auditorily, bolstering the idea that numerosity processing is modality independent.

***Edited and reviewed by:***
*Lutz Jäncke,*
*University of Zurich, Switzerland*

****Correspondence:***
*Michele Fornaciai*
*mfornaci@sissa.it*

***Specialty section:***
*This article was submitted to*
*Cognitive Neuroscience,*
*a section of the journal*
*Frontiers in Human Neuroscience*

***Received:*** *21 February 2022*
***Accepted:*** *25 February 2022*
***Published:*** *07 April 2022*

***Citation:***
*Fornaciai M, Park J and Arrighi R (2022) Editorial: The Functional and Neural Mechanisms of Numerosity Processing: From Perception to Cognition.*
*Front. Hum. Neurosci. 16:880645.*
*doi: 10.3389/fnhum.2022.880645*

Adaptation is not the only process that can bias numerosity perception. Numerosity is indeed intertwined with several continuous magnitudes (e.g., area, density) that could potentially interfere with it. Castaldi et al. addressed the influence of non-numerical attributes in numerosity perception in the context of working memory (WM) resources. When WM resources are deprived, numerosity perception becomes more vulnerable to interference, suggesting that WM plays a role in preventing non-numerical attributes from biasing numerosity perception.

Another interesting perceptual distortion is the "connectedness" illusion, in that connecting pairs of items in a display strongly reduces perceived numerosity. Is this connectedness illusion an automatic, passive process, or does it involve an active segmentation? Pomè et al. show that connectedness requires attentional resources, suggesting that this form of perceptual organization is likely an active process.

Grouping of visual elements (for example by connecting pairs of items) is not always detrimental to numerosity perception. Indeed, clustering a visual array into small groups (i.e., no more than 4 items) improves the discrimination of numerosity–an effect called "groupitizing." Is groupitizing a purely visual mechanism, or does it involve an amodal mechanism as in the case of adaptation? Anobile et al. addressed this question, showing that the clustering of an auditory sequence of tones in small sub-groups significantly improves numerosity discrimination.

This groupitizing phenomenon however relies on another important mechanism, which is the *exact* estimation of very small ($\leq 4$) numerosities. This mechanism is called "subitizing" and is in contrast with the approximate estimation of larger numerosities. Fu et al. used EEG to investigate the encoding of approximate and subitizable numerosities during memory retention, showing a clear difference in their signatures. Interestingly, the signature of small numerosity processing resembles the typical pattern of EEG activity observed in WM tasks, indicating the role of WM in subitizing.

The study of the brain mechanisms involved in numerosity perception is also important in light of robust evidence that it is closely linked to higher-level cognitive functions. For example, formal mathematical abilities have been often observed to correlate with approximate numerosity estimation, suggesting a potential role of numerosity perception as a precursor of mathematics. Here, Tokita and Hirota addressed the relation between approximate numerosity and numeracy in adults across different numerosity judgement tasks. The results show that approximate numerical abilities are significantly related to numeracy irrespective of task, consistent with the idea that there exist overlapping processing mechanisms between numerosity and math.

Moreover, Ma et al. further investigated the resilience of the link between numerosity and math to auditory sensory deprivation (i.e., early deafness). Similarly to vision, the results show that this link holds even in the absence of the auditory input, suggesting that the relation between math and numerosity develops in a way that is independent of any specific sensory modality.

Finally, Szkudlarek et al. investigated the intuitive mathematical abilities of children prior to the actual acquisition of mathematical knowledge. To do so, the authors tested the ability to perform approximate divisions using numerosity stimuli, finding that even children that could not perform simple divisions were still able to do this perceptually-driven divisive operation. These findings suggest that this form of "intuitive arithmetic" precedes mathematical education, and it could represent a mechanism mediating the relationship between numerosity and math.

Overall, this Article Collection not only provides an overview of the multifaceted fields of numerosity perception and numerical cognition but also provides novel insights into the mechanisms of numerosity processing and its relationship with mathematical abilities. The many findings reported in this collection point to three overarching ideas: (1) although rooted in low-level perception, numerosity processing recruits amodal mechanisms abstracted from sensory processing; (2) numerosity processing likely involves an active mechanism requiring attentional and WM resources as well as top-down inputs; and (3) numerosity perception and intuitive arithmetic abilities are likely related to mathematical abilities during development and in adulthood.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## FUNDING

This work was supported by the European Union's Horizon 2020 Research and Innovation Programme under the Marie Sklodowska-Curie grant agreement No. 838823 NeSt to MF and by the Italian Ministry of Education University and Research under the PRIN2017 program (Grant No. 2017XBJN4F-EnvironMag) to RA.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 09 June 2021
doi: 10.3389/fnhum.2021.688144

# Approximate Number Sense in Students With Severe Hearing Loss: A Modality-Neutral Cognitive Ability

**Hailin Ma[1,2†], Xiaoou Bu[2,3†], Emily M. Sanford[4], Tongao Zeng[2] and Justin Halberda[4*]**

[1] College of Education, Shanxi Normal University, Xi'an, China, [2] Plateau Brain Science Research Center, Tibet University, Lhasa, China, [3] Faculty of Education, EastChina Normal University, Shanghai, China, [4] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD, United States

The Approximate Number System (ANS) allows humans and non-human animals to estimate large quantities without counting. It is most commonly studied in visual contexts (i.e., with displays containing different numbers of dots), although the ANS may operate on all approximate quantities regardless of modality (e.g., estimating the number of a series of auditory tones). Previous research has shown that there is a link between ANS and mathematics abilities, and that this link is resilient to differences in visual experience (Kanjlia et al., 2018). However, little is known about the function of the ANS and its relationship to mathematics abilities in the absence of other types of sensory input. Here, we investigated the acuity of the ANS and its relationship with mathematics abilities in a group of students from the Sichuan Province in China, half of whom were deaf. We found, consistent with previous research, that ANS acuity improves with age. We found that mathematics ability was predicted by Non-verbal IQ and Inhibitory Control, but not visual working memory capacity or Attention Network efficiencies. Even above and beyond these predictors, ANS ability still accounted for unique variance in mathematics ability. Notably, there was no interaction with hearing, which indicates that the role played by the ANS in explaining mathematics competence is not modulated by hearing capacity. Finally, we found that age, Non-verbal IQ and Visual Working Memory capacity were predictive of ANS performance when controlling for other factors. In fact, although students with hearing loss performed slightly worse than students with normal hearing on the ANS task, hearing was no longer significantly predictive of ANS performance once other factors were taken into account. These results indicate that the ANS is able to develop at a consistent pace with other cognitive abilities in the absence of auditory experience, and that its relationship with mathematics ability is not contingent on sensory input from hearing.

**Keywords: approximate number sense, mathematics, cognition, hearing loss, domain general**

**Edited by:**
*Michele Fornaciai,*
*International School for Advanced Studies (SISSA), Italy*

**Reviewed by:**
*Ariel Starr,*
*University of Washington, United States*
*Mathieu Guillaume,*
*Stanford University, United States*

***Correspondence:**
*Justin Halberda*
*halberda@jhu.edu*

*†These authors have contributed equally to this work and share first authorship*

**Specialty section:**
*This article was submitted to Cognitive Neuroscience, a section of the journal Frontiers in Human Neuroscience*

***Received:** 30 March 2021*
***Accepted:** 12 May 2021*
***Published:** 09 June 2021*

**Citation:**
*Ma H, Bu X, Sanford EM, Zeng T and Halberda J (2021) Approximate Number Sense in Students With Severe Hearing Loss: A Modality-Neutral Cognitive Ability. Front. Hum. Neurosci. 15:688144. doi: 10.3389/fnhum.2021.688144*

## INTRODUCTION

Mathematical competence is essential to a wide range of activities in most modern cultures. Previous studies suggest that math ability predicts a variety of long-term consequences such as job attainment and success (Rivera-Batiz, 1992), socio-economic status (Ritchie and Bates, 2013), and health care decisions (Reyna et al., 2009). A wealth of research suggests that individual differences in math abilities depend on many factors, including home learning environment (LeFevre et al., 2010), teacher characteristics (Klibanoff et al., 2006; Beilock et al., 2010), and domain-general skills such as IQ, working-memory, and inhibitory control (Rohde and Thompson, 2007; Gilmore et al., 2013). Recent evidence suggests that there is also an innate, non-symbolic sense of quantity that gives rise to our basic numerical intuitions. A component of this broader number sense emerges from an evolutionarily and ontogenetically ancient Approximate Number System (ANS) which is present in human infants in the first year of life (Xu and Spelke, 2000; Izard et al., 2009; Feigenson et al., 2013; Starr et al., 2013), as well as exhibited by other non-verbal populations including monkeys, fish, rats, chicks, and birds (Feigenson et al., 2004; Agrillo et al., 2012). The ANS is a mental system of approximate number representations that is activated during symbolic and non-symbolic number tasks, which can be modeled by a series of Gaussian curves organized on a mental number line (Gallistel and Gelman, 1992; Dehaene et al., 2003; Piazza et al., 2004; Mazzocco et al., 2011a,b). The key signature of the ANS is that it represents numerical information in an imprecise way, with the imprecision in its representations increasing with the numerosity. Indexing this signature, and the acuity of the ANS, can be captured by a Weber fraction ($w$) which varies between individuals, where a smaller Weber fraction corresponds to higher precision (Pica et al., 2004; Halberda et al., 2008; Halberda and Feigenson, 2008; Piazza et al., 2010; Libertus et al., 2011; Odic et al., 2014). In accordance with Weber's Law, the difficulty of discriminating two numerosities depends on their ratio rather than their absolute difference (Piazza et al., 2004). For example, it is equally difficult to distinguish which of 8 vs. 16 is larger as it is to distinguish which of 16 vs. 32 is larger. In humans, ANS acuity increases with age, peaking at around 30 years of age (Halberda et al., 2012).

## Math Achievement and ANS in Students With Hearing Loss (SHL) and Students With Normal Hearing (SNH)

Deaf individuals are generally considered to be lagging behind hearing peers in mathematical tasks across a wide age range (Ansell and Pagliaro, 2006). They show delays in abstract counting and scores on standardized tests (e.g., arithmetical problem solving, logical reasoning, and understanding of fractional concepts) (e.g., in 2–3.5-year-olds, Pagliaro and Kritzer, 2010). Mitchell (2008) found that deaf students are significantly below grade level, exiting high school with about a 5th–6th grade level of mathematical achievement. Previous research has shown that because of the impoverished language environments, their hearing losses and limited access to wide-ranging numerical experience, many Students with Hearing Loss (SHL) are deficient in early quantitative concepts (Nunes, 2004; Kritzer, 2009a; Pagliaro and Kritzer, 2010; Pixner et al., 2014). Madalena et al. (2015) found that SHL with an early exposure to a sign language show better performance than those with a late exposure to the same language. Home environment may differ between typically hearing families and families with a SHL. Indeed, typically hearing families increase the probability of occurrence of informal and natural interactions involving numerical knowledge unconsciously by questioning, asking for clarification, or providing additional information in daily life activities (Kritzer, 2009b; Levine et al., 2010).

The ANS is often assumed to relate to arithmetic performance throughout childhood, adolescence and the adult years and current ANS acuity predicts future math ability (Halberda et al., 2008; Mazzocco et al., 2011b; Libertus et al., 2013). After controlling for scientific ability, writing ability and computer proficiency, the correlation between ANS acuity and mathematical ability of subjects aged 11–85 remained significant across the lifespan (Halberda et al., 2012). In addition, ANS acuity contributes to individual differences not only in the general population, but also in some special groups. Young adults with William's Syndrome performed poorly on both symbolic math and ANS tasks (Libertus et al., 2014), while Wang et al. (2017) found that ANS acuity was linked to symbolic math performance in gifted adolescents. Others found that students with specific math impairment (dyscalculia) performed significantly more poorly on the ANS task than their typically developing peers; in other words, less precise ANS representations are related to difficulty in mathematics broadly (Geary et al., 2008; Piazza et al., 2010; Bull et al., 2011; Skagerlund and Träff, 2016).

As mentioned above, SHL show a range of mathematical difficulties but whether the mechanism of this difficulty is the same as that of students with normal hearing is not known. For instance, in SHL, it may be that the innate ANS representations are as precise as their peers, while the mapping between ANS and more complicated mathematical concepts is delayed due to reduced access to linguistic and mathematical input. If the differences found between SHL and SNH in mathematics performance are due to their differences in experience rather than a difference in their innate ANS representations, there remains a question of whether it is due to a general lack of auditory input, related to delays in access to language or higher-level math concepts, or due to fundamental differences in information processing among SHL (Bull, 2008). In the present study, we aimed to document the potential relationship between hearing loss (and the many factors that covary with it) and ANS acuity.

## ANS and Domain-General Abilities

Because of the potential importance of domain-general abilities to both formal mathematics success and developing ANS acuity, we considered multiple examples of such abilities in the present study.

## Inhibition

Inhibition is thought to be important to performance in ANS tasks. Performance on trials where spatial characteristics can vary widely, e.g., in stimuli that are congruent or incongruent with numerical information. For instance, Clayton et al. (2015) found that people are much more accurate on trials where the larger set numerically also has the larger convex hull (congruent trials) than on trials where the opposite is the case (incongruent trials). Other non-numerical features that can influence responses on number tasks include surface area, diameter, perimeter, and density (Dakin et al., 2011; Gebuis and Reynvoet, 2011).

In order to reduce the extent to which subjects in ANS experiments can rely on non-numerical cues, one strategy has been to use two interleaved groups of stimuli: one where the average dot size is constant across both sets (such that the more numerous set also has a larger e.g., total dot area), and one where the total dot area is constant across both sets (such that the more numerous set has a smaller average dot size; Halberda et al., 2012). By mixing presentations of trials from these two stimulus sets together, neither average dot size nor total area is a reliable predictor of the number of objects throughout the experiment. If a subject tends to rely on a continuous feature such as dot size, then they will show different performance on size congruent and incongruent trials: this feature will help them respond correctly on congruent trials but will result in worse performance on incongruent trials, unless they are able to selectively suppress that signal on incongruent trials. Therefore, in order to consistently perform well on both congruent and incongruent trials, one must exert inhibitory control not unlike that required for a Stroop task. Relative differences in performance between congruent and incongruent trials may reflect differences in inhibitory control across subjects. Given that SHL are often reported to have inhibitory difficulties (Titus, 1995; Traxler, 2000), it may be that part of the source of SHL's mathematical challenges come from a relative lack of inhibitory control.

## Visual Working Memory

Visual-spatial processing is important for number perception, possibly because of the important role it plays in the formation of set representations from visual sets (Paul et al., 2017). In fact, visual form perception and visual short-term memory have been found to fully account for the relationship between ANS acuity and arithmetic performance in some instances (Zhang et al., 2019). This may be particularly important in young children, who appear to use visuospatial strategies when performing mental arithmetic more than older children (McKenzie et al., 2003), and where visual-spatial short-term memory span increases from 3 to 8 years of age (Pailian et al., 2016), and where visual-spatial short-term memory span has been found to be selectively predictive of math success in young children (Bull et al., 2008).

Further support that visual-spatial processing and working memory are important for number perception comes in the form of co-occurring challenges with number processing and working memory. Deficits in visual-spatial working memory have been found to be associated with numerical magnitude processing weaknesses in children with mathematical learning disabilities (Andersson and Östergren, 2012). Children with developmental dyscalculia have shown math-specific impairments as well as deficits in visuo-spatial working and short-term memory and inhibitory control (Szucs et al., 2013). Notably, ANS acuity differences between typically developing children and children with developmental dyscalculia have been found to be more extreme on size-incongruent trials than size-congruent trials. Because of the role that visual working memory plays in extracting numerical information from visual scenes, it may be extremely important to investigate in situations where ANS acuity varies between populations (Bugden and Ansari, 2015).

Conflicting evidence suggests that visual working memory capacity cannot fully explain numerical deficits. For instance, Peng et al. (2017) found that numerical knowledge mediates the relationship between ANS performance and early arithmetic abilities, above and beyond that which is explained by visuospatial processing. Additionally, research on children born extremely preterm found ANS acuity deficits that were not explainable on the basis of working memory or attention abilities (Libertus et al., 2017). Further research is necessary to investigate the extent to which visual working memory capacity can explain ANS acuity differences between populations. Given that deaf children have been shown to have deficits in visual working memory (López-Crespo et al., 2012), this question is particularly relevant for the current study.

## Attention Network

Numerical processing involves the deployment of attention, more so for subitizing than for large number processing (Anobile et al., 2012). In fact, some studies have found that estimation of large numbers is relatively unaffected by tasks with conflicting attentional demands (Burr et al., 2010). Nonetheless, spatial attention has sometimes been found to be involved in ANS task performance (Anobile et al., 2012).

When studying attention related to other cognitive abilities, the attentional system is sometimes divided into three separate components: alerting, orienting, and executive control attention networks (Fan et al., 2009). The alerting portion refers to the ability to increase attention at the expected onset of a new stimulus. The orienting attention network is thought to explain the ability to select a particular target for attention among a variety of inputs, whether intentional or through attention capture. Finally, the executive control network is thought to detect and resolve conflicts between co-occurring mental computations.

Attentional network development has been a topic of particular interest in deaf children (Daza and Phillips-Silver, 2013). The development of the alerting network is thought to be impaired in the absence of auditory stimulation, while some components of the orienting attention network are enhanced, such as moving and engaging. The executive control network has been found to develop along a similar trajectory to that of hearing children.

There is known to be a strong relationship between number processing, math ability, and attention (Anobile et al., 2013). Like ANS perception, performance on attention tasks has been found to predict symbolic math achievement in children and was also predictive of ANS ability (Anobile et al., 2013). Attentional deficits may be implicated in math-specific disabilities such as developmental dyscalculia. Therefore, we are interested in whether similar attentional deficits impact the numerical processing of SHL, and whether these deficits can be traced to specific attentional networks.

## Summary: Motivations for the Current Work

Considering the ways in which school mathematics abilities might be related to the ANS, and vice versa, it is likely to be a highly interdependent relationship. The Defective Number Module Hypothesis, perhaps too simply, suggests that mathematical deficits may have their roots in innate difficulties processing non-symbolic number; for instance, an impairment of the ANS has been proposed as the origin of dyscalculia, a mathematics-specific learning disability (Butterworth, 2005; Mazzocco et al., 2011a). Of course, an effect in the opposite direction might also occur. Such a relationship might be explored in SHL; not because SHL necessarily have dyscalculia themselves; but rather, because the reduced exposure of SHL to numerical concepts in early development may lead to similar problems (Swanwick et al., 2005). A relatively small number of studies have investigated the performance of SHL on specific areas of mathematics (Pagliaro and Kritzer, 2010). The current study expands this by focusing both on symbolic mathematics and non-symbolic numerical processing.

In Experiment 1, we tested whether SHL's responses to the ANS task conform to Weber's law, and investigated whether their acuity is affected by size congruency manipulations (e.g., Clayton et al., 2015). We expected that SHL's ANS responses will follow Weber's law, and that they will perform better on size-congruent than size-incongruent trials – just as SNH.

In Experiment 2, we compared SHL to a population of SNH, to test whether effects such as size congruency influence ANS acuity similarly between the two groups. It is possible that congruency manipulations would be especially detrimental to SHL, since they may have particular difficulties with inhibition (Titus, 1995; Traxler, 2000), and inhibition ability is thought to play an important role in mitigating the influence of size congruency on number responses (Clayton and Gilmore, 2015; Norris and Castronovo, 2016). We then explored the extent to which ANS acuity predicts mathematics ability when taking into account other factors such as inhibitory control, visual working memory capacity, and attention network performance. If SHL perform like other students their age, we would expect to see Weber fraction uniquely account for mathematics ability, above and beyond the contributions of these other factors (Halberda et al., 2008; Chen and Li, 2014; Schneider et al., 2016).

# EXPERIMENT 1

## Materials and Methods

### Participants

One hundred and forty-four students with hearing loss (mean age = 13.58 years, $SD$ = 2.34, range = 8–18 years; 60 females) from 6 special education schools participated in the study. All were enrolled in the third grade to ninth grade. SHL were prelingually deaf students and exhibited severe (71–90 dB) and profound hearing loss (>91 dB). All of them were right-handed, with normal or corrected-to-normal vision and no history of neurological or psychiatric illness.

### ANS Acuity

We administered a version of Panamath (Psychophysical Assessment of Number-Sense Acuity; www.panamath.org), a non-symbolic numerical comparison task, to assess the acuity of children's ANS. The two spatially intermixed arrays of blue and yellow dots were presented for 1,200 ms followed by a 200 ms backward mask, followed by a blank gray screen until the response was completed. Students were asked to judge whether more of the dots were blue or yellow. There were between 5 and 21 dots in each array, the ratios were categorized into 4 ratio bins: 1.14, 1.2, 1.33, and 2, with 20 trials in each ratio bin, yielding a total of 80 trials. To avoid subjects from relying on the cumulative area of dots, on half of the trials dots were size-confounded, and on the other half of the trials dots were size-controlled. Notice that Panamath does not systematically control for all possible non-numerical cues (e.g., convex hull is only partially controlled via the total area and dot size manipulations). Our aims were to test for Weber's law and (in Experiment 2) to test for the relationship of ANS acuity to formal math abilities. Our interest in inhibitory control was test here only by our area manipulation.

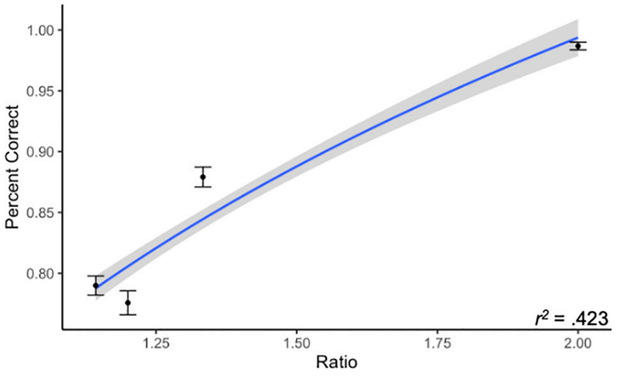


**FIGURE 1 |** Accuracy by trial ratio. We found that subjects in Experiment 1 conformed to Weber's law, where accuracy increases as a function of ratio. Line represents best-fitting logarithmic relationship between ratio and percent correct. Gray region represents 95% CI.

## Results and Discussion

Overall, subjects in Experiment 1 had relatively high accuracy on the ANS task ($M$ = 85.8%, $SD$ = 6.2%). We confirmed that accuracy improved as a logarithmic function of increasing ratio, as is expected with data conforming to Weber's law (Dehaene, 2003). We evaluated this by performing a linear regression predicting subjects' average accuracy (on both trial types) from the logarithm of trial ratio. We found that this model significantly predicted accuracy, β = 0.651, $t(574) = 20.55$, $p < 0.001$ (see **Figure 1**). This result indicates that, among these subjects, accuracy was dependent upon the trial difficulty as determined by comparison ratio, consistent with Weber's law.

Next, we were interested in whether ANS performance improved with age. We were also interested in whether performance was better on size-congruent trials than size-incongruent trials. Both of these effects have been found repeatedly in previous research on the ANS (e.g., Halberda et al., 2008, 2012; Clayton et al., 2015; Smets et al., 2015).

We used subjects' accuracy (for all trials, as well as separately for size-congruent and size-incongruent trials) to different ratios ($r$) to fit their Weber fraction ($w$) according to the following psychophysical model, used extensively in previous ANS research

(Pica et al., 2004; Cantlon and Brannon, 2006; Halberda and Feigenson, 2008; Halberda et al., 2008, 2012; Piazza et al., 2010; Libertus et al., 2011, 2013, 2014; Odic et al., 2013, 2014; DeWind et al., 2015; DeWind and Brannon, 2016; Starr et al., 2017; Wang et al., 2017):

$$probability\ correct = 1 - \frac{1}{2} erfc\left(\frac{r-1}{w\sqrt{2}\sqrt{1+r^2}}\right)$$

The model was fit to each subjects' data using Maximum Likelihood Estimation (MLE) in R. Previous research has indicated that accuracy and response time may index different abilities (e.g., Halberda et al., 2012), and because we were interested in the amount of internal noise in our subjects' number representations, we focused on using accuracy-based Weber fractions to test our hypotheses.

In this model, a smaller Weber fraction corresponds to higher accuracy and therefore better performance. On average, the subjects in this study had a mean Weber fraction of 0.168, which is in line with previous research on ANS acuity among 14-year-olds (the mean age of our participants), who have been found to have Weber fractions ranging from 0.119 to 0.567 (Halberda et al., 2008).

To evaluate whether performance improved with age, we performed a linear regression predicting Weber fraction (based on all trials) from subject age, expecting to see a negative linear trend (indicating that performance improved with age). Indeed, that was what we found: increasing age significantly predicted a decline in Weber fraction, $\beta = -0.212$, $F(1, 142) = 6.69$, $p = 0.011$, $R^2 = 0.04$.

Next, we investigated whether subjects performed differently on the size-confounded versus size-controlled trials, expecting that subjects would have higher Weber fractions (i.e., worse performance) on size-controlled trials than size-confounded trials. A paired t-test confirmed that subjects had smaller Weber fractions and therefore performed better on the size-confounded ($M = 0.15$, $SD = 0.09$) than size-controlled ($M = 0.19$, $SD = 0.12$) trials, $t(143) = 4.11$, $p < 0.001$.

This preliminary study demonstrates our ability to work with SHL in the relevant schools, and replicates several key findings from the literature on the ANS.

# EXPERIMENT 2

## Materials and Methods

### Participants

In Experiment 2, we focused on a subgroup of the children from Experiment 1 and also ran a new group of age-relevant controls. In order to focus on effects related to symbolic math development, we relied on the Chinese Rating Scale of Pupil's Mathematic Abilities (C-RSPMA; Wu and Li, 2005) which is normed for children in primary school. For this reason, we restricted our SHL sample to children in primary school with complete datasets as well as a new group of control children with complete data sets. Ninety-seven SHL ($M_{age} = 12.58$ years, $SD = 1.95$, range = 8–18 years; 38 females) from 6 special education schools and 97 SNH ($M_{age} = 10.36$ years, $SD = 1.24$, range = 8–12 years; 47 females) from 1 normal primary school in Sichuan, China, participated in the study. All were enrolled in the third grade to sixth grade. The SNH students were approximately matched to the SHL in grade level (although SHL were on average older than SNH and had a much wider age range, as is typical in SHL). SHL were prelingually deaf students and exhibited severe (71–90 dB) and profound hearing loss (>91 dB). All subjects were right-handed, with normal or corrected-to-normal vision and no history of neurological or psychiatric illness. **Table 1** shows detailed demographic information on all participants.

### Tasks and Procedure

***Chinese rating scale of pupil's mathematic abilities***

The Chinese Rating Scale of Pupil's Mathematic Abilities (C-RSPMA; Wu and Li, 2005) based on the Germany Rating Scale of Pupil's Mathematic Abilities established by Heidelberg University was used to assess the primary students' basic mathematical competencies.

C-RSPMA is composed of 11 subtests divided into two broad categories. One category tests mathematics operation such as addition, subtraction, multiplication, division, blank filling and comparisons. The other category focuses on skills in spatial vision and logical thinking, with tasks such as figure writing, length estimation, block counting, graph counting and figure connection. For these 11 subtests, students were required to answer as many items as possible within the stipulated time (1–3 min, dependent on different subtests). The Cronbach's alpha is above 0.7, split-half reliability coefficient is 0.83.

***Non-verbal IQ***

To evaluate children's non-verbal IQ, we administered the combined Raven's Test (CRT-CC3; Wang et al., 2007). This test contains 72 matrices of increasing difficulty, and a correct answer yielded one point. Students were required to identify the missing element that best completes a pattern from six or eight alternatives.

***Inhibition***

The Flanker Task was used to measure inhibitory control (Eriksen and Eriksen, 1974; see **Figure 2A**). This task measures inhibitory control by requiring subjects to respond in the direction of a central arrow while ignoring the sometimes-conflicting direction of the arrows on either side of it. Each trial started with a

**TABLE 1** | Experiment 2 participant demographic information.

| | SHL | | | SNH | | |
|---|---|---|---|---|---|---|
| | Age (M ± SD) | Sex | | Age(M ± SD) | Sex | |
| Group | | Male | Female | | Male | Female |
| 3 grade | 11.57 ± 1.93 | 14 | 9 | 8.70 ± .56 | 10 | 13 |
| 4 grade | 12.35 ± 1.90 | 15 | 8 | 9.84 ± .36 | 11 | 11 |
| 5 grade | 12.14 ± .94 | 12 | 10 | 10.83 ± .39 | 12 | 11 |
| 6 grade | 13.93 ± 1.93 | 18 | 11 | 11.72 ± .45 | 17 | 12 |

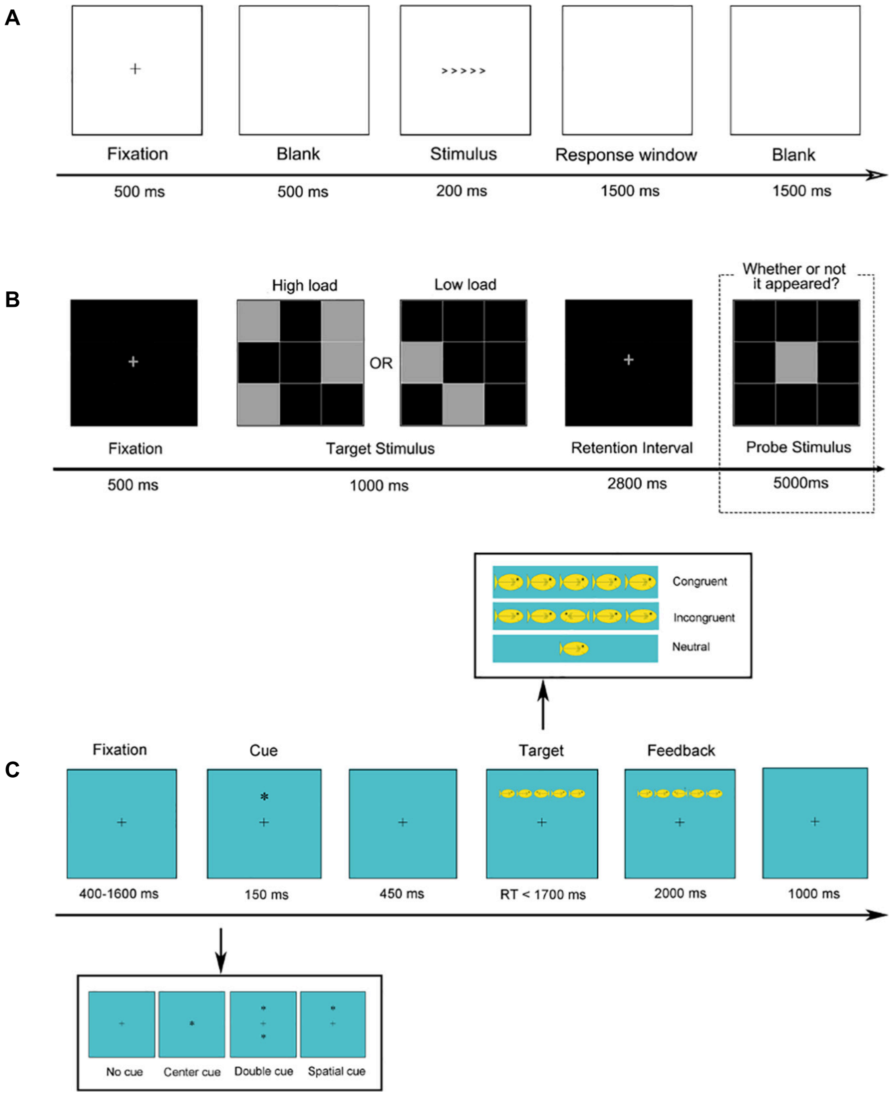


**FIGURE 2 |** The flowchart of the experimental paradigms. **(A)** The flanker task. **(B)** Visual delayed match-to-sample task. **(C)** Attention network task.

fixation cross presented centrally for 500 ms followed by a blank screen for 500 ms, after which the target and flanking stimuli appeared. These stimuli were presented for 200 ms followed by a response window until a response was made up to 1,500 ms later. A blank screen of 1,500 ms separated each trial. Half of the trials were congruent (<<<<< or >>>>>), whereas the other half were incongruent (e.g., <<><< or >><>>). Students were instructed to respond as accurately and as quickly as possible to indicate the direction of the centrally presented target arrow by key press. This task contained a practice block with 12 trials and two experimental blocks with 60 trials each.

### *Visual working memory*

We used a visual delayed match-to-sample task to measure visual working memory (Dong et al., 2014; **Figure 2B**). A fixation cross was presented for 500 ms followed by a target stimulus, which was a grid that had some squares highlighted (high load condition: 4/9 highlighted; low load condition: 2/9 highlighted). The target stimulus was visible for 1,000 ms, followed by a blank screen for 2,800 ms (where the participant had to hold the locations of the target squares in memory). Finally, a probe stimulus appeared, which consisted of the same grid but with only one square highlighted. Students indicated by button press whether or not the probe square appeared in one of the same locations as was highlighted in the target stimulus. The probe stimulus was visible for up to 5,000 ms. This task contained a practice block with 10 trials and one experimental block with 60 trials.

### *Attention network*

The attention network test is used to measure the efficiency of the three aspects of attentional networks (i.e., alerting, orienting, and conflict; Fan et al., 2002; Rueda et al., 2004; **Figure 2C**). Each trial began with a fixation presented at the center of the screen for a random duration between 400 and 1,600 ms, after which the cue stimulus appeared for 150 ms. Subsequently, the fixation was again presented for 450 ms followed by a target stimulus which appeared for a maximum duration of 1,700 ms, followed by feedback for 2,000 ms. Finally, a fixation of 1,000 ms separated each trial. This task consisted of one practice block of 12 trials and two experimental blocks involving 60 trials each.

The ANT includes four cue conditions (no cue, central cue, double cue, and spatial cue) and three target conditions (congruent, incongruent, and neutral). The target stimulus was a single yellow fish or a horizontal row of five yellow fish which were presented about 1°either above or below fixation. Each fish subtended 0.58° of visual angle and was separated from neighboring fish by 0.21°. The five fish subtended a total of 8.84°. Students were instructed to respond to the direction that a central fish was facing by button press.

### ANS acuity

The procedure to evaluate ANS acuity was the same as Experiment 1.

### Data preparation

The data preparation and Weber fraction modeling for the ANS results were identical to those used in Experiment 1.

The C-RSPMA was scored following standard protocol to calculate a Mathematics score for each subject (Wu and Li, 2005).

For the task measuring Non-verbal IQ, the final raw Raven test scores were converted to a standard IQ score according to the norm for Chinese children.

For the task measuring Inhibitory Control, we computed a score based on response time in the Flanker Task. An index of inhibitory control for each subject was calculated using the following formula over mean response times in the two conditions: Score = $RT_{incongruent} - RT_{congruent}$. This single value represents how much longer it took the subject to respond to incongruent trials than to congruent trials, and therefore a lower value corresponds to better inhibitory control.

For the visual working memory task, a composite score was created for working memory performance by combining results from both accuracy and response time. Across all subjects, we z-scored average response times on high memory load trials (correct responses only), average response times on low memory load trials (correct responses only), average accuracy on high memory load trials, and average accuracy on low memory load trials. This resulted in each subject having four values that indicated how well, relative to other subjects, they performed on each of these four indices of performance. We then averaged these four z-scores for each subject to get a single composite score of performance on the working memory task relative to other subjects in the sample.

For the Attention Network Task, we calculated a separate score for the efficiency of the three attentional networks based on response times to different cue conditions. The efficiency of three attentional network scores based on the RTs were calculated using the following formula (see **Figure 2C** for cue conditions): Alerting effect = $RT_{no-cue} - RT_{double-cue}$, Orienting effect = $RT_{center-cue} - RT_{spatial-cue}$, and Conflict effect = $RT_{incongruent} - RT_{congruent}$.

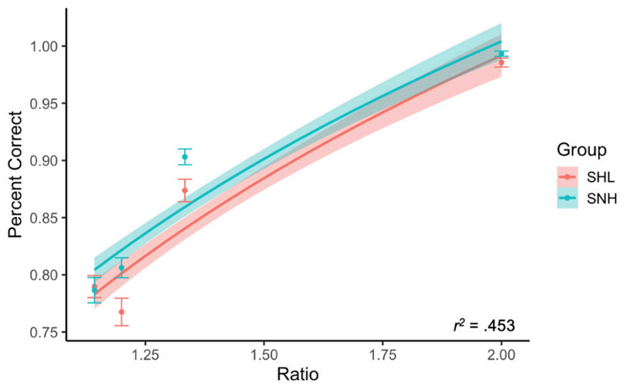


**FIGURE 3 |** Accuracy by trial ratio and group membership. As expected by Weber's law, accuracy increased as a function of trial ratio. SNH were, on average, more accurate than SHL. There was no interaction between these two variables. Lines represents best-fitting logarithmic relationship between ratio and percent correct. Shading corresponds to 95% CI.

# Results

## ANS Performance

Once again, our subjects performed fairly well in terms of accuracy on the ANS task ($M$ = 86.3%, $SD$ = 5.6%). We once again evaluated whether accuracy was dependent upon trial ratio to confirm that our results were consistent with Weber's law. We used multiple regression predicting accuracy from the logarithm of trial ratio, group membership (SNH or SHL), and their interaction. We expected that both groups would show a significant influence of trial ratio on accuracy, and that there would be no interaction between the two variables. This was confirmed: the logarithm of trial ratio significantly predicted accuracy, $\beta = 0.685$, $t(772) = 18.25$, $p < 0.001$. There was also a significant effect of group membership, where SNH ($M$ = 87.2%, $SD$ = 4.6%) had slightly higher accuracy on average than SHL ($M$ = 85.4%, $SD$ = 6.3%), $\beta = 0.076$, $t(772) = 2.85$, $p = 0.004$. There was no interaction between the two, indicating that trial difficulty impacted both groups the same relative amount, $p = 0.579$ (see **Figure 3**).

Next, we were interested in performance as indexed by model-fitted Weber fractions. Overall, with one Weber fraction fit to each subject's responses to all trials, our subjects had similar Weber fractions to those we found in Experiment 1 ($M$ = 0.162, $SD$ = 0.065).

We were interested in whether Weber fraction was affected by group membership (SHL vs. SNH) and size congruity (congruent vs. incongruent trials). To test this, we again fit each subject's responses with a Weber fraction, separately for size congruent and incongruent trials. Then we conducted a two-way ANOVA predicting Weber fraction from hearing and size congruency, with group membership as a between-subjects variable and size congruency as a within-subjects variable. Both main effects were significant. Consistent with the results from

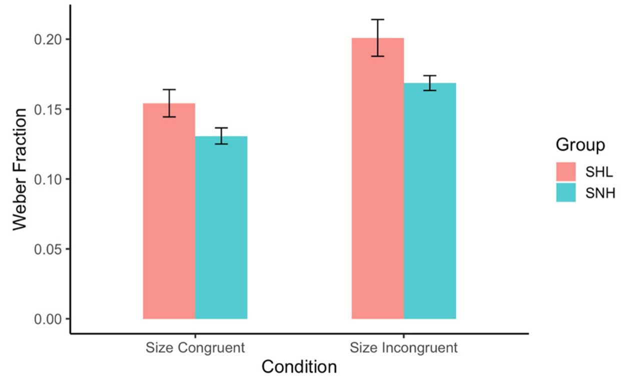


**FIGURE 4 |** Weber fractions by group membership and size-congruity. Weber fractions were lower (corresponding to more precise responses) when subjects were responding to size-congruent trials than when they were responding to size-incongruent trials. Also, SNH had higher acuity than SHL. There was no interaction between these two variables. Error bars represent standard error.

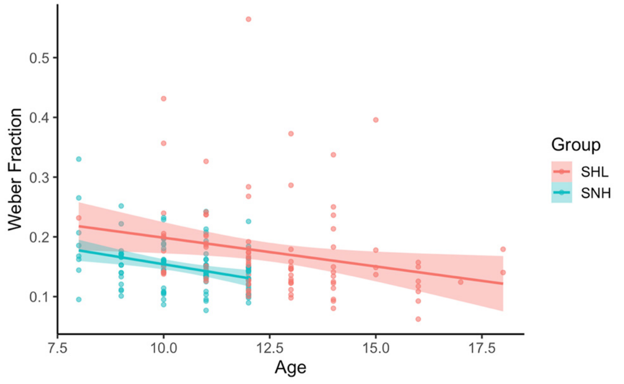


**FIGURE 5 |** Weber fraction by age and group membership. In both groups, increasing age was linked with decreasing Weber fractions. Line corresponds to best-fitting linear relationship between Age and Weber fraction for each group. Shading corresponds to 95% CI.

Experiment 1, we found that acuity was better on size-congruent trials ($M$ = 0.142, $SD$ = 0.080) than on size-incongruent trials ($M$ = 0.185, $SD$ = 0.100), $F(1, 382) = 5.044$, $p = 0.025$, across the two groups. For group membership, we found that SHL ($M$ = 0.174, $SD$ = 0.079) had larger Weber fractions than SNH ($M$ = 0.150, $SD$ = 0.043), $F(1, 382) = 7.34$, $p = 0.007$, indicating that SNH had slightly better acuity. Importantly, there was no significant interaction between these two factors, $F(1, 382) = 0.29$, $p = 0.634$ (see **Figure 4**). This indicates that size-congruency impacted performance equally for subjects regardless of group membership—which runs counter to the expectation that difficulty with inhibition would drive especially worse performance for SHL on size-incongruent trials.

We also verified whether Weber fraction varied with age in this sample. We investigated this by performing a linear regression predicting Weber fraction (collapsed across congruency conditions) from age and group membership. Consistent with previous research, we found that the model significantly predicted Weber fractions, $F(3, 190) = 7.03$, $p < 0.001$, $R^2 = 0.09$ (see **Figure 5**). Both group membership, $\beta = -0.366$, $t(190) = 4.24$, $p < 0.001$, and age, $\beta = -0.293$, $t(190) = 2.96$, $p = 0.003$, significantly predicted Weber fractions, while their interaction was not significant, $p = 0.741$. Within both groups, increasing age was linked to decreasing Weber fractions (meaning older subjects were more precise in their ANS responses than younger subjects), and the rate of this effect did not differ between the two groups.

## Relationship Between ANS Performance and Other Tasks

Next, we were interested in the extent to which ANS performance could be predicted by performance on other related tasks. We tested each subject in the following domains: Non-verbal IQ (Raven task), Inhibitory Control (Flanker task), Visual Working Memory, and Attention Network strength. Using the data processing techniques described above, this resulted in the creation of the following scores for each subject: Raven score; Inhibitory Control composite score; Working Memory composite score; and Conflict, Alerting, and Orienting ANT scores (for mean scores on each task, see **Table 2**). We then predicted Weber fractions from this group of variables, as well as Hearing group and Age. We included interaction terms between Hearing group and each other variable to evaluate whether the pattern of results differed between SHL and SNH.

On the whole, this model explained significant variance in ANS performance, $F(15, 178) = 3.233$, $p < 0.001$, $R^2 = 0.148$. We found that Age, $\beta = -0.306$, $p = 0.002$, Non-verbal IQ, $\beta = -0.238$, $p = 0.029$, and Visual Working Memory score, $\beta = -0.201$, $p = 0.046$, were each significant predictors when other variables were taken into account. Score on Conflict ANT was marginally significant, $\beta = -0.150$, $p = 0.096$, and no other variables were significant, $ps > 0.116$. Increases in each of these variables corresponded to decreases in Weber fractions, indicating that students who were older and had higher Non-verbal IQ, Visual Working Memory capacity, and scores on the Conflict ANT tended to have better ANS acuity. Notably, none of the interactions between group membership and other variables were significant, $ps > 0.240$, indicating that the relationship between ANS and other task performance was similar among SHL and SNH.

**TABLE 2 |** Mean scores on each task by hearing group.

| | SHL | | SNH | |
|---|---|---|---|---|
| | **Mean** | ***SD*** | **Mean** | ***SD*** |
| Non-verbal IQ (Raven) | 48.60 | 6.12 | 53.28 | 6.04 |
| Inhibitory control | 158.05 | 78.92 | 165.87 | 80.29 |
| Visual working memory | −0.352 | 0.79 | 0.352 | 0.58 |
| Conflict ANT | 131.41 | 67.63 | 104.63 | 55.36 |
| Alerting ANT | 29.89 | 42.77 | 21.09 | 39.98 |
| Orienting ANT | −1.47 | 52.90 | 8.47 | 47.38 |

Interestingly, hearing group membership was no longer predictive of ANS performance when the other variables were included, $\beta = -0.178$, $p = 0.116$. However, due to the decreased power associated with the large number of predictors included in this model, we caution against a strong interpretation of this result.

### Relationship Between ANS and Mathematics Performance

Finally, we were interested in the degree to which ANS performance could account for variability in formal mathematics scores ($M = 193.05$, $SD = 62.06$), above and beyond that which could be accounted for by other related abilities. We did this by utilizing the suite of predictors tested in the previous section (Hearing group; Age; Raven score; Inhibitory Control composite score; Working Memory composite score; and Conflict, Alerting, and Orienting ANT scores), and used linear regression to determine whether ANS performance predicted Mathematics performance once these variables were taken into account. As in the previous section, the only interactions included in this model were between hearing group membership and each other variable, to determine whether these variables had different explanatory power among SHL compared to SNH.

For our first model, we regressed Mathematics score over the suite of these predictor variables, excluding ANS performance. This model significantly predicted Mathematics ability, $F(15, 178) = 27.99$, $p < 0.001$, $R^2 = 0.68$ (see **Table 3** for standardized coefficients). Of the predictors, only the Attentional Network scores did not significantly explain some variance in Mathematics ability; group membership, Age, Non-verbal IQ, Inhibitory Control, and Visual Working Memory capacity all contributed to explaining Mathematics performance. SNH ($M = 232.88$, $SD = 41.88$) had significantly higher Mathematics scores on average than SHL ($M = 153.22$, $SD = 52.76$). Increasing Age, Non-verbal IQ, and Visual Working Memory capacity corresponded to increases in Mathematics Score. Interestingly, an increase in Inhibitory Control score corresponded to a *decrease* in Mathematics score. No interactions with group membership were significant, $p$s $> 0.191$, indicating that the influence of each variable on Mathematics score was similar for both groups.

We then compared this model to a second model that included the same predictors and additionally included ANS performance as indexed by Weber fraction. This model also explained a significant amount of variance in Mathematics score, $F(17, 176) = 26.28$, $p < 0.001$, $R^2 = 0.69$. Weber fractions were significantly predictive of Mathematics score even when other variables were taken into account, $t(176) = 3.03$, $p = 0.003$. Once Weber fraction was added to the model, Age, Visual Working Memory capacity and Conflict ANT score became marginally significant predictors of variance in Mathematics score (likely due to the shared variance between these predictors and ANS performance found in the previous section). Hearing group, Inhibitory Control and Non-verbal IQ remained significant predictors (see **Figure 6** for the individual relationship between each predictor and Mathematics score). As in the previous model, there was no interaction between group membership and any of the other predictors, $p$s $> 0.111$.

We then checked that the second model explained significantly more variance than the first model, to confirm that ANS task performance explained additional variance in our subjects' Mathematics scores. An ANOVA comparing these two models significantly favored the second model, and therefore the inclusion of ANS task performance, over the first model, $F(2, 176) = 4.70$, $p = 0.010$. ANS ability uniquely explained variance in Mathematics score beyond that which was explained by other predictors, and did so similarly for both SHL and SNH.

**TABLE 3 |** Standardized coefficients from regressions predicting mathematics score.

| Predictor | Model 1 β | Model 2 β |
|---|---|---|
| Hearing group | 0.559*** | 0.540*** |
| Age | 0.149* | 0.102. |
| Non-verbal IQ | 0.395*** | 0.359*** |
| Inhibitory control | −0.125* | −0.122* |
| Visual working memory | 0.146* | 0.115. |
| Conflict ANT | −0.078 | −0.101. |
| Alerting ANT | −0.071 | −0.078 |
| Orienting ANT | 0.040 | 0.035 |
| Weber fraction | | −0.154** |

*p < 0.1, *p < 0.05, **p < 0.01, ***p < 0.001.*

*Model 1β, the first model beta-coefficient; Model 2β, the second model beta-coefficient.*

## DISCUSSION

To summarize our results, we found that students with hearing loss (SHL) had lower ANS acuity than control subjects (SNH)—even though SHL tended to be a bit older. The magnitude of this effect was decreased when other factors were taken into account (such as Non-verbal IQ and Visual Working Memory capacity), indicating that the difference in ANS performance that we observed may be at least partially due to other factors that tend to vary between these groups, rather than due solely to the imprecision of the ANS representations themselves. All students showed a tendency to perform better on size-confounded than size-controlled trials, consistent with a role for inhibitory control. But, a specific role for reduced inhibitory control to drive especially low ANS acuity in SHL did not bear out. We found that many factors contributed uniquely to performance on the Math test, and most importantly, even when taking these other potential contributing factors into account, the precision of the ANS (Weber fraction) still accounted for significant variance in Math score. Therefore, we conclude that the ANS's contribution to Math ability in children goes above and beyond that which can be accounted for by other measures such as Inhibitory Control, Working Memory capacity, and Attention Network performance, and, to the extent that we find unique variance between ANS and symbolic math ability above and beyond these factors, these abilities may play only a minor role in modulating the link between the ANS and symbolic math ability.

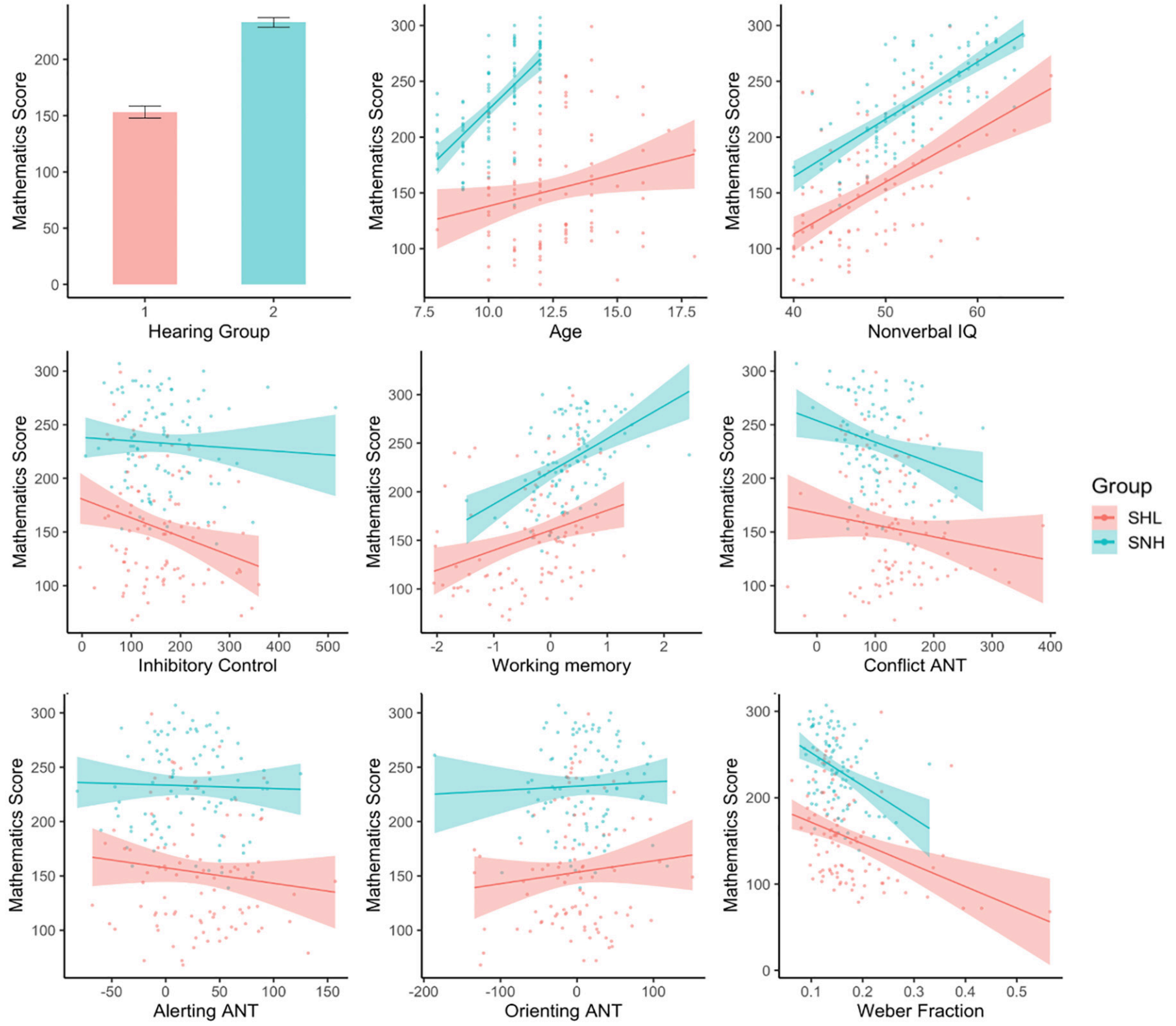


**FIGURE 6 |** Predicting mathematics score. When other variables were taken into account, only group, Non-verbal IQ, Inhibitory Control, and Weber fraction significantly predicted Mathematics score.

The present study adds further support for the claim that ANS abilities relate to school math abilities in children, consistent with previous meta-analyses on the topic (Chen and Li, 2014; Schneider et al., 2016). Here, we observed this effect in two large samples of students, controlling for many relevant factors. We also saw that this effect is important both for typically developing children and students with hearing loss (SHL). That we saw accuracy patterns consistent with Weber's law in our SHL (and only a small difference in Weber fraction between SHL and SNH when controlling for other factors) suggests that the ANS is able to develop somewhat normally in the absence of auditory input. SHL tended to have lower scores on many of the facilities tested in the present studies, which raises the possibility that the ANS deficits we saw are not specific deficits, but rather due to general developmental challenges that arise for deaf children, such as late-onset language exposure or reduced access to early mathematics education (Swanwick et al., 2005; Bull, 2008).

Combining this with the existing result of normal functioning of the ANS in blind participants (Kanjlia et al., 2018) supports the suggestion that the ANS is a domain general cognitive system with representations that abstract away from any particular modal signal. Although size-congruency influenced ANS performance in this sample (and convex hull was not controlled for), given that ANS representations develop in individuals with vastly different sensory experiences, we argue that the content of these shared representations must be something that is preserved across modalities (*see also* Halberda, 2019). That is, if the ANS is able to develop in both blind individuals and SHL, and given that links between the ANS and math ability are observed in both populations, it appears that the ANS abstracts away from particular modal content. Nonetheless, the mechanism underlying congruency effects, and whether they occur at the extraction or response stage, remains a fruitful path for future study.

As with many previous demonstrations, the present results suggest a picture of the ANS as a domain general cognitive system that supports non-symbolic numerical intuitions and relates to symbolic math abilities.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Tibet University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

XB and HM: study concept and design. XB, ES, and TZ: acquisition and analysis or interpretation of data. XB, ES, JH, and HM: drafting of the manuscript. HM and JH: obtained funding. ES, XB, JH, and HM: critical revision of the manuscript for important intellectual content. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Educational Scientific Research Projects of Tibet (XZJKY19104) and a McDonnell Scholar Award to JH. This material is also based upon work supported by the National Science Foundation Graduate Research Fellowship under grant no. DGE1746891.

## REFERENCES

Agrillo, C., Piffer, L., Bisazza, A., and Butterworth, B. (2012). Evidence for two numerical systems that are similar in humans and guppies. *PLoS One* 7:e31923. doi: 10.1371/journal.pone.0031923

Andersson, U., and Östergren, R. (2012). Number magnitude processing and basic cognitive functions in children with mathematical learning disabilities. *Learn. Indiv. Differ.* 22, 701–714. doi: 10.1016/j.lindif.2012.05.004

Anobile, G., Stievano, P., and Burr, D. C. (2013). Visual sustained attention and numerosity sensitivity correlate with math achievement in children. *J. Exp. Child Psychol.* 116, 380–391. doi: 10.1016/j.jecp.2013.06.006

Anobile, G., Turi, M., Cicchini, G. M., and Burr, D. C. (2012). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vision Res.* 74, 102–109. doi: 10.1016/j.visres.2012.06.005

Ansell, E., and Pagliaro, C. M. (2006). The relative difficulty of signed arithmetic story problems for primary level deaf and hard-of-hearing students. *J. Deaf. Stud. Deaf Educ.* 11, 153–170. doi: 10.1093/deafed/enj030

Beilock, S. L., Gunderson, E. A., Ramirez, G., and Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *P. Natl. Acad. Sci. U.S.A.* 107, 1860–1863. doi: 10.1073/pnas.0910967107

Bugden, S., and Ansari, D. (2015). Probing the nature of deficits in the 'Approximate Number System' in children with persistent Developmental Dyscalculia. *Dev. Sci.* 19, 817–833. doi: 10.1111/desc.12324

Bull, R. (2008). "Deafness, numerical cognition, and mathematics," in *Deaf Cognition: Foundations and Outcomes*, eds M. Marschark and P. C. Hauser (Oxford: Oxford University Press), 170–200. doi: 10.1093/acprof:oso/9780195368673.003.0006

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

Bull, R., Marschark, M., Sapere, P., Davidson, W. A., Murphy, D., and Nordmann, E. (2011). Numerical estimation in deaf and hearing adults. *Learn. Individ. Differ.* 21, 453–457. doi: 10.1016/j.lindif.2011.02.001

Burr, D. C., Turi, M., and Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *J. Vis.* 10:20. doi: 10.1167/10.6.20

Butterworth, B. (2005). "Developmental dyscalculia," in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (New York: Psychology Press), 455–467.

Cantlon, J., and Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychol. Sci.* 17, 401–406. doi: 10.1111/j.1467-9280.2006.01719.x

Chen, Q., and Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: a meta-analysis. *Acta Psychol.* 148, 163–172. doi: 10.1016/j.actpsy.2014.01.016

Clayton, S., and Gilmore, C. (2015). Inhibition in dot comparison tasks. *ZDM* 47, 759–770. doi: 10.1007/s11858-014-0655-2

Clayton, S., Gilmore, C., and Inglis, M. (2015). Dot comparison stimuli are not all alike: the effect of different visual controls on ANS measurement. *Acta Psychol.* 161, 177–184. doi: 10.1016/j.actpsy.2015.09.007

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., and Morgan, M. J. (2011). A common visual metric for approximate number and density. *P. Natl. Acad. Sci.* 108, 19552–19557. doi: 10.1073/pnas.1113195108

Daza, M. T., and Phillips-Silver, J. (2013). Development of attention networks in deaf children: support for the integrative hypothesis. *Res. Dev. Disabil.* 34, 2661–2668. doi: 10.1016/j.ridd.2015.05.012

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends Cogn. Sci.* 7, 145–147. doi: 10.1016/S1364-6613(03)00055-X

Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cogn. Neuropsychol.* 20, 487–506. doi: 10.1080/02643290244000239

DeWind, N. K., Adams, G. K., Platt, M. L., and Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition* 142, 247–265. doi: 10.1016/j.cognition.2015.05.016

Dewind, N. K., and Brannon, E. M. (2016). Significant inter-test reliability across approximate number system assessments. *Front. Psychol.* 7:310. doi: 10.3389/fpsyg.2016.00310

Dong, Y. Y., Zhou, R. L., and Gao, X. (2014). The effects of positive emotional arousal on spatial working memory. *Chin. J. Clin. Psychol.* 22, 761–767. doi: 10.16128/j.cnki.1005-3611.2014.05.047

Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon identification of a target letter in a non-search task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267

Fan, J., Gu, X., Guise, K. G., Liu, X., Fossella, J., Wang, H., et al. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain Cogn.* 70, 209–220. doi: 10.1016/j.bandc.2009.02.002

Fan, J., McCandliss, B. D., Sommer, T., Raz, M., and Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. *J. Cogn. Neurosci.* 14, 340–347. doi: 10.1162/089892902317361886

Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002

Feigenson, L., Libertus, M. E., and Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Dev. Perspect.* 7, 74–79. doi: 10.1111/cdep.12019

Gallistel, C. R., and Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition* 44, 43–74. doi: 10.1016/0010-0277(92)90050-R

Geary, D. C., Hoard, M. K., Nugent, L., and Bryd, C. J. (2008). Development of number line representations in children with mathematical learning disability. *Dev. Neuropsychol.* 33, 277–299. doi: 10.1080/87565640801982361

Gebuis, T., and Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behav. Res. Methods* 43, 981–986. doi: 10.3758/s13428-011-0097-5

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., et al. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One* 8:e67374. doi: 10.1371/journal.pone.0067374

Halberda, J. (2019). Perceptual input is not conceptual content. *Trends Cogn. Sci.* 23, 636–638. doi: 10.1016/j.tics.2019.05.007

Halberda, J., and 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. *Dev. Psychol.* 44, 1457–1465. doi: 10.1037/a0012682

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., and Germine, L. (2012). Number sense across the lifespan as revealed by a massive internet-based sample. *P. Natl. Acad. Sci.* 109, 11116–11120. doi: 10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–669. doi: 10.1038/nature07246

Izard, V. R., Sann, C., Spelke, E. S., and Streri, A. (2009). Newborn infants perceive abstract numbers. *P. Natl. Acad. Sci. U.S.A* 106, 10382–10385. doi: 10.1073/pnas.0812142106

Kanjlia, S., Feigenson, L., and Bedny, M. (2018). Numerical cognition is resilient to dramatic changes in early sensory experience. *Cognition* 179, 111–120. doi: 10.1016/j.cognition.2018.06.004

Klibanoff, R. S., Levine, S. C., Huttenlocher, J., Vasilyeva, M., and Hedges, L. V. (2006). Preschool children's mathematical knowledge: the effect of teacher "math talk". *Dev. Psychol.* 42, 59–69. doi: 10.1037/0012-1649.42.1.59

Kritzer, K. L. (2009a). Barely started and already left behind: a descriptive analysis of the mathematics ability demonstrated by young deaf children. *J. Deaf. Stud. Deaf. Edu.* 14, 409–421. doi: 10.1093/deafed/enp015

Kritzer, K. L. (2009b). Families with young deaf children and the mediation of mathematically based concepts within a naturalistic environment. *Am. Ann. Deaf* 153, 474–483. doi: 10.1353/aad.0.0067

LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., et al. (2010). Pathways to mathematics: longitudinal predictors of performance. *Child Dev.* 81, 1753–1767. doi: 10.1111/j.1467-8624.2010.01508.x

Levine, S. C., Suriyakham, L. W., Rowe, M. L., Huttenlocher, J., and Gunderson, E. A. (2010). What counts in the development of young children's number knowledge? *Dev. Psychol.* 46, 1309–1319. doi: 10.1037/a0019671

Libertus, M. E., Feigenson, L., and Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Dev. Sci.* 14, 1292–1300. doi: 10.1111/j.1467-7687.2011.01080.x

Libertus, M. E., Feigenson, L., and Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learn. Individ. Differ.* 25, 126–133. doi: 10.1016/j.lindif.2013.02.001

Libertus, M. E., Feigenson, L., Halberda, J., and Landau, B. (2014). Understanding the mapping between numerical approximation and number words: evidence from Williams syndrome and typical development. *Dev. Sci.* 17, 905–919. doi: 10.1111/desc.12154

Libertus, M. E., Forsman, L., Adén, U., and Hellgren, K. (2017). Deficits in approximate number system acuity and mathematical abilities in 6.5-year-old children born extremely preterm. *Front. Psychol.* 8:1175. doi: 10.3389/fpsyg.2017.01175

López-Crespo, G., Daza, M. T., and Méndez-López, M. (2012). Visual working memory in deaf children with diverse communication modes: improvement by differential outcomes. *Res. Dev. Disabil.* 33, 362–368. doi: 10.1016/j.ridd.2011.10.022

Madalena, S. P., Santos, F. H., Silva, K. D., and Marins, M. (2015). "Arithmetical abilities in Brazilian deaf signers of elementary school," in *Proceedings of the 22nd International Congress on the Education of the Deaf*, (Athens).

Mazzocco, M. M. M., Feigenson, L., and Halberda, J. (2011a). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Dev.* 82, 1224–1237. doi: 10.1111/j.1467-8624.2011.01608.x

Mazzocco, M. M. M., Feigenson, L., and Halberda, J. (2011b). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS One* 6:e23749. doi: 10.1371/journal.pone.0023749

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

Mitchell, R. E. (2008). *Academic Achievement of Deaf Students*. Washington, DC: Gallaudet University Press.

Norris, J. E., and Castronovo, J. (2016). Dot display affects approximate number system acuity and relationships with mathematical achievement and inhibitory control. *PLoS ONE* 11:e0155543. doi: 10.1371/journal.pone.0155543

Nunes, T. (2004). *Teaching Mathematics to Deaf Children*. Oxford: Oxford University Press.

Odic, D., Hock, H., and Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *J. Exp. Psychol. Gen.* 143, 255–265. doi: 10.1037/a0030825

Odic, D., Libertus, M. E., Feigenson, L., and Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Dev. Psychol.* 49, 1103–1112. doi: 10.1037/a0029472

Pagliaro, C. M., and Kritzer, K. L. (2010). Learning to learn: an analysis of early learning behaviors demonstrated by young deaf/hard-of-hearing children with high/low mathematics ability. *Deaf. Educ. Int.* 12, 54–76. doi: 10.1179/146431510x12626982043723

Pailian, H., Libertus, M. E., Feigenson, L., and Halberda, J. (2016). Visual working memory capacity increases between ages 3 and 8 years, controlling for gains in attention, perception, and executive control. *Atten. Percept. Psychophys.* 78, 1556–1573. doi: 10.3758/s13414-016-1140-5

Paul, J. M., Reeve, R. A., and Forte, J. D. (2017). Taking a(c)count of eye movements: multiple mechanisms underlie fixations during enumeration. *J. Vis.* 17:16. doi: 10.1167/17.3.16

Peng, P., Yang, X., and Meng, X. (2017). The relationship between approximate number system and early arithmetic: the mediation role of numerical knowledge. *J. Exp. Child Psychol.* 157, 111–124. doi: 10.1016/j.jecp.2016.12.011

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., and Lucangeli, D. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition* 116, 33–41. doi: 10.1016/j.cognition.2010.03.012

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44, 547–555. doi: 10.1016/j.neuron.2004.10.014

Pica, P., Lemer, C., Izard, V., and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503. doi: 10.1126/science.1102085

Pixner, S., Leyrer, M., and Moeller, K. (2014). Number processing and arithmetic skills in children with cochlear implants. *Front. Psychol.* 5:1479. doi: 10.3389/fpsyg.2014.01479

Reyna, V. F., Nelson, W. L., Han, P. K., and Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychol. Bull.* 135, 943–973. doi: 10.1037/a0017327

Ritchie, S. J., and Bates, T. C. (2013). Enduring links from childhood mathematics and reading achievement to adult socioeconomic status. *Psychol. Sci.* 24, 1301–1308. doi: 10.1177/0956797612466268

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *J. Hum. Resour.* 27, 313–328. doi: 10.2307/145737

Rohde, T. E., and Thompson, L. A. (2007). Predicting academic achievement with cognitive ability. *Intelligence* 35, 83–92. doi: 10.1016/j.intell.2006.05.004

Rueda, M. R., Fan, J., Mccandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan, S. S., Stricker, J., et al. (2016). Associations of non-symbolic and symbolic numerical magnitudeprocessing with mathematical competence: a meta-analysis. *Dev. Sci.* 1–16. doi: 10.1111/desc.12372

Skagerlund, K., and Träff, U. (2016). Number processing and heterogeneity of developmental dyscalculia: subtypes with different cognitive profiles and deficits. *J. Learn. Disabil.* 49, 36–50. doi: 10.1177/0022219414522707

Smets, K., Sasanguie, D., Szücs, D., and Reynvoet, B. (2015). The effect of different methods to construct non-symbolic stimuli in numerosity estimation and comparison. *J. Cogn. Psychol.* 27, 310–325. doi: 10.1080/20445911.2014.996568

Starr, A., Libertus, M. E., and Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *P. Natl. Acad. Sci. U.S.A.* 110, 18116–18120. doi: 10.1073/pnas.1302751110

Starr, A., Dewind, N. K., and Brannon, E. M. (2017). The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition.* 168, 222–233. doi: 10.1016/j.cognition.2017.07.004

Swanwick, R., Oddy, A., and Roper, T. (2005). Mathematics and deaf children: an exploration of barriers to success. *Deaf. Educ. Int.* 7, 1–21. doi: 10.1179/146431505790560446

Szucs, D., Devine, A., Soltesz, F., Nobes, A., and Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex* 49, 2674–2688. doi: 10.1016/j.cortex.2013.06.007

Titus, J. C. (1995). The concept of fractional number among deaf and hard of hearing students. *Am. Ann. Deaf.* 140, 255–262. doi: 10.1353/aad.2012.0582

Traxler, C. B. (2000). The Stanford Achievement Test, 9th edition: national norming and performance standards for deaf and hard-of-hearing students. *J. Deaf. Stud. Deaf. Edu.* 5, 337–348. doi: 10.1093/deafed/5.4.337

Wang, D., Di, M., and Qian, M. (2007). A Report on the third revision of Combined Raven's Test (CRT-C3) for children in China. *Chin. J. Clin. Psychol.* 15, 559–568. doi: 10.16128/j.cnki.1005-3611.2007.06.008

Wang, J. J., Halberda, J., and Feigenson, L. (2017). Approximate number sense correlates with math performance in gifted adolescents. *Acta Psychol.* 176, 78–84. doi: 10.1016/j.actpsy.2017.03.014

Wu, H. R., and Li, L. (2005). Development of Chinese rating scale of pupil's mathematic abilities and study on its reliability and validity. *Chin. J. Public Health* 21, 473–475. doi: 10.11847/zgggws2005-21-04-58

Xu, F., and Spelke, E. S. (2000). Large number discrimination in 6-month old infants. *Cognition* 74, B1–B11. doi: 10.1016/S0010-0277(99)00066-9

Zhang, Y., Liu, T., Chen, C., and Zhou, X. (2019). Visual form perception supports approximate number system acuity and arithmetic fluency. *Learn. Indiv. Differ.* 71, 1–12. doi: 10.1016/j.lindif.2019.02.008

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Human Neuroscience

ORIGINAL RESEARCH
published: 21 June 2021
doi: 10.3389/fnhum.2021.687321

# Groupitizing Improves Estimation of Numerosity of Auditory Sequences

**Giovanni Anobile[1]*, Elisa Castaldi[1,2], Paula A. Maldonado Moscoso[1], Roberto Arrighi[1] and David Burr[1]**

[1] Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy,
[2] Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy

Groupitizing is a recently described phenomenon of numerosity perception where clustering items of a set into smaller "subitizable" groups improves discrimination. Groupitizing is thought to be rooted on the subitizing system, with which it shares several properties: both phenomena accelerate counting and decrease estimation thresholds irrespective of stimulus format (for both simultaneous and sequential numerosity perception) and both rely on attention. As previous research on groupitizing has been almost completely limited to vision, the current study investigates whether it generalizes to other sensory modalities. Participants estimated the numerosity of a series of tones clustered either by proximity in time or by similarity in frequency. We found that compared with unstructured tone sequences, grouping lowered auditory estimation thresholds by up to 20%. The groupitizing advantage was similar across different grouping conditions, temporal proximity and tone frequency similarity. These results mirror the groupitizing effect for visual stimuli, suggesting that, like subitizing, groupitizing is an a-modal phenomenon.

**Keywords: approximate number system, groupitizing, auditory numerosity, calculation, numerosity perception, subitizing**

***Edited by:***
*Giuseppe Giglia,*
*University of Palermo, Italy*

***Reviewed by:***
*Takayuki Kawashima,*
*Teikyo Heisei University, Japan*
*Rosa Rugani,*
*University of Padua, Italy*

****Correspondence:***
*Giovanni Anobile*
*giovanni.anobile@unifi.it*

***Specialty section:***
*This article was submitted to Cognitive Neuroscience, a section of the journal Frontiers in Human Neuroscience*

***Received:*** *29 March 2021*
***Accepted:*** *28 May 2021*
***Published:*** *21 June 2021*

***Citation:***
*Anobile G, Castaldi E, Maldonado Moscoso PA, Arrighi R and Burr D (2021) Groupitizing Improves Estimation of Numerosity of Auditory Sequences. Front. Hum. Neurosci. 15:687321. doi: 10.3389/fnhum.2021.687321*

## INTRODUCTION

Humans exploit various strategies to gauge the number of objects in a set, including serial counting and approximate estimation. Although estimation is relatively fast, it is prone to errors, with response variability (standard deviation of the estimates) tending to scale linearly with the number of objects (Weber Law) (Whalen et al., 1999; Ross, 2003). Interestingly, both serial counting and estimation change characteristics when the set of items is small–between 1 and 4 objects–a range known as *subitizing* (Kaufman et al., 1949). Numerosity judgements within the subitizing range violates Weber law, as people usually do not make estimation errors even when stimuli are presented for just a few milliseconds (Revkin et al., 2008; Choo and Franconeri, 2014). Similarly, serial counting response times are fast and constant within the subitizing range, then steadily increase for higher numerosities, with a clear performance discontinuity around 4 or 5 items (Kaufman et al., 1949). Subitizing was first reported by Jevons (1871), and has since been observed in numerous studies, making it one of the most robust and widely described phenomena in the numerosity literature. The subitizing phenomenon is neither restricted to arrays of items presented simultaneously over a given region of space nor to vision. Indeed, subitizing has been reported for haptic spatial arrays, and for sequences of visual, and auditory stimuli (Riggs et al., 2006;

Repp, 2007; Camos and Tillmann, 2008; Gallace et al., 2008; Plaisier et al., 2009, 2010; Ferrand et al., 2010; Plaisier and Smeets, 2011; Anobile et al., 2019).

Recent studies have described a new phenomenon in numerosity perception, termed *groupitizing*, which shares many characteristics with the subitizing phenomenon. Groupitizing can be defined as "the ability to capitalize on grouping information to facilitate enumeration processes" (Starkey and McCandliss, 2014). When an array of more than ∼4 objects (above the subitizing limit) is spatially clustered into sub-groups each containing few items, with both number of groups and items per group falling within the subitizing range, the counting speed robustly increases compared with unstructured arrays (Wender and Rothkegel, 2000; Starkey and McCandliss, 2014). Signatures of grouping strategies in numerosity perception have also been observed in young chicks. Birds spontaneously prefer arrays grouped into clusters (defined by colors and shapes) containing the same number of items (Loconsole et al., 2021). Discrimination also improves when objects are presented within groups (Rugani et al., 2017).

While groupitizing has been studied much less than subitizing, the advantage in numerosity processing appears to be consistent and robust. For example, counting speed increases for objects randomly scattered over a given space but grouped by color proximity (Ciccione and Dehaene, 2020). Groupitizing also lowers perceptual thresholds (as defined by the normalized standard deviation of estimations) for approximate numerosity estimation of briefly presented stimuli (Anobile et al., 2020): clustering dot arrays into separate groups by spatial or color proximity leads up to 20% improvement in the precision of numerosity estimates. The groupitizing advantage was not restricted to spatial numerosity (items presented simultaneously) but also generalized to temporal numerosity. For example, Anobile et al. (2020) presented sequences of flashes that were all colored the same ("unstructured condition") or grouped by color proximity (e.g., two red, two yellow, two blue). Estimation errors followed Weber's law in both conditions, suggesting that participants did not count the items but estimated their numerosity approximatively. Most importantly, sensory precision was again improved by groupitizing up to about 15%.

Both subitizing and groupitizing share a similar reliance on attentional resources. When participants were asked to estimate the numerosity of dot arrays within the subitizing range under a condition of attentional deprivation (dual-task paradigm), the classical subitizing advantage on sensory thresholds completely vanished, with precision thresholds increasing to match those of higher numerosities (Vetter et al., 2008; Anobile et al., 2012, 2019). Similarly depriving attentional resources via a concurrent visual dual task induced significant detrimental effects on sensory thresholds for grouped arrays relatively to unstructured arrays (Maldonado Moscoso et al., 2020), suggesting that like subitizing, groupitizing relies on attentional resources.

While groupitizing has been demonstrated across different formats (spatial arrays and temporal sequences), for both counting and estimations tasks, it has mainly been studied within the visual domain. The only study (to the best of our knowledge) that has investigated the effect of stimuli grouping in another sensory modality (audition) reported increased accuracy for regular sequences of sounds organized in small equal groups (structured sequences) relative to unstructured sequences (Hoopen and Vos, 1979b). The results showed that grouping stimuli (with elements in a group not exceeding 5) improved numerical estimation accuracy, but only for short ISIs (Hoopen and Vos, 1979b). However, these results were based on error rates, an index that confounds precision and bias, and does not consider error magnitude, and are therefore difficult to relate to modern studies showing perceptual advantages of groupitizing.

The aim of the present study is to examine whether groupitizing is a general phenomenon that occurs in senses other than vision: specifically in audition. We devised an experimental paradigm in which auditory stimuli were grouped in two different ways, to mimic as much as possible grouping cues exploited in previous visual experiments. Participants estimated the numerosity of a series of tones clustered either by proximity in time (mimicking the visual grouping by spatial cues) or by similarity in frequency (mimicking the visual grouping by color). The hypothesis is straightforward: if groupitizing is at least partially rooted into the subitizing system we expect to observe lower numerosity estimation thresholds when auditory groupitizing is facilitated.

## MATERIALS AND METHODS

### Participants

Fourteen adults participated in the study: four authors (GA, EC, RA, PM) and ten naïve students from the School of Psychology of Florence with little or no experience of psychophysical experiments (mean age = 29 years, standard deviation = 6 years, range = 19–45 years). The sample size was based on previous studies on groupitizing (Anobile et al., 2020; Maldonado Moscoso et al., 2020), all participants had normal or corrected-to-normal vision, and none had mathematical or other learning disorders, nor did any have substantial musical training. The research was approved by the local ethics committee ("*Commissione per l'Etica della Ricerca,*" University of Florence, July 7, 2020, n. 111) and informed consent was obtained from all participants prior to testing.

### General Procedure

Stimuli were generated and presented with PsychToolbox routines (Kleiner et al., 2007) for Matlab (ver. R2017b, The Mathworks, Inc.[1]). Participants sat 57 cm from a 15″ screen monitor (60 Hz), in a quiet and dimly light room. Stimuli were temporal sequences of 50 ms pure tones ramped on and off with 20-ms raised cosine ramps, presented with an intensity of 80 dB (at the sound source) and digitized at a sample frequency of 8192 Hz. Sounds were presented through high-quality headphones (Microsoft LifeChat LX-3000).

Each trail started with the participant observing a gray blank screen on which appeared a red central fixation point (2 deg of diameter). After 200 ms, a sequence of sounds was played,

---

[1] http://mathworks.com

with the fixation point that remained visible during the whole presentation. At the end of the auditory sequence, the fixation point turned green, to prompt the participants to provide a response (**Figure 1A**). On each trial, participants were asked to verbally report the number of perceived sounds, which was recorded by the experimenter via a computer keyboard. There was no time pressure on responses, but participants were asked to respond as accurately as possible. Each sequence could contain 5–12 tones, and participants were informed about the numerosity range. As the interval between the first and last sounds was always kept constant, each sequence lasted 1.4 s independently of the number of tones. As a consequence, numerosity correlates with temporal frequencies ranging from 3.5 Hz (for numerosity 5) to 8.5 Hz (for numerosity 12). As the purpose of this study was to investigate approximate numerosity estimation and not serial counting, this frequency range was chosen based on previous studies (Anobile et al., 2018, 2020) showing that in these regimes participants cannot serially count the items one-by-one, but they have to rely on approximate estimations (obeying Weber's law).

The experiment comprised three main conditions (tested in separate sessions) in which sound sequences were manipulated to either facilitate perceptual grouping or not (details in the stimuli section). Participants (except the four authors) were not informed about the experimental conditions and were left free to choose the best strategy to solve the task. For each condition, the testing phase was preceded by a familiarization session of 22 trials (not included in the analyses). During familiarization, all numerosities were randomly presented without feedback. After the familiarization phase, the testing phase started. For each of the three experimental conditions, each participant performed around 25 trials for each numerosity (for a total of 8338 data points across all the experiments and participants). The sessions order was randomized across participants and participants had a break of ∼10 min after each session.

### Auditory Stimuli

Participants were tested in three different conditions: (1) unstructured sequence of tones, (2) sequences grouped by tone frequency, or (3) sequences grouped by inter stimuli interval (ISI) (**Figure 1B**).

The unstructured sequences were built in two steps. On each trial, the whole sequence was divided into regular intervals (total duration/numerosity), with all consecutive pair of sounds demarking an identical ISI. The ISIs of these regular patterns for each numerosity were: N5 = 287 ms, N6 = 220 ms, N7 = 175 ms, N8 = 142 ms, N9 = 118 ms, N10 = 100 ms, N11 = 85 ms; N12 = 72 ms (average = 150 ms, SD = 73 ms). Then to reduce the regularity of the tone sequences, a small temporal jitter (around 10% of the ISI for regular patterns of that numerosity) was applied to the timing of each tone (excluding the first and the last), by increasing or reducing the ISI between two consecutive impulses (sign of the perturbation randomly selected trial by trial for each tone). On any given trial, all tones were defined by an identical frequency randomly selected out of three possibilities: 400, 700, or 1000 Hz.

The temporal structures of the sequences grouped by tone frequency were identical to those used for the unstructured stimuli, except for the frequency of the tones in the sequence: the tones were divided into groups of impulses of identical frequency. Each group of tones had frequencies of 400, 700, or 1000 Hz. The sequence clustering followed the groupitizing rules: the total sequence was divided into two or three groups, each containing two to four tones (see **Figure 1B** for an example of numerosity nine clustered into three groups of three tones each). Each numerosity yielded a given number of possible clusters: N5 (2 + 3 or 3 + 2), N6 (3 + 3 or 2 + 2 + 2), N7 (3 + 2 + 2 or 2 + 2 + 3 or 3 + 4), N8 (4 + 4 or 2 + 2 + 2 + 2), N9 (4 + 3 + 2 or 3 + 3 + 3), N10 (4 + 4 + 2 or 3 + 4 + 3), N11 (4 + 4 + 3), N12 (4 + 4 + 4). On every trial, for the selected numerosity, we randomly selected one of the possible patterns (e.g., for N = 8 the choice was between four groups of two tones or two groups of four tones). Finally, to limit the possibility of solving the task by simply memorizing the correspondence between a given numerosity and a sequence of sound frequencies, we arbitrarily defined up to six different frequency configurations for each numerosity. For example, on each trial in which numerosity "six" was presented, the frequency of the sounds in the sequence was defined by one of the following pattern: [1,000, 10,00, 700, 700, 400, 400] or [1,000, 1,000, 400, 400, 700, 700] or [700, 700, 1000, 1000, 400, 400] or [700, 700, 700, 400, 400, 400] or [400, 400, 400, 1,000, 1,000, 1,000], or [1,000, 1,000, 1,000, 700, 700, 700] Hz.

The sequences grouped by inter-stimulus interval (ISI) were also built in two steps. First the whole sequence (1.4 s) was divided into 12 identical intervals (with 12 corresponding to the highest numerosity tested). Then some of the slots were selected to create temporally separate tone clusters (see **Figure 1B** for an example of numerosity nine clustered into three groups of three tones each). In this condition, we did not apply any temporal jitter to the sequences. The tone clusters were created according to the groupitizing rules: 2, 3, or 4 groups each containing few items (from 1 to 5). The only exception was the numerosity eleven that was created by a group of 5 and a group of 6 tones. The ISI between groups ranged between 140 to 942 ms. To keep the conditions balanced, the numerosity12 was played but as no clustering could have been applied (all slots in the sequence were used), this numerosity was eliminated from the analyses. The temporal clusters were: N5 (2 + 3 or 2 + 1 + 2), N6 (3 + 3 or 2 + 2 + 2), N7 (2 + 3 + 2 or 3 + 1 + 3), N8 (4 + 4 or 2 + 2 + 2 + 2), N9 (2 + 3 + 4 or 3 + 3 + 3), N10 (5 + 5 or 4 + 2 + 4), N11 (5 + 6), N12 (no clusters). On every trial, for each numerosity, we randomly selected one of the two possible patterns (e.g., for N8 four groups of two tones or two groups of four tones). On each trial, all the tones in the sequence were defined by three possible frequencies: 400, 700, or 1,000 Hz.

## Data Analyses

We first checked for response outliers. Separately for each participant, condition and numerosity, we eliminated trials below or above 3 SD of accuracy or response time (∼2% of the trials for each condition for a total of 190 trials). Given that in the ISI condition only numerosities from 5 to 11 provided clustering cues, numerosity twelve was not included in the analyses. For each participant, numerosity and condition we separately calculated the average perceived numerosity and the standard

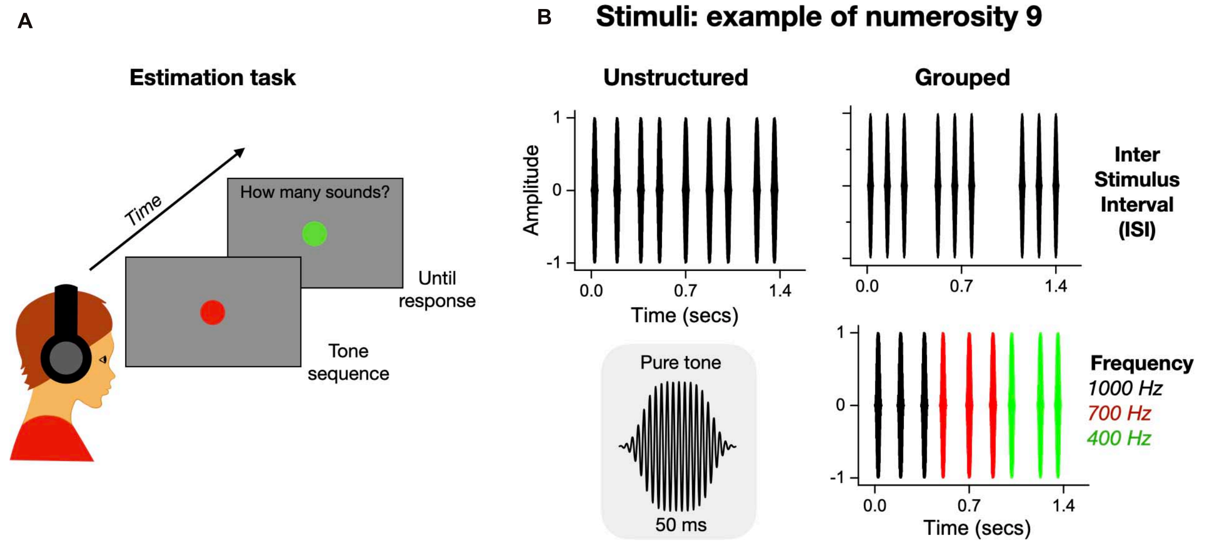


**FIGURE 1 |** Illustration of the procedure and stimuli. **(A)** Illustration of the numerosity estimation task. Participants kept gaze on a red central fixation point while a sequence of tones was played for 1.4 s. After the auditory stimulus had finished, the fixation point color changed from red to green, signaling to verbally report the perceived numerosity. **(B)** Example of auditory stimuli for numerosity nine in the three experimental conditions: unstructured, grouped by ISI and grouped by frequency. The gray insert shows the waveform of a single pure tone.

deviation of the responses. Sensory precision was measured by normalizing the standard deviation by the physical numerosity to obtain a Coefficient of variation (CV), a dimensionless index of precision that allows comparison and averaging of performance across different numerosities.

$$CV = \frac{\sigma}{N} \tag{1}$$

where $N$ is numerosity and σ standard deviation of responses to that numerosity. The percentage of advantage of the CVs in the grouping compared with unstructured condition was indexed as the percent improvement:

$$Grouping\ advantage\,(\%) = \frac{CVu - CVg}{CVu} \times 100 \tag{2}$$

Where $CV_u$ and $CV_g$ are the Coefficients of variation for the unstructured and grouped conditions, respectively.

Data were analyzed by repeated measures ANOVA (3 conditions × 7 numerosities) and *post-hoc* *t*-tests. *P*-values (two-tailed) were corrected for multiple comparisons with the Bonferroni method ($p_{\text{bonf}}$). Effect sizes associated with ANOVA were reported as $\eta^2$, and those associated with *post-hoc* *t*-tests as Cohen's d. *T*-tests were supplemented with Bayesian statistics, calculating Bayes Factors, the ratio of the likelihood of the alternative to the null hypothesis, and reporting them as base 10 logarithms. By convention, LogBF $>$ 0.5 is considered substantial evidence in favor of the alternative hypothesis and LogBF $< -0.5$ substantial evidence for the null hypothesis. Absolute values greater than 1 are considered strong evidence, and greater than 2 definitive. Data were analyses using JASP (ver. 0.8.6 2018) and Matlab (ver. R2017b) software.

# RESULTS

## Effect of Auditory Groupitizing on Perceived Numerosity

We first evaluated the effect of grouping on perceived numerosity. **Figure 2** shows average responses separately for the three experimental conditions, as a function of physical numerosity. To statistically test differences across conditions, we performed a repeated measures ANOVA with numerosity (7 levels, from N5 to N11) and grouping condition (3 levels) as within subject factors. The main effect of numerosity was obviously significant [$F(6, 78) = 445$, $p < 0.001$, $\eta^2 = 0.97$], but there was no significant effect of "grouping condition" [$F(2, 26) = 2.09$, $p = 0.14$, $\eta^2 = 0.14$]. The condition-by-numerosity interaction was statistically significant [$F(12, 156) = 4.73$, $p < 0.001$, $\eta^2 = 0.26$]. To explore this interaction, we performed a series of *post-hoc* *t*-tests contrasting, for each numerosity, the responses in the unstructured condition against those for grouping by ISI or frequency separately. None of the numerosity estimates in the grouping conditions significantly differed from the unstructured condition after Bonferroni correction (min $p_{\text{bonf}} = 0.11$ for the contrast N5 unstructured Vs. N5 frequency, all the other $p_{\text{bonf}} > 0.42$; highestLogBF = 0.84 for the same comparison, all the other $-0.54 <$ LogBF $<$ 0.33). Overall, these results indicate that auditory grouping had no strong effect on average perceived numerosity of auditory stimuli.

## Effect of Auditory Groupitizing on Sensory Precision

Having verified that average perceived numerosity did not change with grouping, we focused on sensory precision, indexed by

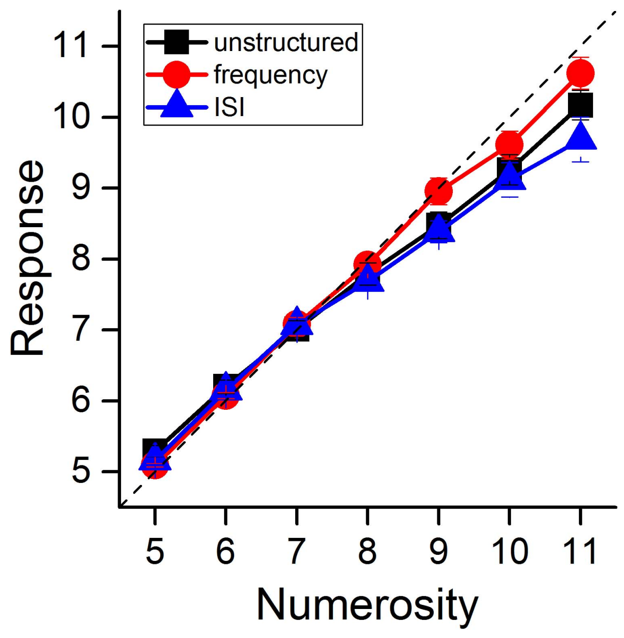


**FIGURE 2 |** Perceived auditory numerosity. Average perceived numerosity as a function of physical numerosity for the three experimental conditions (black squares: unstructured stimuli, red circles: stimuli grouped by frequency, blue triangles: stimuli grouped by ISI). Error bars are ± 1 SEM.

the Coefficient of variation (standard deviations normalized by numerosity; see Materials and Methods). **Figure 3A** shows the average Coefficient of variation as a function of numerosity, for all three experimental conditions. It is evident on inspection that unstructured stimuli (black squares) yielded higher Coefficients of variations (less precision) than the two grouping conditions. **Figure 3B** shows the Coefficient of variation averaged across numerosities and participants for the unstructured (black) and the two grouping conditions (red: frequency, blue: ISI).

Repeated measure ANOVA with numerosity (7 levels, from N5 to N11) and grouping condition (3 levels) as within subject factors revealed a main effect of condition [$F(2, 26) = 7.83$, $p = 0.002$, $\eta^2 = 0.38$]. The factor numerosity was also statistically significant [$F(6, 78) = 8.6$, $p < 0.001$, $\eta^2 = 0.40$], while the condition-by-numerosity interaction was not [$F(12, 156) = 1.76$, $p = 0.06$, $\eta^2 = 0.12$]. *Post hoc t*-tests on conditions revealed that both grouping by frequency ($t = 3.4$, $p_{\mathrm{bonf}} = 0.014$, Cohen's $d = 0.9$, LogBF = 2.19) and by ISI ($t = 4.1$, $p_{\mathrm{bonf}} = 0.004$, Cohen's $d = 1.1$, LogBF = 4.9) significantly improved sensory precision compared to the unstructured condition. The two grouping conditions did not differ between each other ($t = 1.2$, $p_{\mathrm{bonf}} = 0.7$, Cohen's $d = 0.32$, LogBF = –0.52).

Although the condition-by-numerosity interaction in the ANOVA was not statistically significant, to test whether different strategies (such as counting) may have been used to solve at high and low numerosities, we further investigated the dependence on numerosity by dividing the data into high and low numerosities (greater or less than 7.5). The improvement with groupitizing was strong and significant for both ranges [N5–7: mean = 25%, $t_{(83)} = 4.8$, $p < 0.001$, Cohen's $d = 0.53$, LogBF = 3.2; N9–11: mean = 15%, $t_{(83)} = 3.59$, $p < 0.001$, Cohen's $d = 0.39$, LogBF = 1.6]. The size of the effect was statistically indistinguishable in the two numerical ranges [$t_{(83)} = 1.58$, $p = 0.12$, Cohen's $d = 0.17$, LogBF = –0.40].

**Figure 4A** shows the Coefficient of variation (CV) averaged across all numerosities for all participants, plotting CV measured in the two grouped conditions (frequency in red and ISI in blue) against that for the unstructured condition. Despite large inter-individual variability in thresholds, and in the improvement induced by grouping, the majority of data points fall below the equality line, indicating that most of the participants (with no obvious differences between naïve and authors, see filled and empty small data points in **Figure 4A**) estimated numerosity of the grouped stimuli with higher precision than the ungrouped. On average, grouping the stimuli by ISI improved precision by about 25% and grouping by frequency by 15% (with improvement defined by eqn. 2). These robust effects nicely compliment with those previously reported in vision for both, temporal sequences, or spatial arrays (improvement of about 15 and 20%, respectively) as shown by **Figure 4B**. An ANOVA on the grouping advantage across stimuli formats and grouping strategy revealed that the effects were statistically indistinguishable [$F(4, 75) = 0.88$, $p = 0.47$].

## DISCUSSION

The aim of this study was to investigate whether and to what extent groupitizing effects occur in audition. The results revealed that auditory grouping cues had no measurable effect on average perceived numerosity, but they decreased estimation thresholds by up to 20%, similar to the advantage previously reported for spatial arrays. The groupitizing advantage occurred for both grouping conditions, both when groups were defined by manipulating the temporal proximity of the tone, as well as when they were defined by similarity of tone frequency.

These results mirror what has been recently reported in the visual domain, both for arrays of stimuli presented simultaneously over a given region of space and for sequences of flashes (Anobile et al., 2020), suggesting that groupitizing may reflect the activity of one or more a-modal and cross-format systems. Most evidence suggests that groupitizing depends on subitizing, an attention-dependent mechanism for fast and accurate enumeration of small quantities, combined with arithmetical strategies. Participants probably parse the array into subitizable samples, which can be precisely enumerated by leveraging on the subitizing precision, and summed together to estimate total numerosity. In support to this hypothesis, Starkey and McCandliss (2014) showed that children with higher arithmetical abilities took greater advantage from groupitizing in a dot counting task. Ciccione and Dehaene (2020) further generalized these results to the adult population by showing a stronger groupitizing advantage for mathematics university students compared with humanities and psychology. And arithmetical abilities in adults are better correlated with

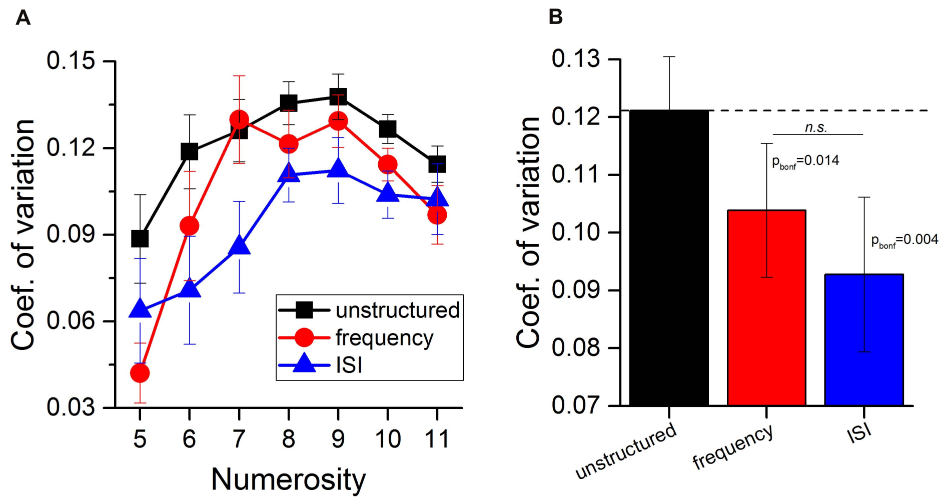


**FIGURE 3 |** Groupitizing affects precision of estimation of auditory stimuli. **(A)** Average Coefficient of variation as a function of numerosity for the three experimental conditions (black squares: unstructured stimuli, red circles: stimuli grouped by frequency, blue triangles: stimuli grouped by ISI). **(B)** Coefficients of variation averaged across numerosity levels and participants. Black Error bars show ± 1 SEM.

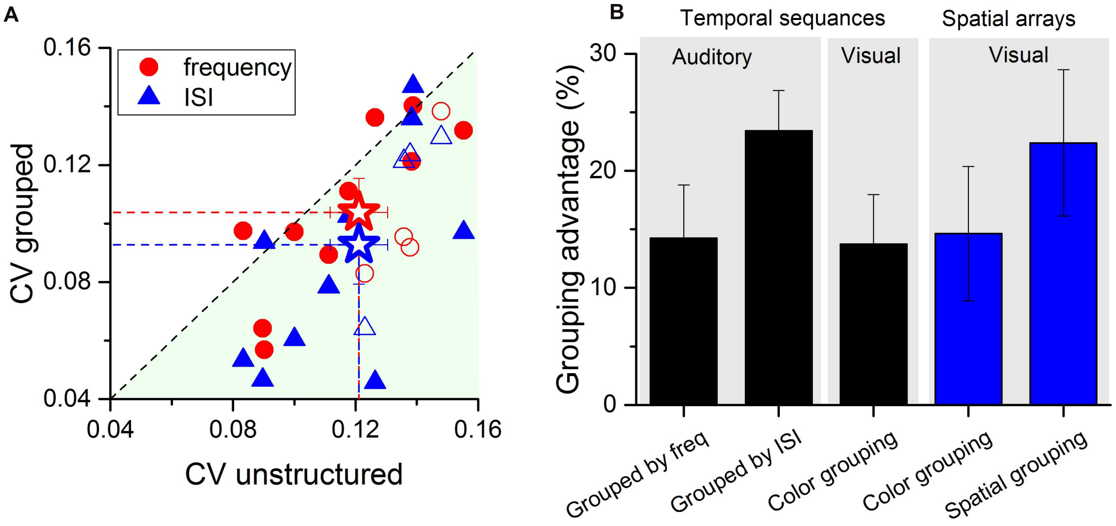


**FIGURE 4 |** Individual coefficients of variation for the three conditions. **(A)** Scatter plot of Coefficient of variation (CV) in the grouped conditions (red circles: stimuli grouped by frequency, blue triangles: stimuli grouped by ISI) plotted against those measured in the unstructured condition. For the grouping by ISI, the average CV was 0.09 (blue star and dashed line), for the grouping by frequency was 0.10 (red star and dashed line), both lower than the average CV in the unstructured condition (0.12). For almost all participants (naïve filled circles and triangles, authors open circles and triangles) CVs for grouped stimuli were lower than those for unstructured stimuli. Error bars are ± 1 SEM. **(B)** Groupitizing advantage on sensory precision across stimuli formats and sensory modalities. The first two bars report the grouping advantage for auditory stimuli (current study) grouped by frequency or by ISI (compared with unstructured stimuli). The other data show results from a previous study investigating groupitizing effects in vision (Anobile et al., 2020). Data are publicly available at Anobile et al. (2020). From left to right: grouping temporal sequences by color; grouping spatial arrays by color; grouping spatial arrays by spatial proximity. Error bars show ± 1 SEM.

numerosity thresholds for grouped than ungrouped stimuli (Maldonado Moscoso et al., 2020).

(Hoopen and Vos, 1979a,b) looked at the effect of grouping of tones on perception, initially to study attentional switching. They found that at some ISIs, grouping caused underestimation of numerosity (Hoopen and Vos, 1979b), which we did not observe here. However, at ISIs compatible with those of this study, they also reported no underestimation in numerosity. They further found that for short ISIs, where counting was not possible, accuracy improved in the grouped condition. Although differences in the experimental procedures (such as using regular rather than randomized ISIs and that their measure of accuracy confounds bias and precision) make it difficult to relate in detail their study with the current study, their findings are broadly consistent with those reported here.

In our study the interstimulus intervals were generally short, making it difficult to count the stimuli: on debriefing, all participants reported that they guessed at the numerosity, and did not attempt to count them (although this was not expressly forbidden). If counting were possible, it would have occurred for the lower rather than the higher numerosities, as total stimulus duration was constant (1.4 s), and ISI varied accordingly, from 287 ms for $N = 5$–85 ms for $N = 11$. However, when we separated the data into high and low numerosities (greater or less than 7.5), we found that both ranges showed strong and highly significant groupitizing effects, with no significant difference between the two ranges. We therefore conclude that the results are unlikely to reflect counting strategies.

Over the last few years there has been increasing interest in the association between numerosity perception and mathematics. A considerable body of evidence suggests that numerosity perception may represent an early non-symbolic foundational capacity for the development of symbolic arithmetic skills (Halberda et al., 2008; Piazza, 2010; Chen and Li, 2014; Fazio et al., 2014; Schneider et al., 2017). However, despite much evidence supporting this fascinating idea, many studies have failed to find significant correlations, or causal training effects between numerosity and arithmetic (De Smedt et al., 2013; Sasanguie et al., 2014; Caviola et al., 2020; Bugden et al., 2021). The literature on this topic is contradictory, and the reasons for reported failures in correlations between numerosity perception and arithmetic are still largely unclear and debated.

One possibility is that groupitizing is the link between numerosity perception and math: people with strong arithmetic skills may take advantage of natural clustering in random arrays and use a combination of grouping and arithmetical strategies to solve the numerosity task. This in turn could drive (even partially) the correlation with math scores. A recent study found that visual and auditory subitizing capacities do not correlate with mental calculation abilities (Anobile et al., 2019). Similarly, numerosity discrimination thresholds for very high numerosity do not correlate with arithmetical abilities (Anobile et al., 2016). The fact that arithmetical abilities correlate only with intermediate numerosities (Burr et al., 2017) might be because these numerosities are ideal for groupitizing. Numerosities within the subitizing range are (by definition) immediately and accurately perceived holistically, with no need to apply arithmetic strategies to combine different subsets. On the other hand, very high numerosities might be difficult to segment and cluster into a small (subitizable) number of subgroups. Furthermore, numerosity discrimination thresholds in the estimation range predict arithmetical abilities in primary school children for spatial arrays (dots), but not for auditory or visual sequences (Anobile et al., 2018). This could reflect lower natural clustering for temporal sequences compared with spatial arrays, or the existence of multiple systems with different relationships with the development of formal arithmetic. Future research should investigate whether auditory groupitizing relates to arithmetical abilities to the same extent as visuo-spatial groupitizing does.

Clinical research may also contribute to clarifying whether the link between numerosity perception and arithmetical skills is mediated by groupitizing. Groupitizing could be studied in developmental dyscalculia, and the effectiveness of training programs promoting the use of groupitizing strategies (inducing mental arithmetical procedures), rather than generally boosting numerosity discrimination *per sè*, should be quantitatively evaluated.

Subitizing limits for auditory sequences are thought to be lower than those for spatial vision (Repp, 2007; Anobile et al., 2019), possibly because the stimuli are one-dimensional rather than two-dimensional. Indeed some studies suggest that the limit could be as low as three (Repp, 2007), while the visual limit is usually considered to be four (Jevons, 1871; Kaufman et al., 1949; Atkinson et al., 1976). However, there is no sharp cutoff for subitizing, and the limit depends somewhat on definition. For example, Piazza et al. (2011) define the limit operationally by fitting estimation errors with a Gaussian error function, and taking the 50% point as the numerosity limit. With this definition, the visual limit for spatial subitizing is around six while the auditory sequential limit is five (Anobile et al., 2019). So, while the limit for auditory sequences is probably slightly less than that for visual sequences, it is likely that the participants were able to subitize reasonably well even the longest clusters of four items. This is supported by the fact that the average groupitizing effects for auditory sequences were very comparable with those previously reported for spatial arrays (∼20% see **Figure 4B**).

In addition to the average values being different, the limits for spatial and temporal subitizing do not correlate with each other, suggesting separate systems (Anobile et al., 2019). It is therefore possible that auditory and spatial visual groupitizing are subserved by different mechanisms, but this issue would need to be specifically addressed in future studies. Research is also needed to explore the brain networks underlying groupitizing and how they relate to those supporting subitizing (Piazza et al., 2002; Ansari et al., 2007; He et al., 2014) and arithmetic calculation (Castaldi et al., 2020).

Counting or estimating the number of visual objects may appear a very simple and basic task compared to many other human capacities. However, the strategies used to solve these tasks may be much more complex and informative than previously thought. Groupitizing, as subitizing, is an example of this complexity and informative power.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: Zenodo (http://doi.org/10.5281/zenodo.4638767).

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Commissione per l'Etica della Ricerca, University of Florence, July 7, 2020, n. 111. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

GA, EC, and PM performed the testing and data collection. GA performed the data analysis. All authors contributed to the study concept, experimental design, interpretation of results, manuscript preparation, and approved the final version of the manuscript.

## FUNDING

This research was funded from the European Union (EU) and Horizon 2020—Grant Agreement no. 832813—ERC Advanced "Spatio-temporal mechanisms of generative perception—GenPercept"; from the Marie Skłodowska-Curie grant agreement No. 885672—DYSC-EYE-7T—"The neural substrate of numerical cognition in dyscalculia revealed by eye tracking and ultra-high field 7T functional magnetic imaging"; and from Italian Ministry of Education, University, and Research under the PRIN2017 program (Grant no. 2017XBJN4F—"EnvironMag" and Grant no. 2017SBCPZY—"Temporal context in perception: serial dependence and rhythmic oscillations").

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.687321/full#supplementary-material

## REFERENCES

Anobile, G., Arrighi, R., and Burr, D. C. (2019). Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities. *J. Exp. Child. Psychol.* 178, 86–103. doi: 10.1016/j.jecp.2018.09.017

Anobile, G., Arrighi, R., Castaldi, E., Grassi, E., Pedonese, L., Moscoso, P. A. M., et al. (2018). Spatial but not temporal numerosity thresholds correlate with formal math skills in children. *Dev. Psychol.* 54, 458–473. doi: 10.1037/dev0000448

Anobile, G., Castaldi, E., Maldonado, M. P. A., Burr, D. C., and Arrighi, R. (2020). "Groupitizing": a strategy for numerosity estimation [Data set]. *Zenodo*. doi: 10.5281/zenodo.4292106

Anobile, G., Castaldi, E., Turi, M., Tinelli, F., and Burr, D. C. (2016). Numerosity but not texture-density discrimination correlates with math ability in children. *Dev. Psychol.* 52, 1206–1216. doi: 10.1037/dev0000155

Anobile, G., Turi, M., Cicchini, G. M., and Burr, D. C. (2012). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vis. Res.* 74, 102–109. doi: 10.1016/j.visres.2012.06.005

Ansari, D., Lyons, I. M., van Eimeren, L., and Xu, F. (2007). Linking visual attention and number processing in the brain: the role of the temporo-parietal junction in small and large symbolic and nonsymbolic number comparison. *J. Cogn. Neurosci.* 19, 1845–1853. doi: 10.1162/jocn.2007.19.11.1845

Atkinson, J., Campbell, F. W., and Francis, M. R. (1976). The magic number 4 +/- 0: a new look at visual numerosity judgements. *Perception* 5, 327–334. doi: 10.1068/p050327

Bugden, S., Szkudlarek, E., and Brannon, E. M. (2021). Approximate arithmetic training does not improve symbolic math in third and fourth grade children. *Trends Neurosci. Educ.* 22:100149. doi: 10.1016/j.tine.2021.100149

Burr, D. C., Anobile, G., and Arrighi, R. (2017). Psychophysical evidence for the number sense. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170045. doi: 10.1098/rstb.2017.0045

Camos, V., and Tillmann, B. (2008). Discontinuity in the enumeration of sequentially presented auditory and visual stimuli. *Cognition* 107, 1135–1143. doi: 10.1016/j.cognition.2007.11.002

Castaldi, E., Vignaud, A., and Eger, E. (2020). Mapping subcomponents of numerical cognition in relation to functional and anatomical landmarks of human parietal cortex. *Neuroimage* 221:117210. doi: 10.1016/j.neuroimage.2020.117210

Caviola, S., Colling, L. J., Mammarella, I. C., and Szûcs, D. (2020). Predictors of mathematics in primary school: Magnitude comparison, verbal and spatial working memory measures. *Dev. Sci.* 23:e12957. doi: 10.1111/desc.12957

Chen, Q., and Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: a meta-analysis. *Acta Psychol. (Amst)* 148, 163–172. doi: 10.1016/j.actpsy.2014.01.016

Choo, H., and Franconeri, S. L. (2014). Enumeration of small collections violates Weber's law. *Psychon. Bull. Rev.* 21, 93–99. doi: 10.3758/s13423-013-0474-4

Ciccione, L., and Dehaene, S. (2020). Grouping mechanisms in numerosity perception. *Open Mind* 4, 102–118. doi: 10.1162/opmi_a_00037

De Smedt, B., Noël, M. P., Gilmore, C., and Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends Neurosci. Educ.* 2, 48–55. doi: 10.1016/j.tine.2013.06.001

Fazio, L. K., Bailey, D. H., Thompson, C. A., and Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *J. Exp. Child. Psychol.* 123, 53–72. doi: 10.1016/j.jecp.2014.01.013

Ferrand, L., Riggs, K. J., and Castronovo, J. (2010). Subitizing in congenitally blind adults. *Psychon. Bull. Rev.* 17, 840–845. doi: 10.3758/pbr.17.6.840

Gallace, A., Tan, H. Z., Haggard, P., and Spence, C. (2008). Short term memory for tactile stimuli. *Brain Res.* 1190, 132–142. doi: 10.1016/j.brainres.2007.11.014

Halberda, J., Mazzocco, M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–668. doi: 10.1038/nature07246

He, L., Zuo, Z., Chen, L., and Humphreys, G. (2014). Effects of number magnitude and notation at 7T: separating the neural response to small and large, symbolic and nonsymbolic number. *Cereb. Cortex* 24, 2199–2209. doi: 10.1093/cercor/bht074

Hoopen, T. G., and Vos, J. (1979a). Attention-switching and grouping in counting interaurally presented clicks. *Acta Psychol. (Amst)* 43, 283–297. doi: 10.1016/0001-6918(79)90037-4

Hoopen, T. G., and Vos, J. (1979b). Effect on numerosity judgement of grouping of tones by auditory channels. *Percept. Psychophys.* 26, 374–380. doi: 10.3758/bf03204162

Jevons, W. S. (1871). The power of numerical discrimination. *Nature* 3, 281–282. doi: 10.1038/003281a0

Kaufman, E. L., Lord, M. W., Reese, T. W., and Volkmann, J. (1949). The discrimination of visual number. *Am. J. Psychol.* 62:27. doi: 10.2307/1418556

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What's new in psychtoolbox-3. *Perception* 36:15.

Loconsole, M., De Agrò, M., and Regolin, L. (2021). Young chicks rely on perceptual grouping to discriminate prime numbers. *bioRxiv* [Preprint] doi: 10.1101/2021.03.04.433923 bioRxiv: 2021.2003.2004.433923,

Maldonado Moscoso, P. A., Castaldi, E., Burr, D. C., Arrighi, R., and Anobile, G. (2020). Grouping strategies in number estimation extend the subitizing range. *Sci. Rep.* 10:14979. doi: 10.1038/s41598-020-71871-5

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends Cogn. Sci.* 14, 542–551. doi: 10.1016/j.tics.2010.09.008

Piazza, M., Fumarola, A., Chinello, A., and Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition* 121, 147–153. doi: 10.1016/j.cognition.2011.05.007

Piazza, M., Mechelli, A., Butterworth, B., and Price, C. J. (2002). Are subitizing and counting implemented as separate or functionally overlapping processes? *Neuroimage* 15, 435–446. doi: 10.1006/nimg.2001.0980

Plaisier, M. A., Bergmann Tiest, W. M., and Kappers, A. M. (2009). One, two, three, many – subitizing in active touch. *Acta Psychol. (Amst)* 131, 163–170. doi: 10.1016/j.actpsy.2009.04.003

Plaisier, M. A., and Smeets, J. B. (2011). Haptic subitizing across the fingers. *Atten. Percept. Psychophys.* 73, 1579–1585. doi: 10.3758/s13414-011-0124-8

Plaisier, M. A., Tiest, W. M., and Kappers, A. M. (2010). Grabbing subitizing with both hands: bimanual number processing. *Exp. Brain Res.* 202, 507–512. doi: 10.1007/s00221-009-2146-1

Repp, B. H. (2007). Perceiving the numerosity of rapidly occurring auditory events in metrical and nonmetrical contexts. *Percept. Psychophys.* 69, 529–543. doi: 10.3758/bf03193910

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., and Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychol. Sci.* 19, 607–614. doi: 10.1111/j.1467-9280.2008.02130.x

Riggs, K. J., Ferrand, L., Lancelin, D., Fryziel, L., Dumur, G., and Simpson, A. (2006). Subitizing in tactile perception. *Psychol. Sci.* 17, 271–272. doi: 10.1111/j.1467-9280.2006.01696.x

Ross, J. (2003). Visual discrimination of number without counting. *Perception* 32, 867–870. doi: 10.1068/p5029

Rugani, R., Loconsole, M., and Regolin, L. (2017). A strategy to improve arithmetical performance in four day-old domestic chicks (*Gallus gallus*). *Sci. Rep.* 7:13900. doi: 10.1038/s41598-017-13677-6

Sasanguie, D., Defever, E., Maertens, B., and Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *Q. J. Exp. Psychol. (Hove)* 67, 271–280. doi: 10.1080/17470218.2013.803581

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., et al. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev. Sci.* 20:e12372. doi: 10.1111/desc.12372

Starkey, G. S., and McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *J. Exp. Child Psychol.* 126, 120–137. doi: 10.1016/j.jecp.2014.03.006

Vetter, P., Butterworth, B., and Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: evidence against a pre-attentive subitizing mechanism. *PLoS One* 3:e3269. doi: 10.1371/journal.pone.0003269

Wender, K. F., and Rothkegel, R. (2000). Subitizing and its subprocesses. *Psychol. Res.* 64, 81–92. doi: 10.1007/s004260000021

Whalen, J., Gallistel, C. R., and Gelman, R. (1999). Nonverbal counting in humans: the psychophysics of number representation. *Psychol. Sci.* 10, 130–137. doi: 10.1111/1467-9280.00120

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 11 August 2021
doi: 10.3389/fnhum.2021.713565

# Evidence for an A-Modal Number Sense: Numerosity Adaptation Generalizes Across Visual, Auditory, and Tactile Stimuli

*Irene Togoli*[1]* *and Roberto Arrighi*[2]

*[1]International School for Advanced Studies (SISSA), Trieste, Italy, [2]Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy*

**Edited by:**
*Joshua Oon Soo Goh,*
*National Taiwan University, Taiwan*

**Reviewed by:**
*Ting-Ting Chang,*
*National Chengchi University, Taiwan*
*Soledad Ballesteros,*
*National University of Distance*
*Education (UNED), Spain*

***Correspondence:**
*Irene Togoli*
*irene.togoli@gmail.com*

**Specialty section:**
*This article was submitted to*
*Cognitive Neuroscience,*
*a section of the journal*
*Frontiers in Human Neuroscience*

***Received:*** *23 May 2021*
***Accepted:*** *16 July 2021*
***Published:*** *11 August 2021*

**Citation:**
*Togoli I and Arrighi R (2021) Evidence for an A-Modal Number Sense: Numerosity Adaptation Generalizes Across Visual, Auditory, and Tactile Stimuli.*
*Front. Hum. Neurosci. 15:713565.*
*doi: 10.3389/fnhum.2021.713565*

Humans and other species share a perceptual mechanism dedicated to the representation of approximate quantities that allows to rapidly and reliably estimate the numerosity of a set of objects: an Approximate Number System (ANS). Numerosity perception shows a characteristic shared by all primary visual features: it is susceptible to adaptation. As a consequence of prolonged exposure to a large/small quantity ("adaptor"), the apparent numerosity of a subsequent ("test") stimulus is distorted yielding a robust under- or over-estimation, respectively. Even if numerosity adaptation has been reported across several sensory modalities (vision, audition, and touch), suggesting the idea of a central and a-modal numerosity processing system, evidence for cross-modal effects are limited to vision and audition, two modalities that are known to preferentially encode sensory stimuli in an external coordinate system. Here we test whether numerosity adaptation for visual and auditory stimuli also distorts the perceived numerosity of tactile stimuli (and vice-versa) despite touch being a modality primarily coded in an internal (body-centered) reference frame. We measured numerosity discrimination of stimuli presented sequentially after adaptation to series of either few (around 2 Hz; low adaptation) or numerous (around 8 Hz; high adaptation) impulses for all possible combinations of visual, auditory, or tactile adapting and test stimuli. In all cases, adapting to few impulses yielded a significant overestimation of the test numerosity with the opposite occurring as a consequence of adaptation to numerous stimuli. The overall magnitude of adaptation was robust (around 30%) and rather similar for all sensory modality combinations. Overall, these findings support the idea of a truly generalized and a-modal mechanism for numerosity representation aimed to process numerical information independently from the sensory modality of the incoming signals.

**Keywords: number sense, numerosity perception, adaptation, tactile perception, cross-modal perception, spatial selectivity**

## INTRODUCTION

Being able to rapidly estimate the number of objects in the surrounding environment is a fundamental ability for most animal species, humans included. For instance, the ability of selecting the location with more food (e.g., the branch of a tree rich in fruit), or the capacity to make a rapid fight or flight decision (i.e., according to how many predators an animal is facing), have clear implications for survival. Humans, as well as many animal species (Meck and Church, 1983; Emmerton et al., 1997; Kilian et al., 2003; Agrillo et al., 2008, 2011; Rugani et al., 2008) are endowed with a "*sense of number*" that allows them to rapidly—albeit approximately—estimate the number of items in the surrounding space: an Approximate Number System" (ANS). Such mechanism has been reported to be evolutionary ancient (Gallistel, 1990; Dehaene, 1997; Hauser et al., 2000) and innate (Antell and Keating, 1983; Izard et al., 2009) although its acuity has been shown to steadily increase with age in humans (Halberda et al., 2012).

Recent electrophysiological and imaging studies in humans support the existence of a dedicated brain system for the representation of approximate numerical magnitude. For example, studies leveraging on functional magnetic resonance imaging (fMRI) have shown numerosity-related activity in several visual regions throughout the brain dorsal stream, starting from low-level visual areas such as V1-V3 up to high-level associative areas in the parietal cortex (Piazza et al., 2004; Fornaciai and Park, 2018a; Castaldi et al., 2019; DeWind et al., 2019). The processing of numerosity has also been shown to be organized in maps, with a graded tuning to different numerosities resembling the topographic organization of visual sensory inputs in retinotopic maps (Harvey et al., 2013; Harvey and Dumoulin, 2017). The idea of numerosity processing being distributed across several visual areas including early visual cortices has been strengthened by EEG studies showing numerosity-specific brain responses soon after the stimulus onset, to suggest that numerosity is processed (at least partially) also in low-level sensory regions (Park et al., 2016; Fornaciai et al., 2017; Fornaciai and Park, 2018a,b).

Crucially, psychophysical studies have shown that numerosity is subject to adaptation. This is of particular importance, as adaptation is usually considered the hallmark of "primary" perceptual attributes such as, in the visual domain, orientation, color, or size. More specifically, Burr and Ross (2008) showed that after sustained exposure to a dot array containing either a large or small number of dots, the numerosity of the stimulus presented immediately after was strongly distorted, resulting in an under- or over-estimation, respectively (Burr and Ross, 2008). This finding, alongside evidence that numerosity perception obeys Weber's law (i.e., the threshold varies proportionally with the number of items), led many authors to consider it as a "*primary visual feature*" (see Anobile et al., 2014; Burr et al., 2018).

Additional studies leveraging on adaptation provided important evidence concerning the nature of the brain mechanisms dedicated to numerosity. For instance, it has been reported that numerosity adaptation affects spatial numerosity (i.e., an array of dots simultaneously presented over a region of space) as well as temporal numerosity (i.e., a sequence of flashes presented over a given interval of time) with adaptation to the latter class of stimuli being able to also distort estimates of the numerosity of arrays of dots. Moreover, numerosity adaptation was found to generalize across the visual and auditory modality: adapting to a series of auditory clicks changed the perceived numerosity of sequences of flashes and vice versa, with the adaptation effect being quantitatively similar to that measured within a single sensory modality (vision or audition; Arrighi et al., 2014). This form of cross-modal adaptation has supported the idea of the existence of a generalized, a-modal, mechanism for numerosity processing, possibly located at the top of the numerosity processing stream (i.e., in parietal associative areas like the intraparietal sulcus; Piazza et al., 2004; Harvey et al., 2013), an idea also supported by both neurophysiological studies in the monkey (Nieder, 2012, 2016) as well as imaging studies in humans (Dormal et al., 2010).

Despite the idea of a generalized sense of number, most of the studies on numerosity perception and in particular those dedicated to numerosity adaptation have been limited to the visual or auditory modality. Only recently, a study from our group (Togoli et al., 2021) investigated numerosity adaptation in touch by measuring to what extent numerosity estimates for tactile stimuli are affected by a sustained exposure to slow or rapid sequences of mechanical impulses on the subjects' finger skin. Adaptation for tactile numerosity turned out in being robust and quantitatively similar to that reported in vision and audition (Togoli et al., 2021). However, so far it has never been investigated whether and to what extent the processing of tactile numerosity affects the processing of numerosity in vision and audition or vice versa. On the one hand, such an interaction should be expected in light of the idea of a truly generalized (or a-modal) number sense meant to process stimulus numerosity regardless of the sensory channels conveying it. On the other hand, it might be that numerosity processing of visual and auditory stimuli converges on a shared mechanism because both systems similarly operate according to an external reference frame exploited to localize and process information of objects in the surrounding environment. Conversely, tactile stimuli are mainly processed *via* a reference frame initially defined in terms of the skin receptors that have been activated by sensory stimulation, which is turned into a spatial reference frame only at a subsequent stage, where sensory information is integrated with body posture—a process termed "tactile remapping." In other words, in case the interference in numerosity perception across sensory modalities only occurs for sensory channels that leverage on a similar coordinate system, it might be expected that the shared numerosity mechanism between vision and audition would not account for the processing of tactile numerosity information.

To test these hypotheses, we measured the interplay between vision, audition, and touch in numerosity perception by leveraging on the technique of adaptation. We measured the accuracy and precision of numerosity estimates for stimuli presented sequentially (temporal numerosity) in vision, audition, and touch and then measured whether and to what extent these estimates were affected by numerosity adaptation to a relatively high or low quantity of stimuli (i.e., either a low- or a high-frequency stream of stimuli sustained for several seconds) of the same or different sensory modality across several combinations. Namely, we tested: (1) the effect of tactile adaptation on tactile numerical estimates, and a series of cross-modal adaptation conditions concerning; (2) the effect of tactile adaptation on auditory numerosity; (3) the effect of auditory adaptation on tactile numerosity; (4) the effect of tactile adaptation on visual numerosity; and (5) the effect of visual adaptation on tactile numerosity. Furthermore, in one experimental condition (tactile adapters; visual test stimuli) we also tested the role of spatial congruency by measuring adaptation aftereffects when adaptor and test stimuli were superimposed (same spatial position) or with a spatial offset (different spatial positions). If the hypothesis of a truly a-modal number sense is correct, we expect adaptation to be effective irrespective of the modality of adaptor and test stimuli, and to be spatially localized to the adapted location (e.g., see Arrighi et al., 2014; Togoli et al., 2021). Conversely, if cross-modal adaptation could only be observed across similarly "distal" modalities such as vision and audition, then we expect the adaptation to tactile stimuli to affect perceived numerosity of tactile impulses but not that of visual or auditory stimuli. Our results show robust and significant numerosity adaptation effects for all combinations of sensory stimuli, supporting the idea of a truly generalized and a-modal numerosity processing system. Moreover, our results also indicate that cross-modal numerosity adaptation is spatially selective as it vanishes when adaptor and test stimuli are presented in different spatial locations.

# MATERIALS AND METHODS

## Participants

A total of 16 right-handed subjects participated in the study. The group was composed of six males and 10 females with ages ranging between 23 and 33 years ($M$ = 26, SD = 2.67). Six participants were included in each of the five experimental conditions of the present study. Note that the total number of participants does not match the summed sample size considering all the conditions because some of the participants were tested in multiple (but not all) conditions (see below "Behavioral Data Analysis" section). The inclusion criteria for the study required participants to have a normal or corrected-to-normal vision, and the absence of neurological, psychiatric and developmental disorders. The participants were tested separately and signed an informed consent form before participating in the study. All the experimental procedures were approved by the local ethics committee (Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer—Firenze FI) and were in line with the Declaration of Helsinki. Note that the sample size of the present study was decided *a priori* based on the cross-modal adaptation effects measured in Arrighi et al. (2014). Namely, we took the average effect size yielded by 2-Hz and 8-Hz adaptation to visual stimuli on numerical estimates of the sequence of sounds as well as the effect of auditory adaptation on visual numerical estimates. Considering this average effect size (Cohen's $d$ = 2.92), a power of 99%, and a two-tailed distribution, the estimated minimum sample size was five subjects.

## Apparatus and Stimuli

The experimental setup included a 17-inches touch screen monitor (resolution 1,280 × 1,024 pixels; refresh rate 60 Hz; LG-FLATRON L1732P), used to present the visual stimuli, and a Clark Synthesis Tactile Sound Transducer (TST429 platinum), positioned behind the screen (in a position corresponding to the location of the visual stimuli on the screen in all experimental conditions except that in which tactile and visual stimuli were presented spatially separated), used to deliver both auditory and tactile stimuli (**Figure 1A**). The tactile sound transducer was composed of a speaker with a rubber ball mounted on top of it, used to convey the speaker vibrations to the hand of the participant. Additionally, the transducer was mounted on an inflatable cushion to avoid the additional noise of vibrations spreading to the table.

All the stimuli used in the different conditions of the present study were generated using Matlab (version R2010a) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), on a computer running Windows 7. The visual stimuli were white discs of 5° diameter, displayed 8° to the left or to the right of the central fixation point (see below "Procedure" section). The auditory and tactile stimuli were both presented through the tactile sound transducer device positioned behind the screen, centered at 8° from the center of the screen. Auditory stimuli were 500-Hz sine waves, with a 5-ms ramp at the onset and offset played at an intensity of around 75 dB. Tactile stimuli were generated through 50-Hz sine waves, a frequency specifically chosen to elicit vibrations to the subjects' skin without being audible through the insulating headphones wore by participants (see below).

In all conditions, the test stimuli were pseudo-random sequences of flashes (i.e., white discs), tones, or vibrotactile pulses (**Figure 1E**), with numerosity ranging from 2 to 20 stimuli. However, during data analysis, we only considered numerosities from 5 to 15 to avoid edge effects at the highest extreme, and the subitizing range (numerosity <5) in the lowest extreme as estimates in the subitizing range are known to be errorless and not susceptible to adaptation (Anobile et al., 2020). Each stimulus in the sequence was presented for 40 ms. To minimize the temporal regularity of the sequence, the ISI between any two consecutive stimuli in each sequence was randomly determined, with the constrain of a minimum ISI of 40 ms between two consecutive stimuli, and an overall sequence duration of 2 s. Adaptor stimuli were similarly pseudo-random sequences of flashes, tones, or vibrotactile pulses (**Figure 1D**). Each stimulus in the sequence lasted for 40 ms. Two different adaptation conditions were defined. In the *low* adaptation condition, adaptor sequences had a frequency of 2 stimuli/s (2 Hz), while in the *high* adaptation condition the adaptor had a frequency of 8 stimuli/s (8 Hz). These adaptation frequencies were chosen to be consistent with previous studies from our group showing robust adaptation effects (Arrighi et al., 2014; Togoli et al., 2021). Note that following previous studies (e.g., Arrighi et al., 2014), in our experimental design we induced adaptation effects *via* a prolonged presentation of a sequence of stimuli presented either with a low (2 Hz) or high (8 Hz) frequency. This technique has been already shown to be highly effective in previous studies concerning perceived numerosity (Arrighi et al., 2014; Anobile et al., 2016; Togoli et al., 2020, 2021), and was also adopted to avoid potential positive (i.e., opposite to adaptation) "serial dependence" effects reported to occur with a shorter stimulus presentation (see for instance Fornaciai and Park, 2019a).

## Procedure

The experiment was performed in a quiet and dimly lit room, with participants wearing insulating headphones throughout the session, which allowed the auditory stimuli to be perceived but prevented the auditory feedback from the tactile stimuli. In all conditions, participants performed a numerosity estimation task of visual, auditory, or tactile impulses in a sequence, after being adapted to either visual, auditory, or tactile stimuli. More specifically, while participants fixated on a central fixation point, the adaptor stimulus was delivered first, followed by the test stimulus after an ISI of 900 ms (**Figure 1B**). In the first trial of each block, the adaptor stimulus was presented for 40 s. In the following trials, we delivered a shorter top-up adaptor stimulus for 6 s. Participants were instructed that the first sequence in each trial was not relevant for the task, while they had to attend to and report the numerosity of the stimuli in the second sequence. At the end of the trial, a virtual number pad appeared on the screen, and participants were instructed to dial the number of stimuli in the sequence they had perceived by using the computer mouse. The response number was displayed on the screen, and participants pressed another button to confirm their response, then the next trial started after 1 s.

Participants performed a total of five conditions (**Figure 1C**) involving a different combination of adaptation and test sensory modalities (tested separately). The conditions were as follows. (1) A purely tactile condition ("Tact-Tact"), in which both adaptor and test stimuli were sequences of tactile impulses. (2) A tactile-auditory condition ("Tact-Aud"), in which the adaptor was tactile, and the test stimulus was a sequence of sounds. (3) An auditory-tactile condition ("Aud-Tact") in which the adaptor was auditory, and the test stimulus was tactile. (4) A visual-tactile condition ("Vis-Tact"), entailing visual adaptation and tactile test stimuli. (5) A tactile-visual condition ("Tact-Vis") with tactile adaptation and visual test stimuli. This last condition was further divided into two different sub-conditions (interleaved within the same blocks), with test stimuli being either spatially matched (Matched position), or presented with a 16° spatial offset (Unmatched position). The two sub-conditions were devised to test for the spatial selectivity of the adaptation effect across the tactile and visual modality.

In the Tact-Tact, Tact-Aud, Aud-Tact, and Vis-Tact conditions, participants performed 7–9 blocks of 20 trials. In the Tact-Vis condition, instead, participants performed five blocks of 40 trials (with the exception of one participant who performed four blocks of 40 trials due to equipment failure). To avoid the different adaptation conditions to interfere with each other, they were performed in different days, with their order randomized across participants. Before the start of each condition, participants were familiarized with the stimuli by performing a few trials without adaptation. No feedback was provided concerning the participants' responses in any of the conditions. Each session took about 120 min, and participants were allowed to take frequent breaks between different blocks.

## Data Analysis

As a measure of accuracy in the numerosity estimation task, we computed for each subject, in each experimental condition, the average numerical estimate for each level of numerosity (5–15). Precision was instead measured in terms of Weber's fraction, defined as the standard deviation of numerical estimates divided by the average estimate ($WF = \sigma_{est}/\mu_{est}$), again computed separately for each subject and condition. To assess the effect of different types of adaptation on numerical estimates, we first performed a series of two-way repeated measures ANOVAs within each condition, with factors "numerosity" (5–15), and "adaptation" (low adaptation vs. high adaptation). Interactions between different factors observed in the ANOVAs were followed up with paired t-tests between low and high adaptation, at each level of the numerosity range. Note that to the purpose of this series of tests, in the Tact-Vis condition we only considered the case in which visual and tactile stimuli were presented in the same spatial position. A comparison between the matched and unmatched sub-conditions was performed separately to assess the spatial selectivity of the effect (see below).

We also assessed subjects' precision in the estimation task in terms of WFs for all conditions (defined by the sensory modality of adapters and test stimuli) and across the two kinds of adaptation (i.e., low vs. high). Statistical tests on precision were carried out with a two-way (independent-samples) ANOVA on WFs averaged across numerosities, with factors "condition" (Tact-Tact, Tact-Aud, Aud-Tact, Tact-Vis, Vis-Tact), and "adaptation" (low vs. high).

Moreover, to better assess the magnitude of effects across different conditions, and compare them directly, we computed an adaptation effect index (AI) as follows:

$$AI = \left( \frac{\overline{PN}_{low} - \overline{PN}_{high}}{\overline{PN}high} \right) * 100$$

where $\overline{PN}_{low}$ represents the average numerical estimate across all numerosities after low adaptation, and $\overline{PN}_{high}$ the average numerical estimate after high adaptation. To compare the effect across different conditions, first, we performed a one-way independent samples ANOVA on the AIs, and then we compared

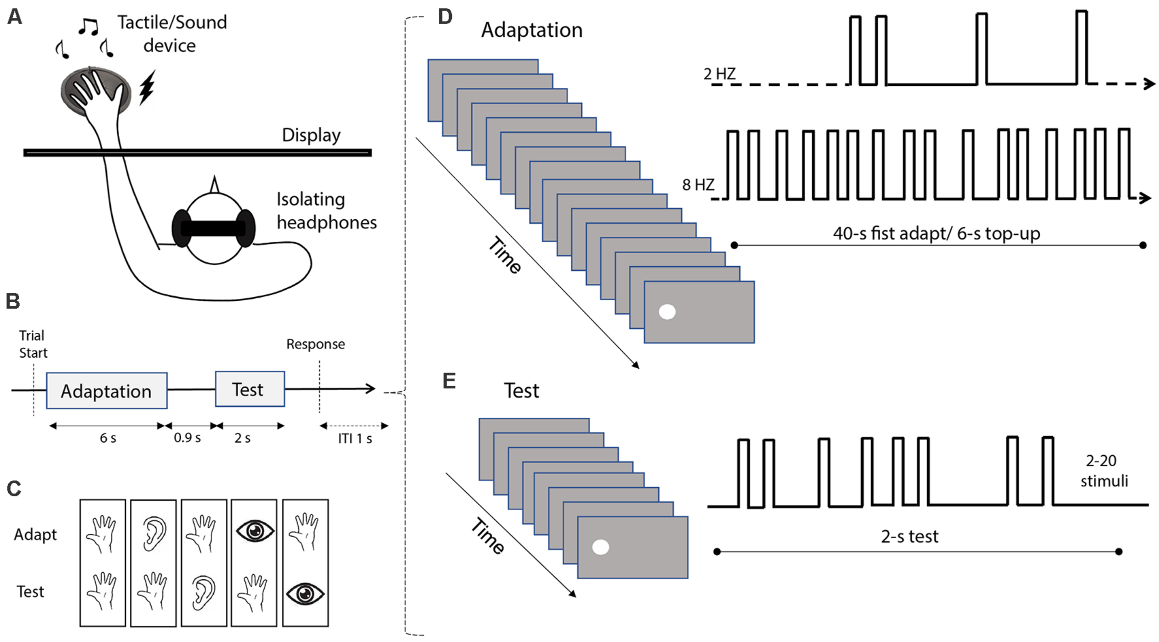


**FIGURE 1 |** Stimuli and experimental procedure. **(A)** Schematic depiction of the experimental setup. Participants seated in front of a monitor screen, with their hands positioned on the audio-tactile device located behind the screen. To avoid receiving auditory feedback during tactile stimulation, the subjects wore isolating headphones. **(B)** Depiction of the trial sequence. Each trial included an adaptation phase (6 s; with the exception of the first trial where the adaptation duration was 40 s) followed by a test phase (2 s), with an inter-stimulus interval of 0.9 s. At the end of each trial, participants were asked to estimate the numerosity of the test stimulus. After providing a response, the next trial started after an inter-trial interval (ITI) of 1 s. **(C)** The different combinations of adaptation and test stimulus modality (from left to right: tactile-tactile, audio-tactile, tactile-audio, visual-tactile, and tactile-visual). **(D)** Example of the adaptation sequence. Two different types of adaptation were tested in each condition: "low" adaptation, involving stimulation at 2 Hz, and "high" adaptation, involving stimulation at 8 Hz. In the first trial, adaptation was delivered for 40 s, while in each following trial we delivered a "top-up" adaptation of 6 s. **(E)** Example of the test stimulus sequence. Each test sequence included 2–20 individual stimuli, delivered within an interval of 2 s. Note that these examples depict the presentation of visual adaptor or test stimuli, but in different conditions, the sequences could involve also sounds or vibrotactile pulses. The examples are not depicted in scale.

individually each condition with a series of independent samples t-tests. To account for multiple comparisons, we applied a false-discovery rate (FDR) procedure with $q = 0.05$.

Finally, in the TactVis condition, we assessed the spatial selectivity of the effect by comparing the adaptation effects when adaptor and test stimuli were superimposed or spatially separated. First, we performed a three-way repeated-measures ANOVA with factors "numerosity" (5–15), "adaptation" (low vs. high), and "test location" (matched vs. unmatched). This test was followed up by *post hoc* tests to address interaction effects. Finally, we directly compared the magnitude of the effect measured in the two conditions in terms of the adaptation index. To do so, we performed two one-sample t-tests against the null hypothesis of zero effect, and a paired t-test comparing the effect in the matched and unmatched conditions.

Note that similarly to previous studies from our group (Arrighi et al., 2014; Anobile et al., 2016, 2020; Togoli et al., 2020, 2021), the adaptation effect here is computed considering two opposite adaptation conditions, rather than considering the difference from a baseline condition without adaptation. Although performing a baseline condition might provide more evidence concerning the adaptation-induced distortion of perceived numerosity compared to the absence of adaptation, it could introduce biases in the estimation of the effect. Indeed, having different sequences of stimuli (i.e., with or without the presentation of the adaptor) might provide different biases through time-order errors (i.e., the systematic under- or over-estimation of the first stimulus in a sequence; see for instance Hellström, 1985). For this reason, we chose not to add a baseline condition, and compute the adaptation effect as the difference between two opposite adaptation conditions.

## RESULTS

In all experimental conditions, we measured subjects' average estimates for each numerosity for both high and low adaptation. **Figure 2** shows data for the pure tactile experiment (Tact-Tact), in which both the adapter and test stimuli were tactile. As a consequence of adaptation to sequences of few tactile impulses (2 Hz), all subjects showed a tendency to overestimate the numerosity of the presented test stimuli (blue data point in **Figure 2A**). On the contrary, after adaptation to sequences entailing numerous stimuli (8 Hz), subjects showed a robust tendency to underestimate the numerosity of the test stimuli (red data points in **Figure 2A**). This pattern of results did hold for all possible combinations of stimulus sensory modalities (**Supplementary Figure 1**) and it is in line with the effects of

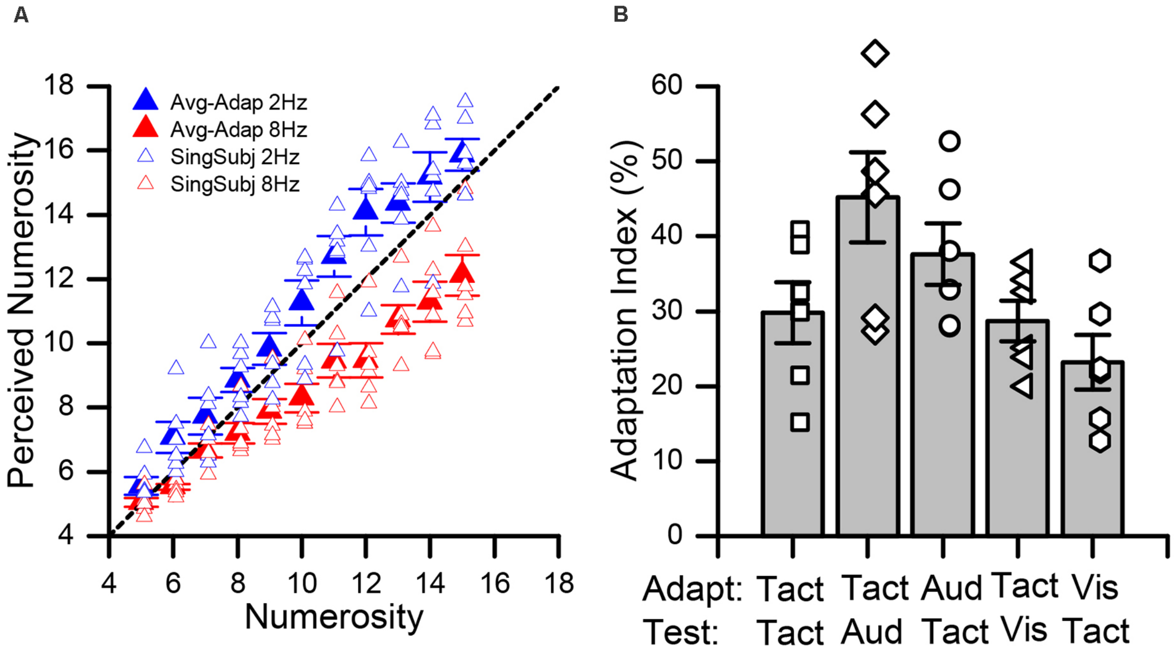


**FIGURE 2 |** Effect of adaptation in the tactile and cross-modal conditions. **(A)** Average numerical estimates for each numerosity (from 5 to 15), in the pure tactile (Tact-Tact) condition. Data relative to the low adaptation condition (2 Hz) are shown in blue whilst those for high adaptation (8 Hz) in red. Individual data (averaged over trials) are shown by the empty symbols whilst bold filled symbols indicate averages across participants. **(B)** Average adaptation effect indexes (AIs) across the different conditions. The empty data points show the adaptation effect for all participants. Error bars represent SEM.

numerosity adaptation reported in previous studies (e.g., Burr and Ross, 2008; Arrighi et al., 2014; Togoli et al., 2021).

To assess the significance of the adaptation effect across the different conditions, we performed a series of two-way repeated measures ANOVAs on the average numerical estimates, with factors "numerosity" (5–15), and "adaptation" (low vs. high). In all the conditions, as expected, we observed a significant main effect of numerosity (TactTact: $F_{(10,55)} = 157.9$, $p < 0.001$, $\eta^2_p = 0.97$; TactAud: $F_{(10,55)} = 127.8$, $p < 0.001$, $\eta^2_p = 0.96$; AudTact: $F_{(10,55)} = 169.2$, $p < 0.001$, $\eta^2_p = 0.97$; TactVis: $F_{(10,55)} = 109.7$, $p < 0.001$, $\eta^2_p = 0.96$ and VisTact: $F_{(10,55)} = 134.5$, $p < 0.001$, $\eta^2_p = 0.96$). Moreover, we also observed a significant main effect of adaptation, again across all the conditions (TactTact: $F_{(1,5)} = 52.5$, $p < 0.001$, $\eta^2_p = 0.91$; TactAud: $F_{(1,5)} = 79.3$, $p < 0.001$, $\eta^2_p = 0.94$; AudTact: $F_{(1,5)} = 89.4$, $p < 0.001$, $\eta^2_p = 0.95$; TactVis: $F_{(1,5)} = 136.9$, $p < 0.001$, $\eta^2_p = 0.96$ and VisTact: $F_{(1,5)} = 38.1$, $p = 0.001$, $\eta^2_p = 0.88$). Furthermore, we also observed in all conditions a significant interaction between the two factors (TactTact: $F_{(10,55)} = 7.6$, $p < 0.001$, $\eta^2_p = 0.60$; TactAud: $F_{(10,55)} = 12.2$, $p < 0.001$, $\eta^2_p = 0.71$; AudTact: $F_{(10,55)} = 15.7$, $p < 0.001$, $\eta^2_p = 0.76$; TactVis: $F_{(10,55)} = 6.7$, $p < 0.001$, $\eta^2_p = 0.57$ and VisTact: $F_{(10,55)} = 5.2$, $p < 0.001$, $\eta^2_p = 0.51$), to suggest differences in the strength of adaptation for different levels of numerosity. Indeed, looking at **Figure 2A** (**Supplementary Figure 1**), it is evident that adaptation is more effective at relatively high numerosities. However, a series of *post hoc* paired *t*-tests (corrected for multiple comparisons with a false discovery rate, FDR, procedure, with $q = 0.05$) within each numerosity showed a statistically significant difference between numerical estimates after low vs. high adaptation for the majority of the tested numerosities with just few exceptions. In the Tact-Tact and Tact-Aud condition, all comparisons were statistically significant (max FDR-adjusted *p*-value = 0.044). In the Aud-Tact and Tact-Vis condition, all comparisons were significant (max FDR-adjusted *p*-value = 0.049 and 0.043, respectively), with the exception of the numerosity level 5 in the Aud-Tact condition (adj-$p = 0.21$) and the numerosity level 6 in the Tact-Vis condition (adj-$p = 0.06$). Finally, in the Vis-Tact condition, all the comparisons were significant (max adj-$p = 0.016$), with the exception of numerosity level 5 and 7 (adj-$p = 0.11$ and 0.12, respectively).

In addition to subjects' accuracy in numerosity estimates (i.e., the mean numerical estimates, reflecting perceived numerosity), we also measured their precision in terms of Weber's fraction (WF; see "Materials and Methods" section). We measured whether there was any difference in precision across the different conditions and as a function of the adaptation frequency (i.e., low vs. high). The average WFs across the different conditions are shown in **Supplementary Figure 2**. To this aim, we performed a two-way (independent samples) ANOVA on WF measures averaged across numerosities, with factor "condition" (Tact-Tact, Tact-Aud, Aud-Tact, Tact-Vis, Vis-Tact), and "adaptation" (low vs. high). The results showed neither a main effect of condition ($F_{(4,25)} = 1.01$, $p = 0.41$), nor a main effect of adaptation ($F_{(1,5)} = 0.006$, $p = 0.94$), and no interaction between the two factors ($F_{(4,25)} = 0.17$, $p = 0.95$).

Given that WFs reflects variability in subjects' responses and this is meant to reflect the noise related to the perceptual process, we can conclude that in none of the conditions the two kinds of adaptation differed in providing a different amount of variability in numerosity processing.

Moreover, in order to obtain a direct comparison of the magnitude of the adaptation effect and compare the effects observed in different conditions, we calculated an adaptation effect index (AI) as the normalized difference between numerical estimates after low and high adaptation, turned into percentage (see formula 1 in the "Data Analysis" section). The average AIs across the different conditions tested are shown in **Figure 2B**. Overall, we observed robust adaptation effects across all conditions. Indeed, a series of one-sample $t$-tests (against the null hypothesis of zero effect; corrected with FDR) showed that the effect is significant in all tested conditions (Tact-Tact: $t_{(5)} = 7.32$, adjusted-$p < 0.001$, Cohen's $d = 2.99$; Aud-Tact: $t_{(5)} = 9.15$, $p < 0.001$, $d = 3.07$; Tact-Aud: $t_{(5)} = 7.52$, $p < 0.001$, $d = 3.72$; Tact-Vis: $t_{(5)} = 10.64$, $p < 0.001$, $d = 4.34$; Vis-Tact: $t_{(5)} = 6.37$, $p = 0.001$, $d = 2.60$). Then, we performed a one-way independent samples ANOVA (with factor "condition") to compare the magnitude of adaptation across all the combinations of sensory modalities of adapting and test stimuli. The results show a significant main effect of condition ($F_{(4, 25)} = 4.1$, $p = 0.01$, $\eta^2_p = 0.40$), suggesting that the adaptation magnitude might actually vary across conditions depending on which modality adaptor and test stimuli belonged to. To further investigate this, we ran a series of pairwise independent-sample $t$-tests comparing the conditions against each other. Again, to account for multiple comparisons, we applied an FDR procedure with $q = 0.05$. The results showed no statistically significant differences across conditions after correcting for multiple comparisons (max $t$-value = 3.14, min adjusted $p$-value = 0.10), suggesting that numerosity adaptation effects across vision, audition, and touch are quite similar in magnitude, regardless of the sensory modality of the adapting and test stimuli.

While the effect of numerosity adaptation within different modalities (visual, auditory, tactile) has been shown to be spatially localized (Arrighi et al., 2014; Togoli et al., 2021), is the cross-modal effect similarly selective for the position of the stimuli? The hypothesis of a truly a-modal numerosity processing system predicts indeed that adaptation should show similar properties—included spatial selectivity—irrespective of the sensory modality of the adaptor and test stimuli, and irrespective of whether the two stimuli belong to the same or different modalities. To address this prediction, we divided the Tact-Vis condition into two sub-conditions. In one condition the visual test stimulus was presented spatially superimposed with the position of the tactile adapter (matched condition), whilst in the other, it was presented with a horizontal spatial offset (unmatched condition). The prediction was straightforward: if the effect is spatially selective, we would expect a significant adaptation effect only when adaptor and test stimuli are presented in a spatially matched position.

The results are shown in **Figure 3**. To assess the effect of adaptation in the matched and unmatched condition, we first performed a three-way repeated measures ANOVA, with factors "numerosity" (5–15), "adaptation" (low vs. high), and "test position" (matched vs. unmatched). The results showed a main effect of numerosity ($F_{(10,55)} = 139.6$, $p < 0.01$, $\eta^2_p = 0.97$), a main effect of adaptation ($F_{(1,5)} = 26.9$, $p = 0.003$, $\eta^2_p = 0.84$), and a main effect of test position ($F_{(1,5)} = 6.6$, $p = 0.004$, $\eta^2_p = 0.57$). A significant two-way interaction was observed between numerosity and adaptation ($F_{(10,55)} = 5.03$, $p < 0.01$, $\eta^2_p = 0.50$), and between adaptation and position ($F_{(1,5)} = 27.16$, $p < 0.01$, $\eta^2_p = 0.84$). On the contrary, no significant interaction

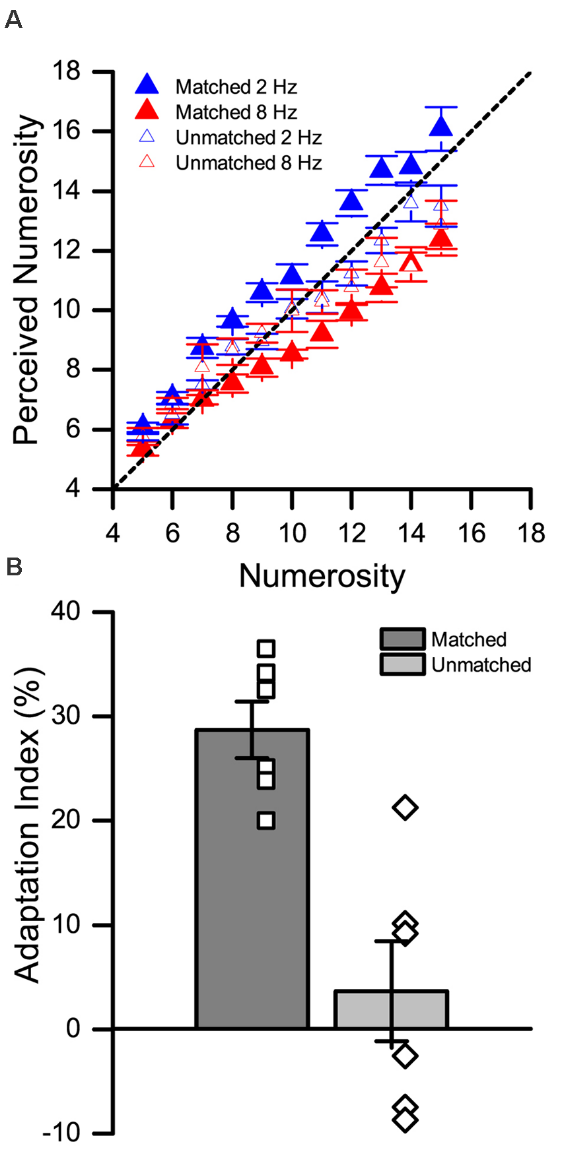


**FIGURE 3** | Spatial selectivity in the cross-modal tactile-visual (Tact-Vis) condition. **(A)** Average numerical estimates for each numerosity presented in the low (blue triangles) and high (red triangles) matched condition. The empty symbols correspond to the unmatched condition and the filled symbols to the matched condition. **(B)** Average adaptation effect indexes (AIs) in the matched (dark gray bar) and unmatched (light gray bar) condition. The empty symbols show the individual estimates of the effect. Error bars represent SEM.

was found between numerosity and position ($F_{(10,55)}$ = 1.57, $p$ = 0.14). Finally, we observed a significant three-way interaction between numerosity, adaptation and position ($F_{(10,55)}$ = 2.21, $p$ = 0.03, $\eta_p^2$ = 0.31).

To address the nature of this three-way interaction, we performed a series of (FDR-corrected) *post hoc* paired $t$-tests comparing numerical estimates after low vs. high adaptation at each numerosity level, separately for the matched and unmatched condition. In the matched condition we observed a statistically significant effect of adaptation (i.e., a significant difference between numerical estimates after low vs. high adaptation) for each numerosity (max adj-$p$ = 0.043) except one (6, adj-$p$ = 0.062). Conversely, in the unmatched condition, we did not observe any significant difference in numerical estimates induced by adaptation at any numerosity (min adj-$p$ = 0.84).

Finally, we also computed the AI for the matched vs unmatched condition, and compared them. As shown in **Figure 3B**, while in the matched condition the adaptation effect is robust, it is almost null when adapting and test stimuli were presented in different spatial locations. A two one-sample t-tests (against zero) showed that while the effect in the matched condition was significantly higher than zero ($t_{(5)}$ = 10.64, $p$ < 0.001, $d$ = 4.34), the effect was not significant in the unmatched condition ($t_{(5)}$ = 0.76, $p$ = 0.48, $d$ = 0.31). In line with that, a paired t-test further showed that the effect in the matched condition is significantly higher compared to the unmatched condition ($t_{(5)}$ = 5.32, $p$ = 0.003, $d$ = 2.17).

## DISCUSSION

In the present study, we tested the idea of a generalized and a-modal mechanism to process numerosity in the human brain by measuring the effect of adaptation across different sensory modalities. Participants were asked to estimate the numerosity of either a sequence of brief flashes, tones, or vibrotactile pulses. Crucially, before the presentation of these test stimuli, participants were adapted to sequences of either flashes, tones, or vibrotactile pulses, at different frequencies entailing a relatively low or relatively high number of events (i.e., low and high adaptation condition, respectively). The conditions tested included a purely tactile condition (tactile adaptation on tactile numerical estimates; Tact-Tact), and a series of cross-modal combinations: tactile adaptation on auditory or visual numerical estimates (Tact-Aud and Tact-Vis, respectively) and auditory or visual adaptation on tactile numerical estimates (Aud-Tact and Vis-Tact, respectively). Overall, our results show robust and significant adaptation effects: a period of 2 Hz stimulation yielded robust overestimation of perceived numerosity of the subsequent test stimulus, while 8 Hz adaptation caused a relative underestimation. Importantly, we show that adaptation aftereffects were quantitatively similar across all the combinations of stimulus sensory modalities.

Despite decades of studies, the brain mechanisms supporting the ability to rapidly and approximately estimate quantities of items—an ability fundamental for survival—remain unclear. In recent years, neuroimaging studies have started to uncover the brain areas and the processing stages linked to numerosity perception. For instance, fMRI studies on visual numerosity perception have shown a pathway for the processing of approximate numerical information starting from the early stations of the visual cortex, towards high-level associative cortices in the parietal cortex. Indeed, although the parietal cortex is the most consistently reported brain region associated with numerosity perception (e.g., Piazza et al., 2004; Dormal and Pesenti, 2009; Harvey et al., 2013; Borghesani et al., 2019; but see Cavdaroglu et al., 2015; and Cavdaroglu and Knops, 2019)—and thus it is considered the core of its processing pathway—other studies have started to uncover the contributions of earlier sensory areas. Indeed, early visual areas such as V1, V2, and V3, have started to be increasingly reported in fMRI as associated with numerosity processing (Fornaciai and Park, 2018a; Castaldi et al., 2019; DeWind et al., 2019). Electroencephalography (EEG) studies further support this idea of a numerosity processing pathway starting from early sensory areas, at least in vision. Namely, it has been recently shown that numerosity-related evoked activity emerges as early as 75–100 ms after the onset of a stimulus (i.e., C1 component), and from areas like V1-V3 (Park et al., 2016; Fornaciai and Park, 2017, 2018a; Fornaciai et al., 2017; Van Rinsveld et al., 2020), and continues through later latencies (i.e., 180–200 ms, P2p component; e.g., Temple and Posner, 1998; Libertus et al., 2007; Hyde and Wood, 2011). All these neuroimaging studies thus provide evidence for the existence of a dedicated brain network for the processing of approximate numerical information.

Studies at the behavioral level further support the existence of brain mechanisms specific to numerosity. For instance, it has been shown that visual perception is more sensitive to numerosity than to other non-numerical visual attributes like texture-density (Anobile et al., 2016; Cicchini et al., 2016), suggesting indeed the existence of dedicated brain mechanisms for numerosity processing (although non-numerical attributes may still contribute to numerosity perception, see for instance Dakin et al., 2011 and Leibovich et al., 2017, for alternative accounts). Furthermore, numerosity perception has also been shown to be modulated by the spatio-temporal properties of the stimuli and by motion (Fornaciai and Park, 2018c; Fornaciai et al., 2018), suggesting again a role for relatively early sensory areas. Most notably, it has been shown that numerosity perception is subject to perceptual adaptation (Burr and Ross, 2008; see Kohn, 2007 for a review on adaptation). Perceptual adaptation is indeed considered the hallmark of a fundamental—*primary*—perceptual attribute (i.e., like for instance orientation, color, contrast, or motion; Burr and Ross, 2008; Grasso et al., 2021; but see Durgin, 2008, for an alternative account). Interestingly, numerosity adaptation has been shown to not be limited to vision, but to also extend to other modalities, like audition (Arrighi et al., 2014; Togoli et al., 2020) and touch (Togoli et al., 2021). Even more striking, is the observation of cross-modal adaptation: adapting to a stream of auditory events can affect the perceived numerosity of visual stimuli, and vice versa (Arrighi et al., 2014; see also

Anobile et al., 2016, 2020, and Togoli et al., 2020; Maldonado Moscoso et al., 2020, for adaptation effects across the motor and sensory domain).

The observation of cross-modal adaptation has suggested the existence of a generalized, a-modal, number sense (Arrighi et al., 2014), most likely implemented at the top of the numerosity processing pathway (i.e., parietal cortex; see for instance Castaldi et al., 2016). These results nicely complement the neurophysiological results in the monkey brain and imaging data in the human brain. For example, neurons in the ventral intraparietal sulcus (IPS) of monkeys have been reported to encode numerosity for both auditory and visual sensory modalities to suggest that numerosity information eventually converges to a more abstract representation (Nieder, 2012, 2016). Similarly, in humans, a right lateralized frontoparietal circuit activated by both auditory and visual number sequences, has been reported (Piazza et al., 2006).

The present results further extend and support the idea of a generalized number sense, by showing that adaptation occurs across a wide range of cross-modal conditions. Previous results have been indeed limited to the auditory and visual modalities—two modalities that are both characterized by the need of processing distal stimuli (i.e., stimuli originating away from the sensory organ transducing their energy). Such similarity between these two modalities raised the question of whether the number sense is truly amodal, as the cross-modal adaptation may remain limited to auditory and visual stimulation. A truly amodal processing system would instead predict similar adaptation effects irrespective of the sensory modality through which adaptor and test stimuli are delivered—even when a quite different modality, like touch, is involved. And this is exactly what our results show: adaptation generalizes across several different cross-modal combinations, and works similarly irrespective of the sensory modality involved. In terms of the brain processing stage probed by adaptation, our results suggest that adaptation occurs at a level in the processing hierarchy at which signals from different sensory modalities interact with each other. In previous studies, cross-sensory interactions have been observed at multiple levels of "uni-sensory" pathways, and as early as the primary sensory cortices of different modalities (e.g., Laurienti et al., 2002; Schroeder and Foxe, 2005; Mishra et al., 2007; Sperdin, 2009; Vasconcelos et al., 2011). However, results from both the present and previous studies congruently suggest that numerosity adaptation mainly occurs in higher-order integrative cortical areas such as the parietal cortex (Castaldi et al., 2016). In line with that, previous results from our group show that numerosity adaptation also generalizes across the perceptual and motor system: adaptation to a series of self-generated tapping movements distorts the perceived numerosity of subsequently-presented visual (Anobile et al., 2016) or auditory (Togoli et al., 2020) stimuli. Thus, numerosity adaptation seems to occur at the converging point of modality-specific sensory pathways and motor signals, making the parietal cortex the best candidate locus for the brain mechanisms involved in numerosity adaptation (e.g., Iacoboni, 2006; Tosoni et al., 2008).

The fact that we did not observe a significant difference across the different adaptation conditions is in line with the idea of a high-level mechanism mediating the number sense. This result is particularly interesting, as one may intuitively expect to observe the stronger and more robust effect in the uni-modal condition (i.e., involving only tactile stimulation). The observation of no significant differences across the different conditions thus supports the idea of a truly generalized, a-modal number sense, whereby the processing of different numerical quantities and adaptation effects are independent of the sensory modality the numerosity information originally belonged to. However, caution is in order when interpreting the non-significance of this result. Indeed, our study was designed to detect a significant adaptation effect against the null hypothesis of zero effect, and not a subtler difference in the level of effect across different conditions, since we did not have a clear a-priori hypothesis concerning this point. Our design may thus lack the necessary power to detect a significant difference across conditions, leaving this point as an open question that should be addressed by future studies.

Furthermore, we also show that tactile adaptation has a spatially-localized effect on visual stimuli, similar to previous studies showing spatially localized adaptation effects in vision and in the tactile modality (Arrighi et al., 2014; Togoli et al., 2020; see also Anobile et al., 2020). In other words, in the Tact-Vis condition, we show that tactile adaptation can affect the perceived numerosity of a visual stimulus only when such a stimulus is presented in the same position as the adaptation. This is particularly important, for two reasons. First, it shows that numerosity processing involves the same spatio-temporal computations in different modalities, and suggests a common encoding of numerical information from the two modalities within a similar topographic representation of external space. Second, it suggests that the effect is perceptual in nature, and not a cognitive or decisional effect, as in this latter case the effect of numerosity adaptation would be expected to occur regardless of the position of the stimuli, with no spatial selectivity (Arrighi et al., 2014).

It is important to note that the generalization across different sensory modalities seems to be a specific property of numerosity adaptation. Indeed, it has been shown that a different effect inducing an attractive bias based on the recent history of stimulation (i.e., serial dependence; Fischer and Whitney, 2014; Fornaciai and Park, 2018c) does not show such generalization. Namely, while serial dependence in numerosity perception entails a spatially-localized effect (Fornaciai and Park, 2018c) and works across sequentially and simultaneously presented visual stimuli (Fornaciai and Park, 2019a), similarly to adaptation, it does not extend between auditory and visual stimuli (Fornaciai and Park, 2019a). However, adaptation and serial dependence likely entail widely different neurophysiological and functional mechanisms (see for instance Fornaciai and Park, 2019b), which may explain this difference. Addressing and comparing these different mechanisms thus represents an interesting open question for future studies.

Finally, another important point to consider is whether the temporal frequency (or rate) *per se* of the stimuli—rather than their numerosity—might have played a role in the observed results. Indeed, our adaptation sequences were defined by different temporal frequencies: 2 Hz (low) vs. 8 Hz (high). However, although numerosity and temporal frequency are potentially confounded in this adaptation design, it is unlikely that temporal frequency adaptation could explain the observed results, for three main reasons. First, we need to consider the relation between the frequency of adaptor and test stimuli. Indeed, while the adaptor stimuli had either a frequency of 2 or 8 Hz, the frequency of the test stimuli (considering that they were presented in a 2-s interval) varied with numerosity, spanning from 2.5 Hz to 7.5 Hz (respectively for 5 and 15 stimuli). If the effect was mediated by temporal frequency, we would thus expect a variable pattern of adaptation effects at different numerosities: the effect should have increased with the difference in frequency between adaptor and test stimuli. Namely, 2 Hz adaptation should be minimally effective on low-numerosity stimuli, while it should have the strongest effect at higher numerosities. The opposite is true for 8 Hz adaptation, which should have the maximum effect at low numerosities and the minimum effect at higher numerosity. However, no such pattern is evident neither in our results (see **Figures 2A**, **3A** and **Supplementary Figure 1**), nor in previous reports leveraging on the same paradigm (Arrighi et al., 2014; Togoli et al., 2021). Second, previous results show no transfer of frequency adaptation across different modalities (Motala et al., 2018), or cross-modal effects that are tightly tuned to the frequency band of the stimuli (i.e., a 4 Hz stimulus is strongly affected by a 5 Hz adaptor, but less so by adaptors of slightly different frequency). Third, temporal frequency adaptation is usually considered a very low-level effect, occurring at the earliest levels of sensory processing like the lateral geniculate nucleus (LGN) in vision (Tan and Yao, 2009), and the primary somatosensory cortex (S1) in touch (Romo and Salinas, 2003). Such early locus of temporal frequency adaptation is thus at odds with the cross-modal transfer observed in the present study. For all these reasons, we believe that the observed results are more in line with a numerosity adaptation effect, rather than temporal frequency adaptation. Nevertheless, another aspect worth it mentioning is that in this specific adaptation protocol the effect does not seem to be modulated by the relative numerosity of the adaptor and test sequences. For instance, one may expect the effect to be modulated by the ratio between adaptor and test (e.g., Piazza et al., 2004). Our results instead show a consistent pattern of adaptation effects in the low and high adaptation conditions, with the magnitude of adaptation roughly increasing with increasing test numerosity. This shows that—in line with previous studies employing a similar methodology (Arrighi et al., 2014; Togoli et al., 2020)—the effect is indeed not modulated by the ratio of the stimuli. If so, we should have instead observed a quite different pattern of effects (i.e., the effect of high adaptation should have peaked at lower test numerosities, and vice versa for the low adaptation). A possibility explaining this feature of the effect might be the relatively long duration of the adaptor stimuli, preventing the visual system from tracking the total numerosity of the adaptor stimuli throughout their presentation interval. However, since we kept the duration of the adaptation sequences constant, our results could not clarify this point, which thus remains another interesting open question for future studies.

Finally, besides the specific mechanisms of numerosity perception, our results are consistent with a broader view of perception as being largely multisensory (e.g., Pascual-Leone and Hamilton, 2001). Stimulation from the external environment is indeed intrinsically multisensory, and object representation has been observed to be systematically facilitated in the presence of multisensory information (e.g., Amedi et al., 2005). Multisensory integration has been shown to affect even the very low-level properties of a stimulus, like for instance the position of a visual flash of light strongly biasing the perceived position of a sound, with multisensory information being integrated in a statistically optimal fashion (i.e., Alais and Burr, 2004). In line with this idea, we show that numerosity—which could be considered a primary perceptual feature (Anobile et al., 2016)—is processed in an intrinsically multi-modal fashion, with the effect of adaptation (e.g., see Kohn, 2007 for a review) occurring independently from the sensory modality of the adapting and test stimuli.

To conclude, our results show that the effect of adaptation on the perceived numerosity of sequential stimuli generalizes across several different cross-modal combinations: the adaptation effect works irrespective of which modality is used to convey adaptor and test stimuli. Our findings thus expand previous results concerning the cross-modal effects of adaptation in numerosity perception and provide novel evidence for the existence of a truly amodal, generalized mechanism for the processing of numerosity.

## DATA AVAILABILITY STATEMENT

All the datasets generated during the experiments described in this manuscript have been uploaded to Open Science Framework (OSF), and are accessible at the following link: https://osf.io/a47js/.

## ETHICS STATEMENT

The protocol of the present study (involving human participants) was reviewed and approved by the Comitato Etico Pediatrico Regionale—Azienda Ospedaliero-Universitaria Meyer—Firenze FI. The participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

IT and RA designed the research, revised and edited the draft, and agreed on the final version of the manuscript. IT programmed the task, piloted the study, performed the research, analyzed the data in interaction with RA, designed the figures in interaction with RA, and wrote the first draft of the article. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the European Research Council (ERC) under the EU's Horizon 2020 research and innovation program (grant agreement no. 832813), by the Italian Ministry of Education, University, and Research under the PRIN2017 program (grant number 2017XBJN4F—"EnvironMag" and by grant number 2017SBCPZY—"Temporal context in perception: serial dependence and rhythmic oscillations").

## ACKNOWLEDGMENTS

We thank Dr. Michele Fornaciai for useful comments on an earlier version of this manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.713565/full#supplementary-material.

## REFERENCES

Agrillo, C., Dadda, M., Serena, G., and Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Anim. Cogn.* 11, 495–503. doi: 10.1007/s10071-008-0140-9

Agrillo, C., Piffer, L., and Bisazza, A. (2011). Number versus continuous quantity in numerosity judgments by fish. *Cognition* 119, 281–287. doi: 10.1016/j.cognition.2010.10.022

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262. doi: 10.1016/j.cub.2004.01.029

Amedi, A., von Kriegstein, K., van Atteveldt, N. M., Beauchamp, M. S., and Naumer, M. J. (2005). Functional imaging of human crossmodal identification and object recognition. *Exp. Brain Res.* 166, 559–571. doi: 10.1007/s00221-005-2396-5

Anobile, G., Cicchini, G. M., and Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychol. Sci.* 25, 265–270. doi: 10.1177/0956797613501520

Anobile, G., Domenici, N., Togoli, I., Burr, D., and Arrighi, R. (2020). Distortions of visual time induced by motor adaptation. *J. Exp. Psychol. Gen.* 149, 1333–1343. doi: 10.1037/xge0000709

Anobile, G., Arrighi, R., Togoli, I., and Burr, D. C. (2016). A shared numerical representation for action and perception. *eLife* 5:e16161. doi: 10.7554/eLife.16161

Antell, S. E., and Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Dev.* 54, 695–701.

Arrighi, R., Togoli, I., and Burr, D. C. (2014). A generalized sense of number. *Proc. Biol. Sci.* 281:20141791. doi: 10.1098/rspb.2014.1791

Borghesani, V., de Hevia, M. D., Viarouge, A., Pinheiro-Chagas, P., Eger, E., Piazza, M., et al. (2019). Processing number and length in the parietal cortex: Sharing resources, not a common code. *Cortex* 114, 17–27. doi: 10.1016/j.cortex.2018.07.017

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vis.* 10, 433–436.

Burr, D. C., Anobile, G., and Arrighi, R. (2018). Psychophysical evidence for the number sense. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170045. doi: 10.1098/rstb.2017.0045

Burr, D., and Ross, J. (2008). A visual sense of number. *Curr. Biol.* 18, 425–428. doi: 10.1016/j.cub.2008.02.052

Castaldi, E., Aagten-Murphy, D., Tosetti, M., Burr, D., and Morrone, M. C. (2016). Effects of adaptation on numerosity decoding in the human brain. *NeuroImage* 143, 364–377. doi: 10.1016/j.neuroimage.2016.09.020

Castaldi, E., Piazza, M., Dehaene, S., Vignaud, A., and Eger, E. (2019). Attentional amplification of neural codes for number independent of other quantities along the dorsal visual stream. *eLife* 8:e45160. doi: 10.7554/eLife.45160

Cavdaroglu, S., and Knops, A. (2019). Evidence for a posterior parietal cortex contribution to spatial but not temporal numerosity perception. *Cereb. Cortex* 29, 2965–2977. doi: 10.1093/cercor/bhy163

Cavdaroglu, S., Katz, C., and Knops, A. (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception. *NeuroImage* 116, 135–148. doi: 10.1016/j.neuroimage.2015.04.019

Cicchini, G. M., Anobile, G., and Burr, D. C. (2016). Spontaneous perception of numerosity in humans. *Nat. Commun.* 7:12536. doi: 10.1038/ncomms12536

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., and Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proc. Natl. Acad. Sci. U S A* 108, 19552–19557. doi: 10.1073/pnas.1113195108

Dehaene, S. (1997). *The Number Sense*. New York, NY: Oxford University Press.

DeWind, N. K., Park, J., Woldorff, M. G., and Brannon, E. M. (2019). Numerical encoding in early visual cortex. *Cortex* 114, 76–89. doi: 10.1016/j.cortex.2018.03.027

Dormal, V., and Pesenti, M. (2009). Common and specific contributions of the intraparietal sulci to numerosity and length processing. *Hum. Brain Mapp.* 47:S111. doi: 10.1002/hbm.20677

Dormal, V., Andres, M., Dormal, G., and Pesenti, M. (2010). Mode-dependent and mode-independent representations of numerosity in the right intraparietal sulcus. *NeuroImage* 52, 1677–1686. doi: 10.1016/j.neuroimage.2010.04.254

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Curr. Biol.* 18, R855–R856. doi: 10.1016/j.cub.2008.07.053

Emmerton, J., Lohmann, A., and Niemann, J. (1997). Pigeons' serial ordering of numerosity with visual arrays. *Anim. Learn. Behav.* 25, 234–244.

Fischer, J., and Whitney, D. (2014). Serial dependence in visual perception. *Nat. Neurosci.* 17, 738–743. doi: 10.1038/nn.3689

Fornaciai, M., and Park, J. (2017). Distinct neural signatures for very small and very large numerosities. *Front. Hum. Neurosci.* 11:21. doi: 10.3389/fnhum.2017.00021

Fornaciai, M., and Park, J. (2018a). Dynamics of numerosity representation in early visual cortex. *J. Vis.* 18:87. doi: 10.1167/18.10.87

Fornaciai, M., and Park, J. (2018b). Early numerosity encoding in visual cortex is not sufficient for the representation of numerical magnitude. *J. Cogn. Neurosci.* 30, 1788–1802. doi: 10.1162/jocn_a_01320

Fornaciai, M., and Park, J. (2018c). Serial dependence in numerosity perception. *J. Vis.* 19:15. doi: 10.1167/18.9.15

Fornaciai, M., and Park, J. (2019a). Serial dependence generalizes across different stimulus formats, but not different sensory modalities. *Vis. Res.* 160, 108–115. doi: 10.1016/j.visres.2019.04.011

Fornaciai, M., and Park, J. (2019b). Spontaneous repulsive adaptation in the absence of attractive serial dependence. *J. Vis.* 19:21. doi: 10.1167/19.5.21

Fornaciai, M., Brannon, E. M., Woldorff, M. G., and Park, J. (2017). Numerosity processing in early visual cortex. *NeuroImage* 157, 429–438. doi: 10.1016/j.neuroimage.2017.05.069

Fornaciai, M., Togoli, I. and Arrighi, R. (2018). Motion-induced compression of perceived numerosity. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-25244-8

Gallistel, C. R. (1990). Representations in animal cognition: an introduction. *Cognition* 37, 1–22. doi: 10.1016/0010-0277(90)90016-D

Grasso, P. A., Anobile, G., and Arrighi, R. (2021). Numerosity adaptation partly depends on the allocation of implicit numerosity-contingent visuo-spatial attention. *J. Vis.* 21:12. doi: 10.1167/jov.21.1.12

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., and Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc. Natl. Acad. Sci. U S A* 109, 11116–11120. doi: 10.1073/pnas.1200196109

Harvey, B. M., and Dumoulin, S. O. (2017). A network of topographic numerosity maps in human association cortex. *Nat. Hum. Behav.* 1:0036. doi: 10.1038/s41562-016-0036

Harvey, B. M., Klein, B. P., Petridou, N., and Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science* 341, 1123–1126. doi: 10.1126/science.1239052

Hauser, M. D., Carey, S., and Hauser, L. B. (2000). Spontaneous number representation in semi-free-ranging rhesus monkeys. *Proc. Biol. Sci.* 267, 829–833. doi: 10.1098/rspb.2000.1078

Hellström, Å. (1985). The time-order error and its relatives: mirrors of cognitive processes in comparing. *Psychol. Bull.* 97, 35–61.

Hyde, D. C., and Wood, J. N. (2011). Spatial attention determines the nature of nonverbal number representation. *J. Cogn. Neurosci.* 23, 2336–2351. doi: 10.1162/jocn.2010.21581

Iacoboni, M. (2006). Visuo-motor integration and control in the human posterior parietal cortex: evidence from TMS and fMRI. *Neuropsychologia* 44, 2691–2699. doi: 10.1016/j.neuropsychologia.2006.04.029

Izard, V., Sann, C., Spelke, E. S., and Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc. Natl. Acad. Sci. U S A* 106, 10382–10385. doi: 10.1073/pnas.0812142106

Kilian, A., Yaman, S., von Fersen, L., and Güntürkün, O. (2003). A bottlenose dolphin discriminates visual stimuli differing in numerosity. *Anim. Learn. Behav.* 31, 133–142. doi: 10.3758/bf03195976

Kohn, A. (2007). Visual adaptation: physiology, mechanisms and functional benefits. *J. Neurophysiol.* 97, 3155–3164. doi: 10.1152/jn.00086.2007

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y.-F., Field, A. S., Stein, B. E., et al. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *J. Cogn. Neurosci.* 14, 420–429. doi: 10.1162/089892902317361930

Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2017). From "sense of number" to "sense of magnitude": the role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* 40:e164. doi: 10.1017/S0140525X16000960

Libertus, M. E., Woldorff, M. G., and Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behav. Brain Funct.* 3:1. doi: 10.1186/1744-9081-3-1

Maldonado Moscoso, P. A., Cicchini, G. M., Arrighi, R., and Burr, D. C. (2020). Adaptation to hand-tapping affects sensory processing of numerosity directly: evidence from reaction times and confidence. *Proc. Biol. Sci.* 287:20200801. doi: 10.1098/rspb.2020.0801

Meck, W. H., and Church, R. M. (1983). A mode control model of counting and timing processes. *J. Exp. Psychol. Anim. Behav. Proc.* 9, 320–334.

Mishra, J., Martinez, A., Sejnowski, T. J., and Hillyard, S. A. (2007). Early cross-modal interactions in auditory and visual cortex underlie a sound-induced visual illusion. *J. Neurosci.* 27, 4120–4131. doi: 10.1523/JNEUROSCI.4912-06.2007

Motala, A., Heron, J., McGraw, P. V., Roach, N. W., and Whitaker, D. (2018). Rate after-effects fail to transfer cross-modally: evidence for distributed sensory timing mechanisms. *Sci. Rep.* 8:924. doi: 10.1038/s41598-018-19218-z

Nieder, A. (2012). Supramodal numerosity selectivity of neurons in primate prefrontal and posterior parietal cortices. *Proc. Natl. Acad. Sci. U S A* 109, 11860–11865. doi: 10.1073/pnas.1204580109

Nieder, A. (2016). The neuronal code for number. *Nat. Rev. Neurosci.* 17, 366–382. doi: 10.1038/nrn.2016.40

Park, J., DeWind, N. K., Woldorff, M. G., and Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cereb. Cortex* 26:bhv017. doi: 10.1093/cercor/bhv017

Pascual-Leone, A., and Hamilton, R. (2001). Chapter 27 The metamodal organization of the brain. *Prog. Brain Res.* 134, 427–445. doi: 10.1016/S0079-6123(01)34028-1

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vis.* 10, 437–442.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44, 547–555. doi: 10.1016/j.neuron.2004.10.014

Piazza, M., Mechelli, A., Price, C. J., and Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: an fMRI study. *Brain Res.* 1106, 177–188. doi: 10.1016/j.brainres.2006.05.104

Romo, R., and Salinas, E. (2003). Flutter Discrimination: neural codes, perception, memory and decision making. *Nat. Rev. Neurosci.* 4, 203–218. doi: 10.1038/nrn1058

Rugani, R., Regolin, L., and Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *J. Exp. Psychol. Anim. Behav. Process* 34, 388–399. doi: 10.1037/0097-7403.34.3.388

Schroeder, C. E., and Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Curr. Opin. Neurobiol.* 15, 454–458. doi: 10.1016/j.conb.2005.06.008

Sperdin, H. F. (2009). Early, low-level auditory-somatosensory multisensory interactions impact reaction time speed. *Front. Integrat. Neurosci.* 3:2. doi: 10.3389/neuro.07.002.2009

Tan, Z., and Yao, H. (2009). The spatiotemporal frequency tuning of lgn receptive field facilitates neural discrimination of natural stimuli. *J. Neurosci.* 29, 11409–11416. doi: 10.1523/JNEUROSCI.1268-09.2009

Temple, E., and Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proc. Natl. Acad. Sci. U S A* 95, 7836–7841. doi: 10.1073/pnas.95.13.7836

Togoli, I., Crollen, V., Arrighi, R., and Collignon, O. (2020). The shared numerical representation for action and perception develops independently from vision. *Cortex* 129, 436–445. doi: 10.1016/j.cortex.2020.05.004

Togoli, I., Marlair, C., Collignon, O., Arrighi, R., and Crollen, V. (2021). Tactile numerosity is coded in external space. *Cortex* 134, 43–51. doi: 10.1016/j.cortex.2020.10.008

Tosoni, A., Galati, G., Romani, G. L., and Corbetta, M. (2008). Sensory-motor mechanisms in human parietal cortex underlie arbitrary visual decisions. *Nat. Neurosci.* 11, 1446–1453. doi: 10.1038/nn.2221

Van Rinsveld, A., Guillaume, M., Kohler, P. J., Schiltz, C., Gevers, W., Content, A., et al. (2020). The neural signature of numerosity by separating numerical and continuous magnitude extraction in visual cortex with frequency-tagged EEG. *Proc. Natl. Acad. Sci. U S A* 117, 5726–5732. doi: 10.1073/pnas.1917849117

Vasconcelos, N., Pantoja, J., Belchior, H., Caixeta, F. V., Faber, J., Freire, M. A. M., et al. (2011). Cross-modal responses in the primary visual cortex encode complex objects and correlate with tactile discrimination. *Proc. Natl. Acad. Sci. U S A* 108, 15408–15413. doi: 10.1073/pnas.1102780108

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 04 October 2021
doi: 10.3389/fnhum.2021.750964

# Can a Single Representational Object Account for Different Number-Space Mappings?

*Arnaud Viarouge[1]* and Maria Dolores de Hevia[2]**

*[1] Université de Paris, LaPsyDÉ, CNRS, Paris, France, [2] Université de Paris, INCC UMR 8002, CNRS, Paris, France*

Numbers are mapped onto space from birth on, as evidenced by a variety of interactions between the processing of numerical and spatial information. In particular, larger numbers are associated to larger spatial extents (number/spatial extent mapping) and to rightward spatial locations (number/location mapping), and smaller numbers are associated to smaller spatial extents and leftward spatial locations. These two main types of number/space mappings (number/spatial extent and number/location mappings) are usually assumed to reflect the fact that numbers are represented on an internal continuum: the mental number line. However, to date there is very little evidence that these two mappings actually reflect a single representational object. Across two experiments in adults, we investigated the interaction between number/location and number/spatial extent congruency effects, both when numbers were presented in a non-symbolic and in a symbolic format. We observed a significant interaction between the two mappings, but only in the context of an implicit numerical task. The results were unaffected by the format of presentation of numbers. We conclude that the number/location and the number/spatial extent mappings can stem from the activation of a single representational object, but only in specific experimental contexts.

**Keywords: number, space, mappings, SCE, spatial–numerical association of response code (SNARC)**

***Edited by:***
*Joonkoo Park,*
*University of Massachusetts Amherst,*
*United States*

***Reviewed by:***
*Rosa Rugani,*
*University of Padua, Italy*
*Elvio Blini,*
*University of Padua, Italy*

****Correspondence:***
*Arnaud Viarouge*
*arnaud.via@gmail.com*
*Maria Dolores de Hevia*
*dehevialola@gmail.com*

***Specialty section:***
*This article was submitted to*
*Cognitive Neuroscience,*
*a section of the journal*
*Frontiers in Human Neuroscience*

***Received:*** *31 July 2021*
***Accepted:*** *13 September 2021*
***Published:*** *04 October 2021*

***Citation:***
*Viarouge A and de Hevia MD*
*(2021) Can a Single Representational*
*Object Account for Different*
*Number-Space Mappings?*
*Front. Hum. Neurosci. 15:750964.*
*doi: 10.3389/fnhum.2021.750964*

## INTRODUCTION

Humans possess an inborn ability to represent, discriminate, and manipulate numerical quantities, an ability that is shared with many other species (Cantlon, 2018). This ability is supported by the approximate number system (or ANS, Burr and Ross, 2008; Dehaene, 2009; Odic and Starr, 2018), which allows us to estimate (and mentally manipulate) the approximate numerosity of a set without using any symbolic knowledge (language or counting). The main signature of this cognitive system is that the variable determining a successful discrimination is the ratio between the two numerosities to compare, so that the larger the ratio between them the better the discrimination, a signature that also governs discrimination for other continuous, perceptual dimensions (Feigenson et al., 2004). This core, numerical cognitive system is thought to show continuity in development, and therefore to support the acquisition of formal math and symbolic numerical representations, with individual differences in numerical acuity predicting, and correlating with, math scores later on in life (Halberda et al., 2008; Starr et al., 2013). Another crucial aspect characterizing numerical representations is their spatial signature. In fact, a widely accepted view on numerical cognition considers numbers as distributions of activation along a spatially oriented mental number line (Restle, 1970; Dehaene, 1992).

Interactions between number and spatial representations have been extensively described in the literature, using a variety of methods: from visuo-spatial tasks, such as line bisection and reproduction tasks, where numbers impact spatial performance (Fischer, 2001; de Hevia et al., 2006, 2008; de Hevia and Spelke, 2009; Viarouge and de Hevia, 2013), to numerical tasks, such as parity judgment or magnitude comparison, where visuo-spatial variables modulate numerical judgments (Dehaene et al., 1993; Fias, 1996; Fischer et al., 2003; Bulf et al., 2014). This bidirectional interaction has been described across ages in humans, from infants who from birth spontaneously associate small numbers to small spatial extents and leftward spatial locations and large numbers to large spatial extents and rightward spatial locations (de Hevia and Spelke, 2010; de Hevia et al., 2014, 2017; Bulf et al., 2016; Di Giorgio et al., 2019), through childhood (Girelli et al., 2000; de Hevia and Spelke, 2009; Patro and Haman, 2012) and up to adulthood (Dehaene et al., 1993; Fischer et al., 2003), and even in non-human animals, such as chicks and chimpanzees, who have been shown to exhibit lateralized spatial associations of numerical quantities similar to humans (Adachi, 2014; Drucker and Brannon, 2014; Rugani et al., 2015).

In general terms, the interactions between numerical and spatial information can be described according to two types of number-space mappings (de Hevia, 2021). On the one hand, numbers can be mapped onto corresponding spatial extents, with larger numbers associated with larger spatial extents. This type of mapping is well exemplified in the Stroop paradigm, also known as the size congruity effect (SCE): when subjects have to compare numerical quantities (be it in a symbolic or non-symbolic format), the information of physical size interacts with the quantity judgment, so that incongruent combinations (e.g., a small quantity occupying a large space) yield worse performance than congruent combinations (e.g., a small quantity occupying a small space). This type of interaction between size and number has also been described in line bisection tasks (e.g., de Hevia et al., 2006; de Hevia and Spelke, 2009) as well as in reproduction tasks (e.g., de Hevia et al., 2008; Viarouge and de Hevia, 2013), providing further support to the idea that number is mapped onto a corresponding physical spatial extent.

On the other hand, numbers can be associated with different, lateralized spatial locations. This phenomenon is represented by the Spatial Numerical Association of Response Codes (SNARC) effect, by which small numbers are associated with the left and large numbers with the right side of space (Dehaene et al., 1993), obtaining an advantage in performance when the response side/or number location and the numerical magnitude are aligned according to a left-to-right oriented representation, with increasing numbers toward the right. This effect might be modulated by factors such as scanning habits (left-to-right vs. right-to-left writing/reading direction: Shaki et al., 2009), and contextual experience factors (Pitt and Casasanto, 2020).

Besides a few exceptions (de Hevia et al., 2006, 2008; Cipora et al., 2020; de Hevia, 2021), most authors usually assume that these two types of mappings (number-extent and number-location) reflect the same representational object, often appealing to the activation of a "mental number line" when interpreting the source of the interaction between numerical and spatial information. In fact, any type of behavioral interaction between numerical and spatial information in a given task is accounted for by the idea that humans might represent different numerosities along an internal spatial continuum: the mental number line.

However, some findings cast doubt on this generally accepted assumption. First, humans at birth show evidence of a dissociation between the number-spatial extent and the number-spatial location mappings. In particular, from birth and during the first year of life, infants spontaneously create mappings between number and spatial extent that can be generalized to the dimension of time (de Hevia and Spelke, 2010; Srinivasan and Carey, 2010; de Hevia et al., 2014). In some conditions, newborns and infants are also able to create mappings between number and brightness (de Hevia and Spelke, 2013; Bonn et al., 2019) and brightness and loudness (Lewkowicz and Turkewitz, 1980), supporting the view that number-spatial extent mapping is not specific to numerical information and extends to other quantitative dimensions. However, while at birth and during infancy humans associate lateralized spatial locations (left vs. right) to different numerosities (small vs. large, respectively), this number-spatial location mapping does not generalize to dimensions other than number, like size and brightness (Bulf et al., 2016; de Hevia et al., 2017), supporting the view that these two number-space mappings reflect distinct cognitive phenomena, at least early in life.

Second, a subcomponent of the number-spatial position mapping, which reflects the spontaneous tendency to mentally organize ordered information along a spatially oriented axis, is extended to any ordered dimension even in infancy (Bulf et al., 2017; Bulf et al., 2022). This type of number-spatial location mapping differs from the lateralized one (i.e., left-small/right-large) in that while the first one is related to the ordinality of the information (i.e., first, second, third. . .), the latter one, as it is also the case for the number-spatial extent mapping, is tied to the information of magnitude (i.e., smaller vs. larger). The fact that numbers can reflect these two properties at the same time, ordinality and cardinality, partially explains why authors often consider that any numerical-spatial effect emanates from a single representational object. However, analogous effects for non-numerical information, such as a SNARC-like effect for items in a grocery list (Previtali et al., 2010), are hardly accounted for by invoking a mental number line.

Finally, the question of an absence or not of a dependency between the number-spatial extent and the number-spatial location mappings has not received, to our knowledge, explicit attention by researchers. However, in one experiment from a larger study investigating the automaticity of the activation of numerical magnitude, Fitousi et al. (2009) observed that the SNARC effect was independent of the SCE in a task where participants had to judge the physical size of Arabic digits, again going against the idea of both number-space mappings reflecting a single construct in adulthood.

The present study is aimed at further investigating the interaction between the number-spatial extent and the number-spatial location interactions within a single experimental session. If both mappings reflect the same psychological construct, i.e., the same representational object (e.g., large number = large size = right side of space), they should interact with each other. As a consequence, the congruency effect related to one mapping will be modulated by the congruency relative to the other mapping. For instance, if the number/location (resp. number/spatial extent) mapping is congruent, then there will be a number/spatial extent (resp. number/location) mapping, so that congruent trials will lead to higher performance relative to incongruent trials. For number/location (resp. number/spatial extent) incongruent mappings, no difference in performance should be observed between number/spatial extent (resp. number/location) congruent and incongruent trials. However, if both mappings reflect distinct representational objects, the congruency relative to one mapping should not impact the congruency effect related to the other mapping. If we assume that both congruency effects are unaffected by each other, this should in turn, lead to an overall additivity pattern, whereby the highest performance is observed in trials showing a congruency for both mappings, and the lowest performance is observed in trials showing an incongruency for both mappings (**Figure 1**).

It is possible that the interaction between the two mappings depends on the format of presentation of numbers. According to the infant literature mentioned above, we hypothesize distinct mappings with non-symbolic representations of numbers. However, with education, digits could be mapped onto multiple spatial dimensions (extent and location) in a more holistic manner, leading to stronger interactions between both mappings when processing symbolic numbers.

To test these predictions, we designed a numerical judgment task containing both number-extent and number-location congruent and incongruent trials. In order to directly assess the role of the format of presentation, each participant performed both a non-symbolic and a symbolic version of the task.

# EXPERIMENT 1

## Methods

### Participants

We recruited 77 adult participants using the Prolific online recruitment platform [47 males (2 data missing), mean age = 30.64 years, *SD* = 9.49 (4 data missing)][1]. They each received a 2.6 euros compensation for their participation. We ran an *a priori* power analysis using G*Power, Version 3.1; Faul et al., 2009 in order to estimate the sample size. The analysis indicated that a sample size of 54 was necessary to detect a medium effect size (Cohen's $f = 0.25$) in a 2 Format (symbolic, non-symbolic) × 2 Location Congruency (congruent, incongruent) × 2 Size Congruency (congruent, incongruent) repeated measures ANOVA with a power of 0.95 ($\alpha = 0.05$).

The internal ethical board of the Faculty of Psychology ruled that in light of the potential risks for the participants of the present study, no formal ethical approval by one of the National Ethical Committees was needed in agreement with the Ethical law governing human research in France. Participation was voluntary after obtaining signed informed consent. All participants were tested in accordance with national and international norms governing the use of human research participants.

### Materials and Procedure

The task was programmed using the jsPsych JavaScript library (de Leeuw, 2015), and the data was collected online through the Cognition.run website. The participants were asked to perform symbolic and non-symbolic numerical comparison tasks. In the symbolic comparison task, Arabic digits (1, 2, 3, 4, 5, 7, 8, 9, 10, 11) were presented at the center of the screen and the participants had to decide whether the numbers were smaller or larger than 6, by pressing either the "d" or the "k" key on their computer's keyboard. In the non-symbolic comparison task, participants had to decide whether the number of dots (6, 7, 8, 9, 10, 40, 44, 48, 52, 56) in centrally presented arrays was smaller or larger than 20, by pressing the "d" or "k" key. In order for the symbolic and non-symbolic tasks to present similar difficulty levels, while avoiding the subitizing range for the non-symbolic task, we used a 1:2 ratio between the reference (20) and the two immediately smaller (10) and larger (40) numbers. This allowed us to use five numbers above the subitizing range (6, 7, 8, 9, 10) and five numbers with matching ratios to 20 (40, 44, 48, 52, and 56). Using the jsPsych calibration plugin, the Arabic digits were set to be presented at a fixed size of either 1 cm × 0.6 cm (small size) or 2 cm × 1.2 cm (large size) at the center of the participants' screen, while the images of the dot arrays were set to be presented at a fixed size of either 5.8 cm × 5.8 cm (small size) or 11.6 cm × 11.6 cm (large size) at the center of the participants' screen. The arrays of dots were generated using Matlab. The dots were randomly arranged on the surface of the images, and had a constant size. Thus, the large images showed arrays of dots twice the size of the dots in the small images, and occupying a space twice as large. Half of the stimuli (Arabic digits and dot arrays) were presented in blue over a black background, while the other half was red over a black background. The color and physical size of the stimuli were counterbalanced across both tasks. Each task (symbolic and non-symbolic) consisted of two blocks, one block for each response/key assignment. The order of the tasks was counterbalanced across participants. The order of the response/key assignments ("larger" on the right first, or "larger" on the left first) was constant across both tasks, and counterbalanced across the participants. In both tasks, a trial started with the presentation of a black screen for 500 ms, followed by a central fixation point for 1 s. Then, the stimulus was presented until the participants gave their response. If no response was given, the next trial began after 5 s of stimulus' presentation (**Figure 2**). Each task started with eight training trials, during which a feedback on accuracy ("correct" printed in green or "incorrect" printed in red) was given for 1 s, followed by a 500 ms black screen. Each block contained 80 trials, half of which were size-congruent (e.g., large number in large physical

---

[1] www.prolific.co

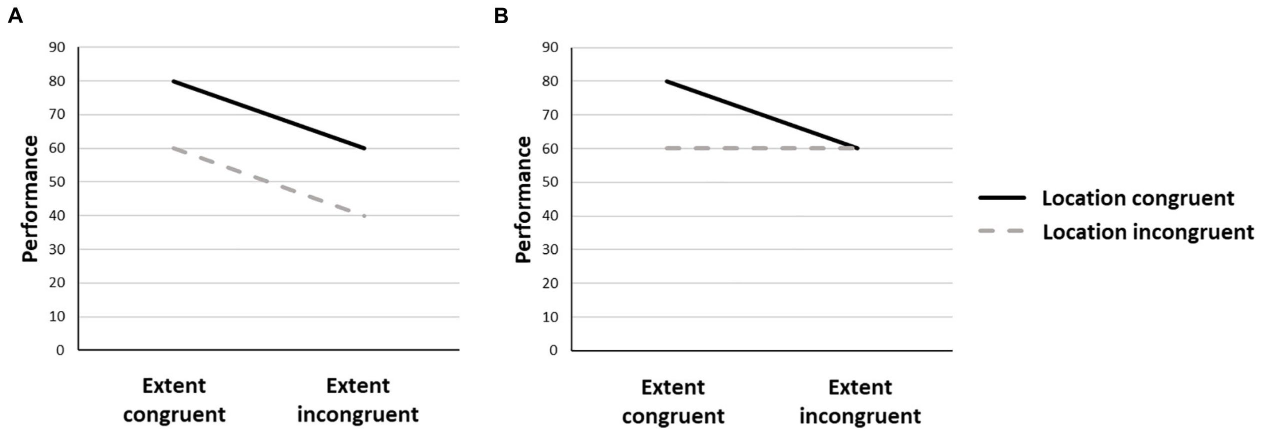


**FIGURE 1 |** Predicted patterns of performances in the case of distinct **(A)** or single **(B)** representational objects accounting for the number/location and number/spatial extent mappings.

size), and the other half were size-incongruent. For each task, one block consisted of location-congruent trials ("small" response on the left, "large" response on the right), while the other block consisted of location-incongruent trials. This led to a total of 320 trials across the two tasks, for a duration of approximately 15 min. The participants were invited to take short breaks if needed between each block and each task.

## Results

The data from 11 participants were removed due to an accuracy rate (AR) below 75% in at least one of the tasks' blocks. Thus, the reported analyses were performed on a sample of 66 participants. For each participant, we removed the trials with response times either below 150 ms, or larger than three standard deviations above the individual's average. This led to removing 1.71% of the total number of trials across the analyzed sample. For each participant, we combined ARs and RTs by computing an inverse efficiency score (IES; Townsend and Ashby, 1978, 1983), that is, by dividing participants' average RTs by their average ARs for each of the eight experimental conditions (2 task × 2 size congruency × 2 location congruency). Note that, in line with the recommendations of Bruyer and Brysbaert (2011), the use of IES was possible due to the high ARs in the tasks (mean = 0.96, *SD* = 0.04 in Experiment 1, mean = 0.96, *SD* = 0.04 in Experiment 2), and the absence of speed-accuracy trade-offs [$r(65) = 0.1$, $p = 0.45$ in Experiment 1, $r(67) = 0.05$, $p = 0.68$ in Experiment 2].

We ran a 2 Task (symbolic, non-symbolic) × 2 Size Congruency (size congruent, size incongruent) × 2 Location Congruency (location congruent, location incongruent) repeated measures ANOVA on the IES. We observed a main effect of Task [$F(1,65) = 4.89$, $p = 0.03$, $\eta^2_p = 0.07$], with lower performance in the symbolic (*IES* = 601, *SD* = 165) than in the non-symbolic task (*IES* = 577, *SD* = 143). The results also showed a classic main effect of Size Congruency [$F(1,65) = 17.21$, $p < 0.001$, $\eta^2_p = 0.21$], with lower performance in the size incongruent (*IES* = 594, *SD* = 153) than in the size congruent (*IES* = 583, *SD* = 156) trials. There was a significant Task × Location Congruency interaction [$F(1,65) = 7.41$, $p = 0.008$, $\eta^2_p = 0.1$]. *Post hoc* Bonferroni corrected comparisons indicated a trend toward an effect of Location congruency ($p = 0.19$) in the symbolic, but not in the non-symbolic task ($p = 1$). Critically, there was a significant difference between the performance in location incongruent trials between the non-symbolic and the symbolic tasks ($p = 0.01$), but not in location congruent trials ($p = 1$), showing that location incongruency significantly worsened performance with Arabic digits but not with dot arrays. There were no other significant effects. In particular, we did not observe a significant interaction involving Size and Location Congruency (*Fs* < 1.07, *ps* > 0.31, **Figure 3**), in line with the prediction in **Figure 1A**. We verified that the same analyses run on RTs yielded similar results, with a significant main effect of Size Congruency [$F(1,65) = 15.98$, $p < 0.001$], and a significant Task × Location Congruency interaction [$F(1,65) = 5.40$, $p = 0.02$]. We only found a significant main effect of Size Congruency when analyzing ARs [$F(1,65) = 6.318$, $p = 0.014$], which could possibly be due to ceiling effects.

## Interim Discussion

We found an effect of congruency between the numerical magnitude and the physical magnitude, for both symbolic and non-symbolic stimuli, so that congruent trials (i.e., when larger numerical quantities were larger in size, and smaller numerical quantities were smaller in size) had a higher performance than incongruent ones (i.e., when numerical quantities and their sizes differed, one being small and the other large). The overall lower performance with digits than with non-symbolic numbers could be due to the ratio differences, which were more pronounced for non-symbolic stimuli, and therefore comparison was easier and faster for this numerical format.

We observed a number/location congruency effect (SNARC) only for digits. While performance for dots was not affected by the location of the response side (no cost nor advantage for location-incongruent and location-congruent trials, respectively), and performance was similar to that for location-congruent trials

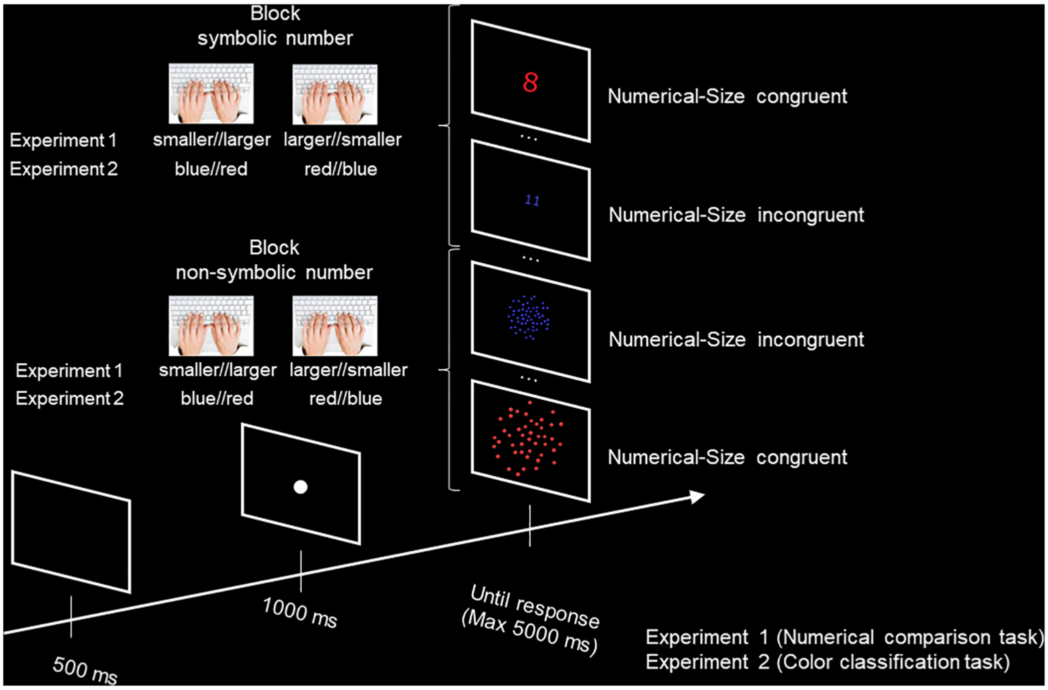


**FIGURE 2 |** Illustration of the methodology used in Experiment 1 (Numerical comparison task) and Experiment 2 (Color classification task).

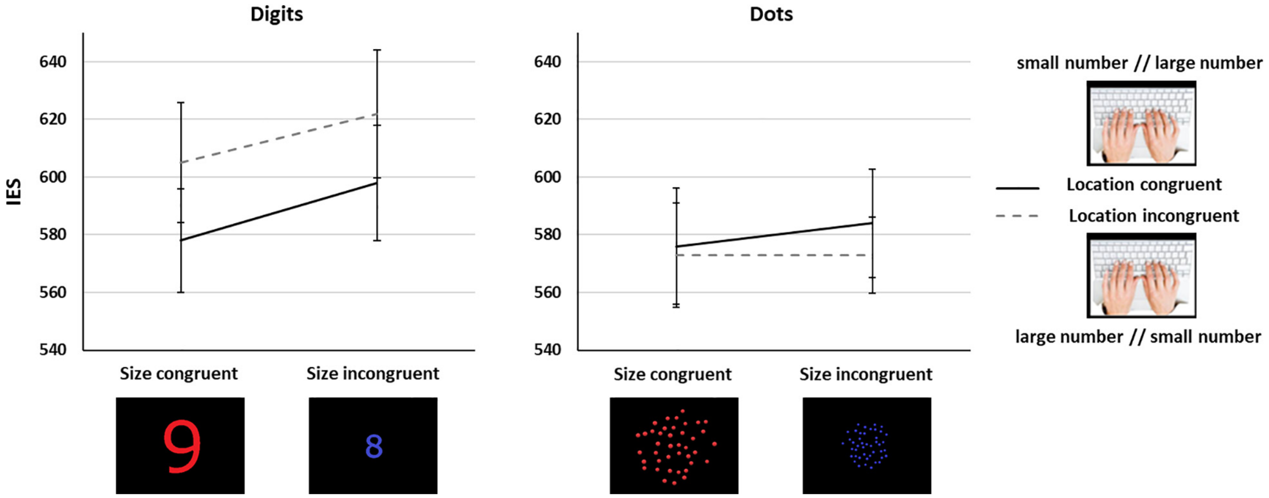


**FIGURE 3 |** Interaction between number/location and number/spatial extent mappings in Experiment 1 for the symbolic (left) and the non-symbolic (right) tasks. Error bars show standard error of the mean (SEM).

in digits, there was a significant performance cost for digit trials, compared to dot trials, in location-incongruent trials. Therefore, the spatial location of the response button affected symbolic but not non-symbolic numerical comparisons. Studies investigating non-symbolic SNARC effects are scarcer than in the symbolic domain, with mixed results (Cleland et al., 2020). However, a few studies in adults have reported SNARC or SNARC-like effects with non-symbolic numbers, especially when numerical magnitude was relevant to the task (Zhou et al., 2016; Nemeh et al., 2018), in contradiction with our results. The heightened importance of spatial location of response buttons for Arabic digits might derive from the fact that adults have abundant exposure to digits arranged on a horizontal (left-right oriented) space, while this spatial layout is rarely, if ever, observed with non-symbolic numerosities. Finally, there was no interaction between the two types of number-spatial mappings, suggesting that both mappings act independently.

Quantity might have a more relevant role for the SCE, while the SNARC effect might be boosted in a task highlighting order. Since we used an explicit numerical comparison task, it is possible that the number-spatial extent mapping was enhanced, in detriment of the number-spatial position mapping, and might account for the absence of location congruency effects for non-symbolic numbers. These results are in line with Fitousi et al.'s (2009) study, in which the size judgment task was also emphasizing the number-spatial extent mapping. Additionally, in our task, the number/spatial extent and number/location congruencies varied at different levels. While spatial extent congruency was manipulated at the trial level, location congruency varied between blocks (based on the response-key assignment). Since we aimed at keeping the task instructions identical across the two numerical formats, we ran the same experiment, except that we engaged participants in implicit non-symbolic and symbolic numerical tasks, by asking them to judge the color of the numerical stimuli. In this way, we avoided favoring a mental representation in terms of quantity or order, and congruencies relative to spatial extent and location both varied at the trial level.

# EXPERIMENT 2

## Methods

### Participants

We recruited 73 new adult participants using the Prolific online recruitment platform (see text footnote 1; 58 males, mean age = 26.56 years, *SD* = 6.92). They each received a 2.6 euros compensation for their participation.

The internal ethical board of the Faculty of Psychology ruled that in light of the potential risks for the participants of the present study, no formal ethical approval by one of the National Ethical Committees was needed in agreement with the Ethical law governing human research in France. Participation was voluntary after obtaining signed informed consent. All participants were tested in accordance with national and international norms governing the use of human research participants.

### Materials and Procedure

The materials and procedure were the same as in Experiment 1, except for the instructions. The participants performed a symbolic and a non-symbolic task, in which they were instructed to decide whether the presented stimulus was blue or red, by pressing one of the two keys on their keyboard. Given these new instructions, the location-congruency of the trials was now counterbalanced within each block. For instance, if the response/key assignment was "red on the left and blue on the right," half of the large numbers (the blue ones) were responded to with the right-hand key (location-congruent), while the other half (the red ones) were responded to with the left-hand key (location-incongruent).

## Results

As in Experiment 1, we removed the data from five participants whose ARs were below 75% in at least one of the experimental blocks, leading to an analyzed sample of 68 participants[2]. 1.98% of the entire dataset was removed due to RTs faster than 150 ms or slower than 3 standard deviations above the individual's average. We ran a 2 Task (symbolic, non-symbolic) × 2 Size Congruency (size congruent, size incongruent) × 2 Location Congruency (location congruent, location incongruent) repeated measures ANOVA on the IES.

We did not observe main effects of Task nor Size Congruency ($Fs < 1$). There was only a trend toward a main effect of Location Congruency [$F(1,67) = 2.38$, $p = 0.13$]. None of the interactions involving Task reached significance ($Fs < 1.59$, $ps > 0.21$). Importantly, and contrary to what was observed in Experiment 1, we found a significant interaction between Size and Location Congruency [$F(1,67) = 6.52$, $p = 0.01$, $\eta^2_p = 0.09$, **Figure 4**]. *Post hoc* Bonferroni corrected comparisons revealed a significant Location Congruency effect only for size congruent trials (location incongruent: *IES* = 494, *SD* = 161, location congruent: *IES* = 485, *SD* = 143, $t = 2.92$, $p = 0.03$), but not for size incongruent trials (location incongruent: *IES* = 489, *SD* = 151, location congruent: *IES* = 491, *SD* = 152, $t = 0.8$, $p = 1$), in line with the prediction in **Figure 1B**. Regarding Size congruency, although performance was lower for size incongruent than size congruent trials only in the location congruent trials, this difference did not reach statistical significance ($p = 0.36$; all other contrasts were not significant, all $ps > 0.6$). As in Experiment 1, we observed similar results with only a significant Size × Location Congruency interaction [$F(1,67) = 4.48$, $p = 0.04$] when using RTs, and no significant effect when using ARs.

## Interim Discussion

When using an implicit task, i.e., a color judgment task, participants did not show anymore a main effect of size congruency, possibly because this effect was weakened when avoiding explicit magnitude coding. However, and in contrast to Experiment 1, we found a location congruency effect only in size congruent trials, so that there was an advantage in performance for smaller quantities responded on the left and larger quantities on the right when the size of stimuli aligned with their numerical magnitude. This effect was present irrespective of the numerical format, involving both symbolic and non-symbolic trials.

# GENERAL DISCUSSION

In this study we aimed at directly investigating the question of whether there is an interaction between two number-space mappings, i.e., the number/spatial extent and the number/spatial location mappings. We reported that, when using a numerical

---

[2] Although power analysis indicated a sample size of 54, no sequential testing was performed and data were only analyzed once all the data were collected ($N = 77$ in Experiment 1, $N = 73$ in Experiment 2). In order to test for the robustness of the interaction in Experiment 2 vs. the absence of interaction in Experiment 1, we ran 100 analyses on random sub-samples of 54 participants. The significant Location × Size Congruency interaction reported in Experiment 2 was observed 74 out of 100 times ($p < 0.001$, binomial test), while it was significant only 1 out of 100 times in Experiment 1. Moreover, the significant Task × Location Congruency interaction reported in Experiment 1 was observed 87 out of 100 times ($p < 0.001$, binomial test).

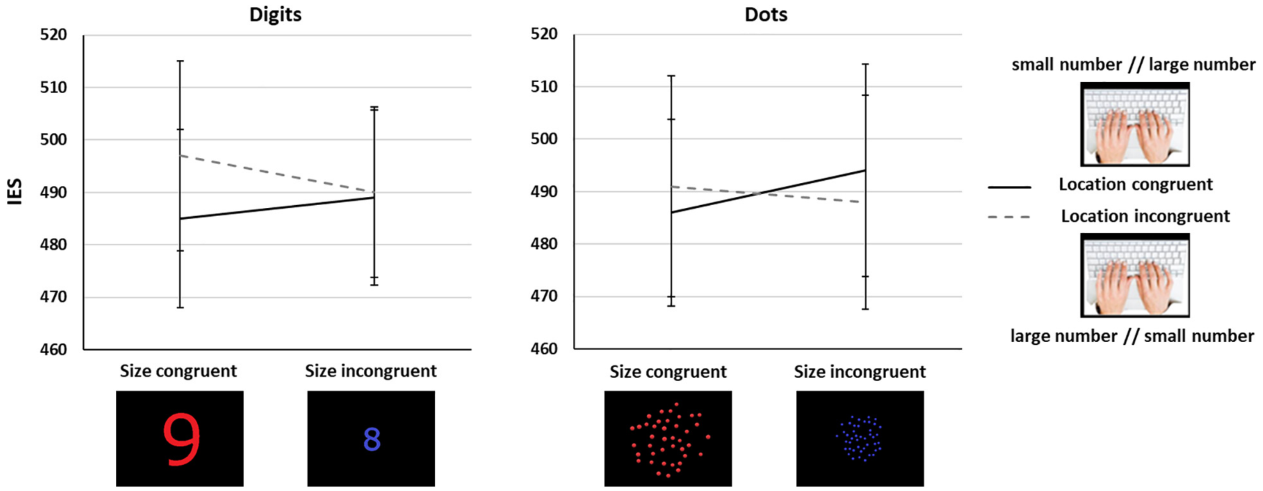


**FIGURE 4 |** Interaction between number/location and number/spatial extent mappings in Experiment 2 for the symbolic (left) and the non-symbolic (right) tasks. Error bars show standard error of the mean (SEM).

magnitude judgment task, the two mappings did not interact: the difference in performance between congruent and incongruent trials relative to one mapping was the same for congruent and incongruent trials relative to the other mapping (in line with the prediction **Figure 1A**). However, in the context of an implicit task, i.e., a color judgment task, using the same stimuli, the two mappings interacted, with stronger congruency effect relative to one mapping when the trials were congruent relative to the other mapping (in line with the prediction **Figure 1B**).

Altogether, our results suggest that both the number/extent and the number/location mappings can tap onto a shared representational object, but that its activation critically depends on the experimental context, and can manifest itself in different ways. Several factors could be at play. First, the reliance on a shared representational object for both mappings could depend on whether magnitude is implicitly or explicitly activated by the task. Both in Fitousi et al. (2009) and in the current study, using a task whereby magnitude (either numerical or non-numerical) was explicitly activated yielded no interaction between the two mappings. As mentioned previously, it is possible that the interaction between both mappings depends on how much the experimental context emphasizes one mapping over the other. By having instructions based on numerical magnitude or physical size, the number/spatial extent mapping might be more activated, preventing the use of a shared representational object. On a related note, previous studies have found that the nature of the number/location mappings, reflected by the SNARC effect, could differ depending on the task (Gevers et al., 2006), with a more categorical association observed in the context of a numerical magnitude judgment task, and a more continuous association in the context of a parity judgment task. It is possible that the two mappings only interact in an experimental context typically eliciting a more continuous number/location mapping.

Second, for an interaction to be observed between the two mappings, another factor could be the level at which the different types of information are being manipulated. When using a magnitude-relevant task, the number/location mapping can only be analyzed by comparing different blocks of trials, corresponding to the different response-key assignments (e.g., "more on the right" vs. "more on the left"). Using a task whereby numerical magnitude is irrelevant, such as in Experiment 2, allows us to analyze the number/location mapping across trials (similarly to the analysis of the number/extent mapping), since the numerical magnitude is independent of the response side. This methodological factor could also contribute to balancing out the weight of each mapping, yielding to their interaction, as observed in Experiment 2. Further studies will be needed to investigate the exact contextual conditions that elicit an interaction between the number/extent and the number/location mappings.

Our results did not show any evidence for an effect of format on the interaction between the two number-space mappings. While Experiment 1 showed an effect of the format of presentation (symbolic vs. non-symbolic) on the number-location mapping, in both experiments the interaction between the two mappings did not differ significantly depending on the format. This could indicate a continuity in the number-space associations between non-symbolic and symbolic representations of magnitude. Developmental studies would help to shed more light on this question.

Altogether, our results support the idea of the existence of a shared spatial representational object with a directionality (from left-to-right in our group of participants), and on which smaller numbers are associated with smaller spatial extents, and larger numbers with larger spatial extents (de Hevia, 2021). In this regard, the mental number line could account both for the number/location and the number/extent mappings. However, and in line with previous studies investigating the SNARC effect in particular (Viarouge et al., 2014), we see that this shared representation can be activated flexibly depending on the experimental context. Thus, while we provide evidence for the existence of a shared representation, our data suggests that its activation is not automatic in any task using numerical stimuli.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors upon request, without undue reservation.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AV and MH equally contributed to the conception and design of the study, wrote the first draft of the manuscript, revised it, and approved the submitted version. AV performed the statistical analysis. Both authors contributed to the article and approved the submitted version.

## FUNDING

AV was supported by an IdEx Université de Paris "Emergence en Recherche" grant. MH was supported by an Agence Nationale de la Recherche Scientifique grant (NUMACT-18-CE28-0017-01). This study contributed to the IdEx Université de Paris ANR-18-IDEX-0001.

## REFERENCES

Adachi, I. (2014). Spontaneous spatial mapping of learned sequence in chimpanzees: evidence for a SNARC-like effect. *PLoS One* 9:e90373. doi: 10.1371/journal.pone.0090373

Bonn, C. D., Netskou, M.-E., Streri, A., and de Hevia, M. D. (2019). The association of brightness with number/duration in human newborns. *PLoS One* 14:e0223192. doi: 10.1371/journal.pone.0223192

Bruyer, R., and Brysbaert, M. (2011). Combining Speed and Accuracy in Cognitive Psychology: is the Inverse Efficiency Score (IES) a Better Dependent Variable than the Mean Reaction Time (RT) and the Percentage Of Errors (PE)? *Psychol. Belg.* 51, 5–13. doi: 10.5334/pb-51-1-5

Bulf, H., Capparini, C., Nava, E., de Hevia, M. D., and Macchi Cassia, V. (2022). Space modulates cross-domain transfer of abstract rules in infants. *J. Exp. Child Psychol.* 213:105270. doi: 10.1016/j.jecp.2021.105270

Bulf, H., de Hevia, M. D., Gariboldi, V., and Macchi Cassia, V. (2017). Infants learn better from left to right: a directional bias in infants' sequence learning. *Sci. Rep.* 7:2437. doi: 10.1038/s41598-017-02466-w

Bulf, H., de Hevia, M. D., and Macchi Cassia, V. (2016). Small on the left, large on the right: numbers orient visual attention onto space in preverbal infants. *Dev. Sci.* 19, 394–401. doi: 10.1111/desc.12315

Bulf, H., Macchi Cassia, V., and de Hevia, M. D. (2014). Are numbers, size and brightness equally efficient in orienting visual attention? Evidence from an eye-tracking study. *PLoS One* 9:e99499. doi: 10.1371/journal.pone.0099499

Burr, D., and Ross, J. (2008). A Visual Sense of Number. *Curr. Biol.* 18, 425–428. doi: 10.1016/j.cub.2008.02.052

Cantlon, J. F. (2018). How Evolution Constrains Human Numerical Concepts. *Child Dev. Perspect.* 12, 65–71. doi: 10.1111/cdep.12264

Cipora, K., Haman, M., Domahs, F., and Nuerk, H.-C. (2020). Editorial: on the Development of Space-Number Relations: linguistic and Cognitive Determinants, Influences, and Associations. *Front. Psychology* 11:182. doi: 10.3389/fpsyg.2020.00182

Cleland, A. A., Corsico, K., White, K., and Bull, R. (2020). Non-symbolic numerosities do not automatically activate spatial–numerical associations: evidence from the SNARC effect. *Q. J. Exp. Psychol.* 73, 295–308. doi: 10.1177/1747021819875021

de Hevia, M. D. (2021). How the human mind grounds numerical quantities on space. *Child Dev. Perspect.* 15, 44–50. doi: 10.1111/cdep.12398

de Hevia, M. D., Girelli, L., Bricolo, E., and Vallar, G. (2008). The representational space of numerical magnitude: illusions of length. *Q. J. Exp. Psychol.* 61, 1496–1514. doi: 10.1080/17470210701560674

de Hevia, M. D., Girelli, L., and Vallar, G. (2006). Numbers and space: a cognitive illusion? *Exp. Brain Res.* 168, 254–264. doi: 10.1007/s00221-005-0084-0

de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., and Streri, A. (2014). Representations of space, time, and number in neonates. *Proc. Natl. Acad. Sci. U. S. A.* 111, 4809–4813. doi: 10.1073/pnas.1323628111

de Hevia, M. D., and Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition* 110, 198–207. doi: 10.1016/j.cognition.2008.11.003

de Hevia, M. D., and Spelke, E. S. (2010). Number-space mapping in human infants. *Psychol. Sci.* 21, 653–660. doi: 10.1177/0956797610366091

de Hevia, M. D., and Spelke, E. S. (2013). Not All Continuous Dimensions Map Equally: number-Brightness Mapping in Human Infants. *PLoS One* 8:e81241. doi: 10.1371/journal.pone.0081241

de Hevia, M. D., Veggiotti, L., Streri, A., and Bonn, C. D. (2017). At Birth, Humans Associate "Few" with Left and "Many" with Right. *Curr. Biol.* 27, 3879–3884.e2. doi: 10.1016/j.cub.2017.11.024

de Leeuw, J. R. (2015). jsPsych: a JavaScript library for creating behavioral experiments in a web browser. *Behav. Res. Methods* 47, 1–12. doi: 10.3758/s13428-014-0458-y

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition* 44, 1–42. doi: 10.1016/0010-0277(92)90049-n

Dehaene, S. (2009). Origins of mathematical intuitions: the case of arithmetic. *Ann. N. Y. Acad. Sci.* 1156, 232–259. doi: 10.1111/j.1749-6632.2009.04469.x

Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371

Di Giorgio, E., Lunghi, M., Rugani, R., Regolin, L., Dalla Barba, B., Vallortigara, G., et al. (2019). A mental number line in human newborns. *Dev. Sci.* 22:e12801. doi: 10.1111/desc.12801

Drucker, C. B., and Brannon, E. M. (2014). Rhesus monkeys (*Macaca mulatta*) map number onto space. *Cognition* 132, 57–67. doi: 10.1016/j.cognition.2014.03.011

Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149

Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002

Fias, W. (1996). The Importance of Magnitude Information in Numerical Processing: evidence from the SNARC Effect. *Math. Cogn.* 2, 95–110.

Fischer, M. H. (2001). Cognition in the bisection task. *Trends Cogn. Sci.* 5, 460–462. doi: 10.1016/S1364-6613(00)01790-3

Fischer, M. H., Castel, A. D., Dodd, M. D., and Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* 6, 555–556. doi: 10.1038/nn1066

Fitousi, D., Shaki, S., and Algom, D. (2009). The role of parity, physical size, and magnitude in numerical cognition: the SNARC effect revisited. *Percept. Psychophys.* 71, 143–155. doi: 10.3758/APP.71.1.143

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., and Fias, W. (2006). Numbers and space: a computational model of the SNARC effect. *J. Exp. Psycho. Hum. Percept. Perform.* 32, 32–44. doi: 10.1037/0096-1523.32.1.32

Girelli, L., Lucangeli, D., and Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *J. Exp. Child Psychol.* 76, 104–122. doi: 10.1006/jecp.2000.2564

Halberda, J., Mazzocco, M. M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–668. doi: 10.1038/nature07246

Lewkowicz, D. J., and Turkewitz, G. (1980). Cross-modal equivalence in early infancy: auditory–visual intensity matching. *Dev. Psychol.* 16, 597–607. doi: 10.1037/0012-1649.16.6.597

Nemeh, F., Humberstone, J., Yates, M. J., and Reeve, R. A. (2018). Non-symbolic magnitudes are represented spatially: evidence from a non-symbolic SNARC task. *PLoS One* 13:e0203019. doi: 10.1371/journal.pone.0203019

Odic, D., and Starr, A. (2018). An Introduction to the Approximate Number System. *Child Dev. Perspect.* 12, 223–229. doi: 10.1111/cdep.12288

Patro, K., and Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006

Pitt, B., and Casasanto, D. (2020). The correlations in experience principle?: how culture shapes concepts of time and number. *J. Exp. Psychol. Gen.* 149, 1048–1070. doi: 10.1037/xge0000696

Previtali, P., de Hevia, M. D., and Girelli, L. (2010). Placing order in space: the SNARC effect in serial learning. *Exp. Brain Res.* 201, 599–605. doi: 10.1007/s00221-009-2063-3

Restle, F. (1970). Speed of adding and comparing numbers. *J. Exp. Psychol.* 83, 274–278. doi: 10.1037/h0028573

Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science* 347, 534–536. doi: 10.1126/science.aaa1379

Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328

Srinivasan, M., and Carey, S. (2010). The long and the short of it: on the nature and origin of functional overlap between representations of space and time. *Cognition* 116, 217–241. doi: 10.1016/j.cognition.2010.05.005

Starr, A., Libertus, M. E., and Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proc. Natl. Acad. Sci. U. S. A.* 110, 18116–18120. doi: 10.1073/pnas.1302751110

Townsend, J. T., and Ashby, F. G. (1978). "Methods of modeling capacity in simple processing systems," in *Cognitive Theory,* Vol. 3, eds J. Castellan, and F. Restle (Hillsdale, NJ: Lawrence Erlbaum), 200–239.

Townsend, J. T., and Ashby, F. G. (1983). *Stochastic Modeling of Elementary Psychological Processes*. Cambridge: Cambridge University Press.

Viarouge, A., and de Hevia, M. D. (2013). The role of numerical magnitude and order in the illusory perception of size and brightness. *Cognition* 4:484. doi: 10.3389/fpsyg.2013.00484

Viarouge, A., Hubbard, E. M., and Dehaene, S. (2014). The Organization of Spatial Reference Frames Involved in the Snarc Effect. *Q. J. Exp. Psychol.* 67, 1484–1499. doi: 10.1080/17470218.2014.897358

Zhou, X., Shen, C., Li, L., Li, D., and Cui, J. (2016). Mental Numerosity Line in the Human's Approximate Number System. *Exp. Psychol.* 63, 169–179. doi: 10.1027/1618-3169/a000324

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 06 October 2021
doi: 10.3389/fnhum.2021.745188

# The Grouping-Induced Numerosity Illusion Is Attention-Dependent

*Antonella Pomè[1]*, Camilla Caponi[1] and David C. Burr[1,2]*

[1]*Department of Neuroscience, Psychology, Pharmacology, and Child Health, University of Florence, Florence, Italy,* [2]*School of Psychology, University of Sydney, Sydney, NSW, Australia*

Perceptual grouping and visual attention are two mechanisms that help to segregate visual input into meaningful objects. Here we report how perceptual grouping, which affects perceived numerosity, is reduced when visual attention is engaged in a concurrent visual task. We asked participants to judge the numerosity of clouds of dot-pairs connected by thin lines, known to cause underestimation of numerosity, while simultaneously performing a color conjunction task. Diverting attention to the concomitant visual distractor significantly reduced the grouping-induced numerosity biases. Moreover, while the magnitude of the illusion under free viewing covaried strongly with AQ-defined autistic traits, under conditions of divided attention the relationship was much reduced. These results suggest that divided attention modulates the perceptual grouping of elements by connectedness and that it is independent of the perceptual style of participants.

**Keywords: numerosity perception, attention, segmentation, autistic quotient, grouping, connectedness**

***Edited by:***
*Joonkoo Park,*
*University of Massachusetts*
*Amherst, United States*

***Reviewed by:***
*Kyle R. Cave,*
*University of Massachusetts*
*Amherst, United States*
*Gisella Decarli,*
*Integrative Neuroscience and*
*Cognition Center (INCC) Université*
*de Paris, France*
*Sandarsh Pandey,*
*University of Massachusetts*
*Amherst, United States in*
*collaboration with reviewer KC*

****Correspondence:***
*Antonella Pomè*
*antonella.pome@unifi.it*

***Specialty section:***
*This article was submitted to*
*Cognitive Neuroscience,*
*a section of the journal*
*Frontiers in Human Neuroscience*

***Received:*** *21 July 2021*
***Accepted:*** *16 September 2021*
***Published:*** *06 October 2021*

***Citation:***
*Pomè A, Caponi C and Burr DC*
*(2021) The Grouping-Induced*
*Numerosity Illusion Is*
*Attention-Dependent.*
*Front. Hum. Neurosci. 15:745188.*
*doi: 10.3389/fnhum.2021.745188*

## INTRODUCTION

To make sense of visual scenes, meaningful perception relies on our ability to quickly and efficiently organize visual information. The visual system groups elements using principles first introduced by Gestalt psychologists, including similarity, proximity, closure, and *common fate* (Wertheimer, 1923). This allows incoming information to be organized and integrated into coherent, whole objects, separate from the backgrounds. Selective attention is another process that influences how we perceive visual information. Attention and perceptual organization are interconnected, affecting visual processing and performance in various tasks and conditions.

Attentional demands in grouping have been investigated over the past decades, but the conclusions have been inconsistent (Ben-Av et al., 1992; Mack et al., 1992; Moore and Egeth, 1997; Kimchi and Razpurker-Apfeld, 2004; Lamy et al., 2006; see Kimchi, 2009, for a review). For example, when observers are engaged in an attentionally demanding task they are unable to report grouping organizations presented in the unattended backdrop of the task-relevant stimulus (Mack et al., 1992). Along these lines, apparent perceptual organization (luminance similarity) of a multielement array is intensified when attended and attenuated when unattended (Barbot et al., 2018), both suggesting that perceiving organization requires attention. In contrast, research using visual illusions whose susceptibility depends on grouping incoming information shows that individuals are susceptible to grouping even when they are unable to explicitly report the elements being grouped (Moore and Egeth, 1997; Driver et al., 2001; Russell and Driver, 2005; Lamy et al., 2006; Kimchi and Peterson, 2008; Shomstein et al., 2010; Carther-Krone et al., 2016). This view is further supported by studies involving patients with neurological disabilities such as hemispatial neglect (Russell and Driver, 2005; Shomstein et al., 2010) and simultanagnosia (Karnath et al., 2000; Huberle and Karnath, 2006), who can implicitly group elements despite difficulties explicitly reporting the global configuration.

In sum, the relationship between perceptual organization and attention is complicated: whereas some forms of perceptual organization can occur without attention (Braun and Sagi, 1991; Kimchi and Peterson, 2008), attention can nevertheless modulate perceptual organization processes.

In the present study, we used the numerosity illusion of *connectedness* to measure perceptual grouping. This illusion taps grouping mechanisms indirectly, without requiring participants to report directly the perceptual organization. When visual items such as circles or squares are grouped together with connecting lines, they appear less numerous (Franconeri et al., 2009; He et al., 2009; Anobile et al., 2017; Pomè et al., 2021a). The connecting lines are equally effective when very thin (Franconeri et al., 2009), or even when illusory (Kirjakovski and Matsumoto, 2016; Adriano et al., 2021). This has been taken as evidence that numerosity operates on segmented objects, defined by grouping, rather than individual local elements. For densely packed items, the effect of connectivity is greatly reduced (Anobile et al., 2017), showing that the effect is limited to the numerosity range of estimation of segregable items, rather than judgments of texture density. It also affects fMRI responses to numbers (He et al., 2015), adaptation to numbers (Fornaciai et al., 2016), and pupillometry (Castaldi et al., 2021).

Moreover, we recently demonstrated that the magnitude of the effect varies according to the perceptual styles of the participants: those scoring high on the self-reported Autistic Quotient questionnaire (AQ) showed a reduced illusory effect compared with participants with lower autistic traits. This is in line with theories that have linked autism with local processing and reduced awareness of the global aspects of stimuli (Pomè et al., 2021a).

The current work investigates whether grouping by connectedness can occur without attention being freely available to judge the numerosity of the stimulus. Recent evidence has shown that depriving visual attentional resources by a concomitant visual or auditory dual-task result in a higher cost in number representation in the small, *subitizing* number range than for larger numerosities (Burr et al., 2010; Pomè et al., 2019). Furthermore, the presentation of a visual cue that increases attentional engagement in a given task facilitates estimation, leading to a less compressive representation of numbers in space compared to when attention is diverted elsewhere (Pomè et al., 2021b). Numerosity benefits from object-based attentional resources, as cuing anywhere within an object gives the same attentional advantage as cuing the precise location of the object, suggesting that attention to number spreads from the cued location throughout the whole cued object (Pomè et al., 2021b).

To investigate the dependence of grouping on attention, we measured the strength of the connectedness illusion (illustrated in **Figure 1A**) during divided attention. If grouping by connectivity is similarly strong it would suggest pre-attentive processing mechanisms responsible for perceptual grouping. On the other hand, if the illusion is reduced with divided attention, it would strongly implicate attention in implicit grouping processes.

# MATERIALS AND METHODS

## Participants

Eighteen neurotypical young psychology students from the University of Florence participated in the experiment [11 females, age: 27.7 ± 2.7 (mean ± SD)]. All were naïve to the goals of the experiment but were experienced psychophysical observers who had all participated in previous psychophysical research. All had normal or corrected-to-normal visual acuity without major visual impairment. This sample size was deemed to be appropriate to obtain a moderate effect size with $\alpha = 0.05$ and power of 0.8. All participants gave written informed consent, and experimental procedures were approved by the local ethics committee ("*Commissione per l'Etica della Ricerca*," University of Florence, July 7, 2020, n. 111), and are in line with the declaration of Helsinki.

## AQ Scores

All participants completed the self-administered Autistic Quotient questionnaire, in either the validated Italian or English versions (Ruta et al., 2012; Ruzich et al., 2015). This contains 50 items, grouped in five subscales: attention switching, attention to detail, imagination, communication, and social skills. For each question, participants read a statement and selected the degree to which the statement best described them: "strongly agree," "slightly agree," "slightly disagree," and "strongly disagree". The standard scoring described in the original article was used: 1 when the participant's response was characteristic of ASD (slightly or strongly), 0 otherwise. Total scores ranged between 0 and 50, with higher scores indicating higher degrees of autistic traits. All except one participant (with AQ 33) scored below 32, the threshold above which a clinical assessment is recommended (Baron-Cohen et al., 2001). The median of the scores was 13.5, with lower and upper quartiles of 8 and 22. Scores were normally distributed, as measured by the Jarque-Bera goodness-of-fit test of composite normality (JB = 1.14, $p = 0.32$).

## Apparatus and Stimuli

The experiment was run in a dimly lit room with stimuli presented on a 13-inch Macintosh monitor with resolution 2,560 × 1,600 pixels; refresh rate 60 Hz. Participants viewed the stimuli binocularly at a distance of 57 cm. The stimuli were generated and presented under Matlab 9.1 using PsychToolbox routines. Dots were small disks of 2.5 mm diameter (subtending 0.25° at 57 cm), half-white, half-black (so that luminance did not vary with a number, providing a potential cue). The stimuli for the numerosity task were two types of random dot-patterns, illustrated in **Figure 1A**: dots were either *isolated* (right image in **Figure 1A**), or with 40% of neighboring dots *connected* to create dumbbell-like shapes (left image). For patches containing isolated dots, dot positions were generated online to respect the sole condition that two items could not be closer than 2.5 mm (0.25°), preventing dot overlap. For the connected patterns, dot position was calculated in two stages: first couples of dots (40% of the total dots of the reference stimulus) were cast and connected *via* a line of the same color, with the constraints that line length was between 10 and 15 mm, with no lines crossing; in the second stage, the remaining 60% of dots were cast with the constraint of not overlapping either the other dots or the connecting lines. The connector line width was 0.5 mm.

Probe stimuli always comprised only *isolated* dots, but the constant-numerosity reference could comprise either isolated (baseline) or *connected* dots. In a particular session, one cloud of dots (the reference, randomly right or left) maintained a particular numerosity across trials, whereas the other (the probe) varied around this numerosity. The number of dots in the probe patch varied according to the QUEST adaptive algorithm (Watson and Pelli, 1983), perturbed with Gaussian noise with a standard deviation 0.15 log units. In separate blocks, three different reference numerosities were tested: 15, 25, 100. The dot stimuli were presented for 500 ms, simultaneously with a visual distractor. The visual distractors (**Figure 1B**) comprised four centrally positioned colored squares (3° × 3°), which could have eight color arrangements. The stimulus was a target if a specific conjunction of color and spatial arrangement was satisfied: two green squares along the right diagonal, or two yellow squares along the left diagonal.

## Procedure

In the single-task condition, participants indicated which of the two peripheral dot clouds contained more dots. In the dual-task condition, participants first responded to the distractor task and then indicated which of the two arrays was more numerous, using the right or left arrow on a keyboard (see **Figures 1C,D**). The order of tasks was pseudorandom across participants. Before starting the experimental condition, all participants were familiarized with the distractor task, in which they were asked to judge whether the central colored square was a target until they attained 75% accuracy; otherwise, the session was repeated. In the main experiment, all trials started with a fixation point presented until the participant pressed a key to start the experiment, and then the primary and secondary stimuli were presented for 500 ms. Participants were tested with three different reference numerosities. For each numerosity, they performed 180 trials, with the order of testing randomized across participants and conditions (connected or isolated), as well as the order of the tasks (single and dual). Participants were asked to maintain fixation on the central stimulus while performing both tasks. To ensure compliance, eye movements were monitored visually by the experimenter during all sessions. We verified that eye movements as small as 2° towards the peripheral stimuli (clouds of dots of 8 degree from central fixation) were readily observable. We reported no cases of observable eye movements under any condition, as may be expected for trained psychophysical observers.

## Data Analyses

For each participant, the proportion of trials in which the probe appeared more numerous than the reference was plotted against the number of reference dots and fit with a cumulative Gaussian error function. The median (the numerosity corresponding to 50% response) gave the point of subjective equality (PSE), and the difference in numerosity required to pass from 50% to 75% defined the JND, a measure of precision. The JND divided by the perceived numerosity yields the Weber Fraction (WF), a dimensionless index of precision that allows comparison of performance across numerosities. Our main analyses compared data across conditions (connected or isolated), tasks (single or dual), and groups of participants: ANOVAs and correlation analyses were complemented with Bayes factors estimation. Bayesian analyses were performed with the software JASP (entered with the per-condition, per-task, and per-subject averages computed in Matlab). Bayes Factors (Rouder et al., 2009) quantify the evidence for or against the null hypothesis as the ratio of the likelihoods for the experimental and the null hypothesis. We express it as the base 10 logarithm of the ratio (LogBF), where negative numbers indicate that the null hypothesis is likely to be true, positive that it is more likely false. By convention, $|\text{LogBF}| > 0.5$ is considered substantial evidence for either the alternate or null hypothesis, >1 strong evidence, and >2 decisive (van den Bergh et al., 2019).

# RESULTS

We tested the effect of attentional load on perceptual grouping over a range of three different numerosities. **Figures 2A,B** show psychometric functions of an example participant for one numerosity (N15), for isolated and connected dots, in the two different attention conditions. For the single-task attentional condition (**Figure 2A**), there was a clear effect of connecting dots: when 40% of the reference dots were connected, the probe PSE was around 11 instead of 15, 27% less than the physical numerosity, agreeing with previous literature (Franconeri et al., 2009; He et al., 2009, 2015; Anobile et al., 2017). However, for the dual-task (**Figure 2B**), the shift was much reduced, only about two elements, or 13%. The point of subject equality (PSE) for the isolated dots in both single and dual was very near the physical numerosity of the reference (N15) in both cases, as to be expected.

Psychometric functions like those of **Figures 2A,B** were fit for each participant, condition and task, from which we extracted estimates of PSE for the various conditions. **Figure 2C** reports the average biases (expressed as percentage change) separately for the single and the dual (color-coded as in **Figures 2A,B**). For both tasks, the baseline biases (gray filled circles) were statistically indistinguishable from zero ($p > 0.5$), as to be expected. The bias of the connected stimuli for the single task was strong, and decreased with numerosity (mean ± SD: N15 = −19.81 ± 8.2; N25 = −18.87 ± 9.1, N100 = −7.91 ± 9.75) confirming our previous results (Anobile et al., 2017; Pomè et al., 2021a). However, the magnitude of the illusion was much less for the dual task condition at the lower numerosities (mean ± SD: N15 = −6.56 ± 8.8; N25 = −6.98 ± 4.2; N100 = −7.99 ± 7.43). This difference is revealed by the statistically significant main effects and interactions between numerosity and task for the connected condition. Two-way repeated measures ANOVA (two attentional conditions, three numerosities: Task $F_{(1,11)} = 26.24$, $p < 0.001$, logBF = 3.40, $\eta_p^2 = 0.7$;

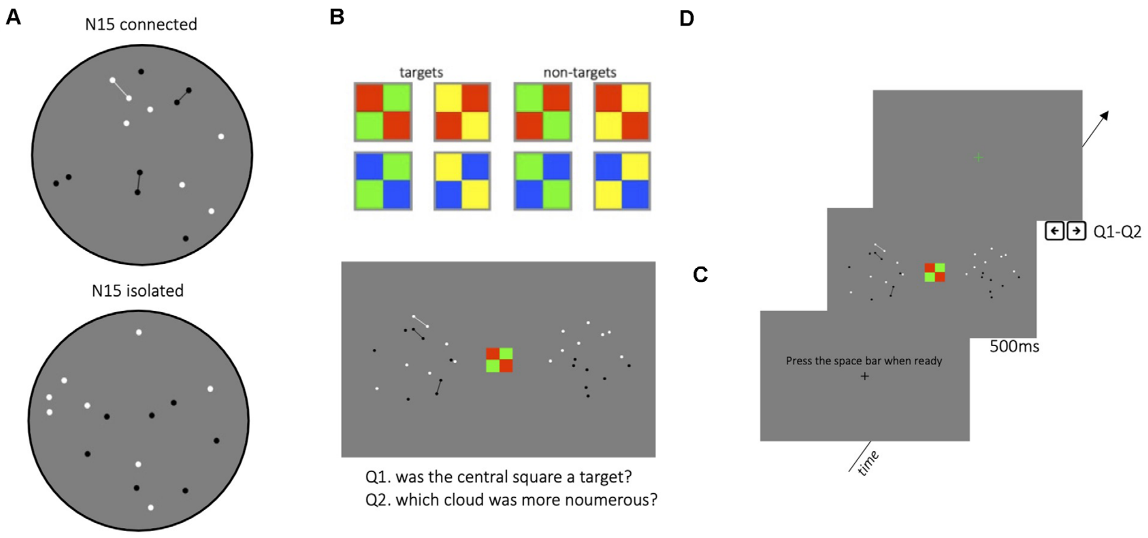


**FIGURE 1 |** Illustration of the stimuli used in the experiment. **(A)** Example stimuli showing the connected condition (upper image), with 40% dots connected by thin lines, and the isolated-dot condition (lower image), with the lines, removed. Although both images have the same numerosity (N15), connecting dots with lines clearly causes the patch to appear less numerous. **(B)** Conjunction stimuli displayed in the center of the screen for the visual-spatial dual-task. The stimulus was a target if it satisfied a specific conjunction of colors and orientations (see the "Apparatus and Stimuli" section for details). **(C,D)** Example of the procedure with a timeline of a trial. Each trial started with a fixation point, followed by two- dot clouds presented together with the distractor, both for 500 ms. In the dual-task condition, participants responded first to the distractor task and then indicated which of the two clouds of dots seemed more numerous. In the single task, they performed only the numerosity task.

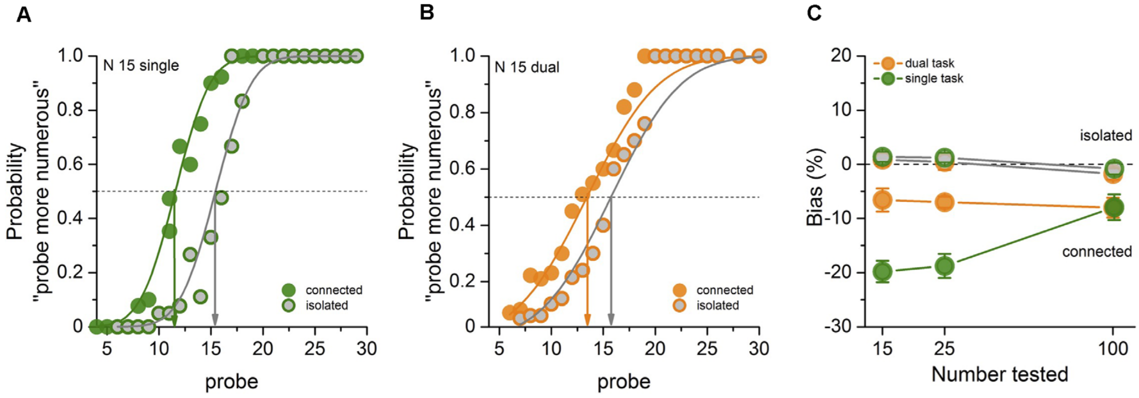


**FIGURE 2 |** Psychometric functions for a typical participant for the single- **(A)** and dual-task **(B)** attentional condition at N15. The graphs plot the proportion of trials where the probe appeared more numerous than the reference (15 dots), as a function of probe numerosity (shown on the abscissa). The vertical arrows show the estimates of the point of subjective equality (PSE), given by the median of the fitted cumulative Gaussian functions. Green and orange respectively refer to single- and dual-task conditions; gray fillings refer to baseline (isolated) conditions, otherwise connected conditions. **(C)** Average PSEs as a function of dot number expressed as the percentage difference from the reference number, color-coded as in **(A,B)**, for both connected and isolated conditions in single- and dual-task.

Numerosity $F_{(1,11)} = 2.95$, $p = 0.07$, logBF = 1.03, $\eta_p^2 = 0.21$; Task × Numerosity $F_{(2,22)} = 5.83$; $p = 0.009$; LogBF = 1.56, $\eta_p^2 = 0.34$).

As has been previously reported, when attention is not deprived, connecting dots of the low-density patterns reduces apparent numerosity considerably, while at higher densities the effect is less obvious. We, therefore, separated the data into low (N15–N25) and high (N100) numerosities to examine in more detail the relationship between biases and numerosities. As we were also interested in the effects of autistic personality traits especially on the condition of divided attention, we divided participants into low AQ (displayed as dark cyan) and high AQ (magenta), based on a median split of their AQ scores (above or below 13.5). **Figure 3A** plots the mean bias for the connected patch at low numerosities against AQ scores for both single and dual task. Results show a good correlation for the single task condition ($r = 0.59$; $p < 0.05$; logBF = 0.66): underestimation of the connected patches decreased with AQ scores. However, the

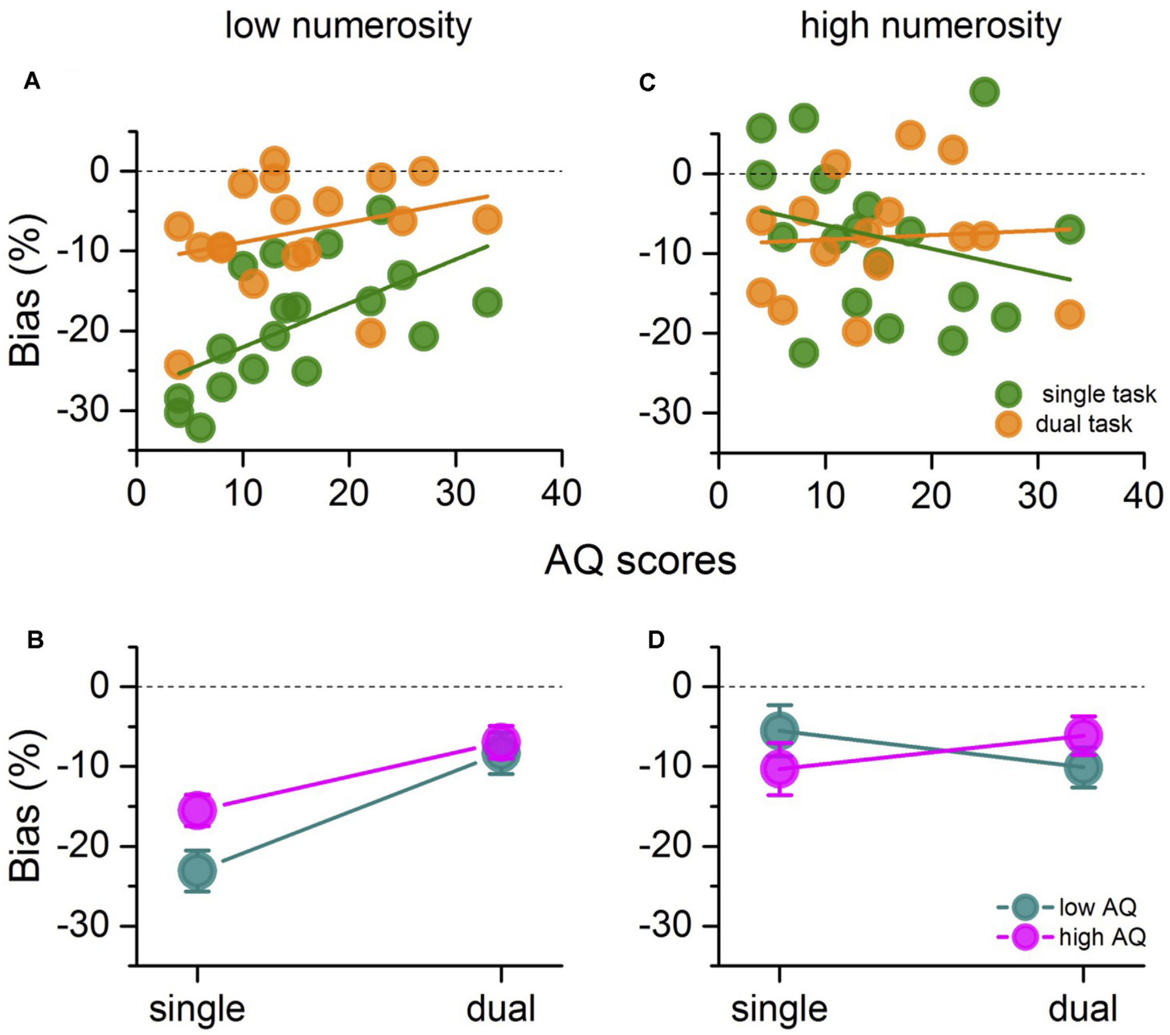


**FIGURE 3 | (A)** Mean bias at low numerosities (N15-N25) plotted against AQ for all participants, for single- (green) and dual-task (orange). Thick green and orange lines show the linear fit of the data. **(B)** Mean underestimation bias for low (cyan) and high AQ (magenta), with error bars = ± 1 SEM, plotted as a function of the two tasks. **(C,D)** Same as in **(A)** and **(B)** but for high numerosity (N100).

dependency on AQ diminished and became insignificant for the dual task condition ($r = 0.30$; $p = 0.2$; logBF $= -0.43$). The Bayes factor is not strong ($|\text{LogBF}| < 0.5$), so it is not clear if the lack of significance results from there being no dependence, or lack of statistical power with the diminished effect. These differences are also evident in the mean underestimation effect for the low- and high-AQ groups shown in **Figure 3B** (mean ± SD: Single Task Low AQ $= -23.08 \pm 7.27$; Single Task High AQ $= -15.51 \pm 5.9$; Dual Task Low AQ $= -8.32 \pm 7.8$; Dual Task high AQ $= -6.95 \pm 6.1$). A two-way repeated measures ANOVA revealed main effect of task ($F_{(1,8)} = 163.6$, $p < 0.001$, logBF $= 0.23$, $\eta_p^2 = 0.95$), but no interaction between AQ and task ($F_{(1,8)} = 2.5$, $p = 0.1$, logBF $= -0.39$, $\eta_p^2 = 0.24$), and no main effect of AQ ($F_{(1,8)} = 5.0$, $p = 0.057$, logBF $= -0.48$, $\eta_p^2 = 0.38$) (although it is approaching significance, mainly driven by the differences in underestimation biases in the single task condition). **Figure 3C** shows the bias at high densities (N100). Here there is no correlation with AQ for either the single or the dual task ($r = -0.25$; $p = 0.3$; logBF $= -0.53$ for the single task and $r = 0.06$; $p = 0.8$; logBF $= -0.70$ for the dual task), and no significant difference between the average bias of the two groups in the two tasks (mean ± SD: Single Task Low AQ $= -5.51 \pm 9.6$; Single Task High AQ $= -10.31 \pm 9.7$; Dual Task Low AQ $= -10.11 \pm 7.5$; Dual Task high AQ $= -6.13 \pm 7.31$), as shown in **Figure 3D** (two way repeated measures ANOVA: main effect of task $F_{(1,6)} = 0.032$, $p = 0.86$, LogBF $= -0.58$, $\eta_p^2 = 0.005$); main effect of AQ $F_{(1,6)} = 0.0006$, $p = 0.98$, logBF $= -0.56$, $\eta_p^2 = 0.0001$; interaction $F_{(1,6)} = 0.33$, $p = 0.58$, logBF $= -0.82$, $\eta_p^2 = 0.053$). For both measures Bayes factors show substantial evidence for no effect.

**Figure 4A** reports the Weber Fraction (WF) of the participants (averaged over all numerosities), given by the SD of the best-fitting Gaussians to the psychometric functions, normalized by the average perceived numerosity. Depriving

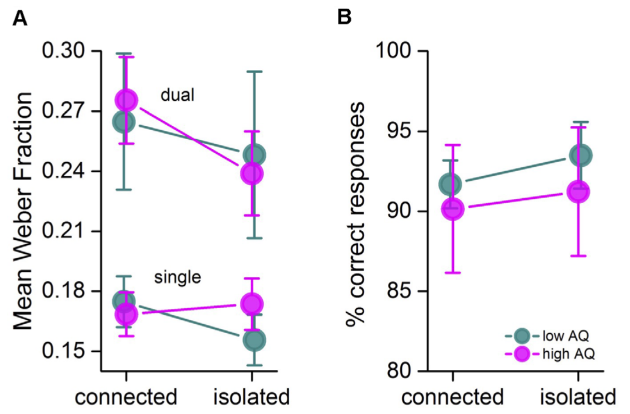


**FIGURE 4 | (A)** Mean Weber fraction for discriminating numerosity in the isolated and connected conditions for the two groups and the two tasks. **(B)** Mean correct responses to the central visual distractor for the connected and isolated condition in the two groups. Color-coded as in **Figure 3**. Error bars = ± 1 SEM.

attention clearly increased thresholds, by about 50%, in line with previous studies with peripheral stimuli (Pomè et al., 2019). However, the costs were similar for the isolated and connected conditions and for the high and low AQ groups (Condition $F_{(1,5)} = 0.93$, $p = 0.38$; logBF $= -0.7$, $\eta_p^2 = 0.15$; Task $F_{(1,5)} = 10.3$, $p = 0.02$ ; logBF $= -1.55$, $\eta_p^2 = 0.67$; AQ $F_{(1,5)} = 0.07$, $p = 0.8$; logBF $= -0.77$, $\eta_p^2 = 0.01$; Condition × task $F_{(1,5)} = 0.29$, $p = 0.62$; logBF $= -0.62$, $\eta_p^2 = 0.05$; Condition × AQ $F_{(1,5)} = 0.28$, $p = 0.62$; logBF $= -1.03$, $\eta_p^2 = 0.05$; Task × AQ $F_{(1,5)} = 0.01$, $p = 0.9$; logBF $= -0.68$, $\eta_p^2 = 0.003$; Condition × Task × AQ $F_{(1,5)} = 20.5$, $p = 0.006$; logBF $= -1.39$, $\eta_p^2 = 0.80$). And similar results were observed even when considering only the low numerosities, which could have affected the results also in terms of precision (Condition $F_{(1,4)} = 0.11$, $p = 0.75$; logBF $= -0.74$, $\eta_p^2 = 0.02$; Task $F_{(1,4)} = 7.12$, $p = 0.056$ ; logBF $= 0.69$, $\eta_p^2 = 0.64$; AQ $F_{(1,4)} = 0.01$, $p = 0.9$; logBF $= -0.72$, $\eta_p^2 = 0.004$; condition × task $F_{(1,4)} = 0.27$, $p = 0.63$; logBF $= -0.62$, $\eta_p^2 = 0.06$; condition × AQ $F_{(1,4)} = 0.36$, $p = 0.57$; logBF $= -1.04$, $\eta_p^2 = 0.08$; Task × AQ $F_{(1,4)} = 0.59$, $p = 0.48$; logBF $= -0.56$, $\eta_p^2 = 0.013$; Condition × Task × AQ $F_{(1,4)} = 5.78$, $p = 0.07$; logBF $= -1.22$, $\eta_p^2 = 0.6$).

We also calculated the percentage of correct responses to the central visual distractor (**Figure 4B**). Performance was very similar in both groups and conditions (excluding the possibility that the results were driven by deteriorated performance on the central distractor). Two-way repeated ANOVA confirmed that none of these effects were statistically significant (all $p > 0.05$, all $\eta_p^2 < 0.2$).

## DISCUSSION

In this study, we examined the importance of visual attention for perceptual grouping in numerosity judgments. Participants judged which of two peripherally presented clouds of dots appeared more numerous, while simultaneously performing a difficult conjunction task. Apparent numerosity was manipulated by connecting dots with thin lines, known to cause underestimation of perceived numerosity, probably by grouping dot-pairs into objects. We replicated previous results (Franconeri et al., 2009; He et al., 2009; Anobile et al., 2017; Pomè et al., 2021a) when participants were free to attend to the dot-stimuli, with participants underestimating numerosity of the connected patch by around 20% for low-moderate numbers; however, the effect almost disappeared when attention was diverted to a concomitant attention-grabbing task, reduced to only 7%.

Together with principles first emphasized by Gestalt psychologists (Wertheimer, 1923) such as proximity, similarity, or common fate, uniform *connectedness* has been suggested as a grouping principle (Palmer and Rock, 1994): connecting a region of uniform visual properties causes it to be organized into a single perceptual unit. Several studies show that connecting dots with lines, as in this study, is a strong grouping facilitator, which dominates other factors, such as proximity and similarity. Connectedness can facilitate visual working memory, by organizing items into pre-packaged "chunks," facilitating encoding of grouped items (Peterson and Berryhill, 2013). Connecting object parts has also been shown to influence the shift of visual attention (Watson and Kramer, 1999), multiple-object tracking (Scholl, 2001), and the performance of patients with visual neglect (Tipper and Behrmann, 1996). Studies have suggested that pairwise connecting of multiple targets significantly alters the spatial distribution of the attentional priority map, increasing the tendency of participants to jointly report or jointly miss elements that belonged to the same object (Gilchrist et al., 1997; Dent et al., 2011).

The present study provides further support that attention modulates the perceptual grouping of elements by connectedness. The results suggest that perceptual grouping affects numerosity estimation only after an attention-dependent grouping mechanism has generated a representation of a perceptual object. This implies that object completion requires sufficient attentional resources deployed to those parts of the visual field that could give rise to the perception of an integrated object; when the allocation of attention is prevented, such as by a concomitant attention-consuming visual task, this cannot proceed. From this perspective, attention may act like a "glue" to bind parts into wholes (Conci et al., 2018), contrary to the view that perceptual grouping may be considered "pre-attentive". Future studies could examine the effects of enhancing attention instead of depriving it of numerosity grouping induced biases.

We have previously shown that the connectedness illusion is strongest for low to medium numerosity densities (Anobile et al., 2017), presumably because the items are less crowded and hence more segregable (Anobile et al., 2015, 2016; Burr et al., 2018). We replicate this effect here. Indeed, the effect of attention on the illusion at high densities was negligible, presumably because it was in any event much reduced.

One possible artifact is eye movements: in the single-task condition, participants could in theory have moved their eyes to foveate the targets, which may have made the thin grouping-lines more salient, whereas this would have been more difficult during the double-task condition. However, we believe this is most unlikely for several reasons. He et al.

(2009) measured connected-induced grouping effects at various exposure durations, and observed, under similar eccentric conditions to ours, that the effect was strong (possibly stronger) at brief, 50-ms durations, too brief for foveation eye-movements to have occurred. Eye-movements are certainly not essential for the effect. In our experiment, while our participants were naïve to the goals of the experiment, they were trained psychophysical observers, well accustomed to maintaining fixation on instruction. To ensure compliance, the experimenter monitored eye-movements visually, after ascertaining that she could detect with perfect accuracy 2° deviations. As the targets were 8° eccentric, it would have been impossible for such a large movement to go unnoticed. We can therefore safely exclude this possibility.

We recently showed that perceiving the numerosity illusion is correlated with perceptual style: participants with higher self-reported autistic traits (AQ) are less susceptible to the connected numerosity illusion, suggesting that they are less susceptible to grouping effects. This is consistent with their having a more detail-oriented perceptual style, which has been linked with autism (Happe and Frith, 2006). In this study, the grouping advantage for low AQ participants disappeared under dual-task conditions. It is difficult to be certain whether the small remaining effect under deprived attention does not depend on AQ, or that the effect has simply become too small to reveal any dependence. The log10 Bayes factor for the correlation was $-0.43$, approaching $-0.5$ (the commonly accepted threshold for demonstrating a null effect), but we remain cautious in interpreting the results. However, we tentatively suggest that the lack of dependence of grouping on AQ when attention is diverted elsewhere is consistent with the same pattern of results for all participants, regardless of the perceptual style.

To conclude, the present study revealed that attention alters the perceived organization of multiple visual elements, furthering our understanding of the way attention modulates grouping by connectedness and impacts visual appearance. Overall, these findings advance our knowledge of the relationship between attention and perceptual organization, two prioritizing mechanisms that help to shape the way we experience our visual world.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, under reasonable request.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Commissione per l'Etica della Ricerca, università di Firenze. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

All authors contributed to the study concept and to the design. Stimuli were designed by AP and CC. Testing and data collection and data analysis were performed by AP and CC. All authors contributed to the interpretation of the results. AP drafted the manuscript and DB provided critical revisions. All authors contributed to the article and approved the submitted version.

## FUNDING

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program, Grant No. 832813 (GenPercept).

## REFERENCES

Adriano, A., Rinaldi, L., and Girelli, L. (2021). Visual illusions as a tool to hijack numerical perception: disentangling nonsymbolic number from its continuous visual properties. *J. Exp. Psychol. Hum. Percept. Perform.* 47, 423–441. doi: 10.1037/xhp0000844

Anobile, G., Cicchini, G. M., and Burr, D. C. (2016). Number as a primary perceptual attribute: a review. *Perception* 45, 5–31. doi: 10.1177/0301006615602599

Anobile, G., Cicchini, G. M., Pomè, A., and Burr, D. C. (2017). Connecting visual objects reduces perceived numerosity and density for sparse but not dense patterns. *J. Num. Cogn.* 3, 133–146. doi: 10.5964/jnc.v3i2.38

Anobile, G., Turi, M., Cicchini, G. M., and Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *J. Vis.* 15:4. doi: 10.1167/15.5.4

Barbot, A., Liu, S., Kimchi, R., and Carrasco, M. (2018). Attention enhances apparent perceptual organization. *Psychon. Bull. Rev.* 25, 1824–1832. doi: 10.3758/s13423-017-1365-x

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., and Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J. Autism Dev. Disord.* 31, 5–17. doi: 10.1023/a:1005653411471

Ben-Av, M. B., Sagi, D., and Braun, J. (1992). Visual attention and perceptual grouping. *Percept. Psychophys.* 52, 277–294. doi: 10.3758/bf03209145

Braun, J., and Sagi, D. (1991). Texture-based tasks are little affected by second tasks requiring peripheral or central attentive fixation. *Perception* 20, 483–500. doi: 10.1068/p200483

Burr, D. C., Anobile, G., and Arrighi, R. (2018). Psychophysical evidence for the number sense. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170045. doi: 10.1098/rstb.2017.0045

Burr, D. C., Turi, M., and Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *J. Vis.* 10:20. doi: 10.1167/10.6.20

Carther-Krone, T. A., Shomstein, S., and Marotta, J. J. (2016). Looking without perceiving: impaired preattentive perceptual grouping in autism spectrum disorder. *PLos One* 11:e0158566. doi: 10.1371/journal.pone.0158566

Castaldi, E., Pome, A., Cicchini, G. M., Burr, D. C., and Binda, P. (2021). Pupil size automatically encodes numerosity. *J. Vis.* 21:2302. doi: 10.1167/jov.21.9.2302

Conci, M., Groß, J., Keller, I., Müller, H. J., and Finke, K. (2018). Attention as the 'glue' for object integration in parietal extinction. *Cortex* 101, 60–72. doi: 10.1016/j.cortex.2017.12.024

Dent, K., Humphreys, G. W., and Braithwaite, J. J. (2011). Spreading suppression and the guidance of search by movement: evidence from negative color carry-over effects. *Psychon. Bull. Rev.* 18, 690–696. doi: 10.3758/s13423-011-0091-z

Driver, J., Davis, G., Russell, C., Turatto, M., and Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition* 80, 61–95. doi: 10.1016/s0010-0277(00)00151-7

Fornaciai, M., Cicchini, G. M., and Burr, D. C. (2016). Adaptation to number operates on perceived rather than physical numerosity. *Cognition* 151, 63–67. doi: 10.1016/j.cognition.2016.03.006

Franconeri, S. L., Bemis, D. K., and Alvarez, G. A. (2009). Number estimation relies on a set of segmented objects. *Cognition* 113, 1–13. doi: 10.1016/j.cognition.2009.07.002

Gilchrist, I. D., Humphreys, G. W., Riddoch, M. J., and Neumann, H. (1997). Luminance and edge information in grouping: a study using visual search. *J. Exp. Psychol. Hum. Percept. Perform* 23, 464–480. doi: 10.1037//0096-1523.23.2.464

Happe, F., and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *J. Autism Dev. Disord.* 36, 5–25. doi: 10.1007/s10803-005-0039-0

He, L., Zhang, J., Zhou, T., and Chen, L. (2009). Connectedness affects dot numerosity judgment: implications for configural processing. *Psychon. Bull Rev.* 16, 509–517. doi: 10.3758/PBR.16.3.509

He, L., Zhou, K., Zhou, T., He, S., and Chen, L. (2015). Topology-defined units in numerosity perception. *Proc. Natl. Acad. Sci. U S A* 112, E5647–E5655. doi: 10.1073/pnas.1512408112

Huberle, E., and Karnath, H.-O. (2006). Global shape recognition is modulated by the spatial distance of local elements—Evidence from simultanagnosia. *Neuropsychologia* 44, 905–911. doi: 10.1016/j.neuropsychologia.2005.08.013

Karnath, H.-O., Ferber, S., Rorden, C., and Driver, J. (2000). The fate of global information in dorsal simultanagnosia. *Neurocase* 6, 295–306. doi: 10.1080/13554790008402778

Kimchi, R. (2009). Perceptual organization and visual attention. *Prog. Brain Res.* 176, 15–33. doi: 10.1016/S0079-6123(09)17602-1

Kimchi, R., and Peterson, M. A. (2008). Figure-ground segmentation can occur without attention. *Psychol. Sci.* 19, 660–668. doi: 10.1111/j.1467-9280.2008.02140.x

Kimchi, R., and Razpurker-Apfeld, I. (2004). Perceptual grouping and attention: not all groupings are equal. *Psychon. Bull. Rev.* 11, 687–696. doi: 10.3758/bf03196621

Kirjakovski, A., and Matsumoto, E. (2016). Numerosity underestimation in sets with illusory contours. *Vis. Res.* 122, 34–42. doi: 10.1016/j.visres.2016.03.005

Lamy, D., Segal, H., and Ruderman, L. (2006). Grouping does not require attention. *Percept. Psychophys.* 68, 17–31. doi: 10.3758/BF03193652

Mack, A., Tang, B., Tuma, R., Kahn, S., and Rock, I. (1992). Perceptual organization and attention. *Cogn. Psychol.* 24, 475–501. doi: 10.1016/0010-0285(92)90016-u

Moore, C. M., and Egeth, H. (1997). Perception without attention: evidence of grouping under conditions of inattention. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 339–352. doi: 10.1037//0096-1523.23.2.339

Palmer, S., and Rock, I. (1994). Rethinking perceptual organization: the role of uniform connectedness. *Psychon. Bull. Rev.* 1, 29–55. doi: 10.3758/BF03200760

Peterson, D. J., and Berryhill, M. E. (2013). The gestalt principle of similarity benefits visual working memory. *Psychon. Bull. Rev.* 20, 1282–1289. doi: 10.3758/s13423-013-0460-x

Pomè, A., Anobile, G., Cicchini, G. M., Scabia, A., and Burr, D. C. (2019). Higher attentional costs for numerosity estimation at high densities. *Atten. Percept. Psychophys.* 81, 2604–2611. doi: 10.3758/s13414-019-01831-3

Pomè, A., Caponi, C., and Burr, D. C. (2021a). Grouping-induced numerosity biases vary with autistic-like personality traits. *J. Autism Dev. Disord.* doi: 10.1007/s10803-021-05029-1. [Online ahead of print].

Pomè, A., Thompson, D., Burr, D. C., and Halberda, J. (2021b). Location- and object-based attention enhance number estimation. *Atten. Percept. Psychophys.* 83, 7–17. doi: 10.3758/s13414-020-02178-w

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., and Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon. Bull. Rev.* 16, 225–237. doi: 10.3758/PBR.16.2.225

Russell, C., and Driver, J. (2005). New indirect measures of "inattentive" visual grouping in a change-detection task. *Percept. Psychophys.* 67, 606–623. doi: 10.3758/bf03193518

Ruta, L., Mazzone, D., Mazzone, L., Wheelwright, S., and Baron-Cohen, S. (2012). The autism-spectrum quotient--italian version: a cross-cultural confirmation of the broader autism phenotype. *J. Autism Dev. Disord.* 42, 625–633. doi: 10.1007/s10803-011-1290-1

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., et al. (2015). Measuring autistic traits in the general population: a systematic review of the Autism-spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Mol. Autism* 6:2. doi: 10.1186/2040-2392-6-2

Scholl, B. J. (2001). Objects and attention: the state of the art. *Cognition* 80, 1–46. doi: 10.1016/s0010-0277(00)00152-9

Shomstein, S., Kimchi, R., Hammer, M., and Behrmann, M. (2010). Perceptual grouping operates independently of attentional selection: evidence from hemispatial neglect. *Atten. Percept. Psychophys.* 72, 607–618. doi: 10.3758/APP.72.3.607

Tipper, S. P., and Behrmann, M. (1996). Object-centered not scene-based visual neglect. *J. Exp. Psycho. Hum. Percept. Perform.* 22, 1261–1278. doi: 10.1037//0096-1523.22.5.1261

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K., et al. (2019). A tutorial on conducting and interpreting a bayesian ANOVA in JASP. *L'Année Psychologique* 120, 73–96. doi: 10.3917/anpsy1.201.0073

Watson, S. E., and Kramer, A. F. (1999). Object-based visual selective attention and perceptual organization. *Percept. Psychophys.* 61, 31–49. doi: 10.3758/bf03211947

Watson, A. B., and Pelli, D. G. (1983). QUEST: a bayesian adaptive psychometric method. *Percept. Psychophys.* 33, 113–120. doi: 10.3758/bf03202828

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt. II. *Psychol. Forsch.* 4, 301–350. doi: 10.1007/BF00410640

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 03 November 2021
doi: 10.3389/fnhum.2021.750417

# Numerosity Perception in Peripheral Vision

**Min Susan Li*†, Clement Abbatecola†, Lucy S. Petro and Lars Muckli**

*Centre for Cognitive Neuroimaging, School of Psychology and Neuroscience, University of Glasgow, Glasgow, United Kingdom*

Peripheral vision has different functional priorities for mammals than foveal vision. One of its roles is to monitor the environment while central vision is focused on the current task. Becoming distracted too easily would be counterproductive in this perspective, so the brain should react to behaviourally relevant changes. Gist processing is good for this purpose, and it is therefore not surprising that evidence from both functional brain imaging and behavioural research suggests a tendency to generalize and blend information in the periphery. This may be caused by the balance of perceptual influence in the periphery between bottom-up (i.e., sensory information) and top-down (i.e., prior or contextual information) processing channels. Here, we investigated this interaction behaviourally using a peripheral numerosity discrimination task with top-down and bottom-up manipulations. Participants compared numerosity between the left and right peripheries of a screen. Each periphery was divided into a centre and a surrounding area, only one of which was a task relevant target region. Our top-down task modulation was the instruction which area to attend – centre or surround. We varied the signal strength by altering the stimuli durations i.e., the amount of information presented/processed (as a combined bottom-up and recurrent top-down feedback factor). We found that numerosity perceived in target regions was affected by contextual information in neighbouring (but irrelevant) areas. This effect appeared as soon as stimulus duration allowed the task to be reliably performed and persisted even at the longest duration (1 s). We compared the pattern of results with an ideal-observer model and found a qualitative difference in the way centre and surround areas interacted perceptually in the periphery. When participants reported on the central area, the irrelevant surround would affect the response as a weighted combination – consistent with the idea of a receptive field focused in the target area to which irrelevant surround stimulation leaks in. When participants report on surround, we can best describe the response with a model in which occasionally the attention switches from task relevant surround to task irrelevant centre – consistent with a selection model of two competing streams of information. Overall our results show that the influence of spatial context in the periphery is mandatory but task dependent.

**Keywords: numerosity perception, peripheral vision, spatial integration, computational modelling, psychophysics**

***Edited by:***
*Roberto Arrighi,*
*University of Florence, Italy*

***Reviewed by:***
*Matteo Toscani,*
*University of Giessen, Germany*
*Nuttida Rungratsameetaweemana,*
*Salk Institute for Biological Studies, United States*

***Correspondence:**
*Min Susan Li*
*susanlee029@gmail.com*

*†These authors have contributed equally to this work and share first authorship*

***Specialty section:***
*This article was submitted to Cognitive Neuroscience, a section of the journal Frontiers in Human Neuroscience*

***Received:*** *30 July 2021*
***Accepted:*** *14 October 2021*
***Published:*** *03 November 2021*

***Citation:***
*Li MS, Abbatecola C, Petro LS and Muckli L (2021) Numerosity Perception in Peripheral Vision. Front. Hum. Neurosci. 15:750417. doi: 10.3389/fnhum.2021.750417*

## INTRODUCTION

Visual resolution decreases toward the periphery of the visual field, compared to foveal vision. Accordingly, while functional brain imaging research using a visual occlusion paradigm shows that the content of a visual scene can be decoded from brain activity patterns in a non-stimulated, peripheral part of the retinotopic visual cortex (Smith and Muckli, 2010; Muckli et al., 2015; Revina et al., 2018; Morgan et al., 2019), human peripheral vision tends to generalise scene information, as evidenced by behavioural phenomena such as crowding (e.g., Balas et al., 2009), the uniformity illusion (Otten et al., 2017), and a higher prominence of gist processing (Larson and Loschky, 2009). This tendency to generalize and blend information is ecologically relevant, if we consider that one of the roles of peripheral vision is to monitor the environment for relevant changes while we use foveal vision to focus on the current task. For example, in rabbits, the upper peripheral visual field is particularly tuned for dark spots on blue skies signifying predator birds (Levick, 1967; Steele-Russell et al., 2012). As humans, we need to cancel out redundant or predictable information in the periphery to save processing power and not be too easily distracted. On the other hand, we need to be made aware of those changes in the environment that are sufficiently salient or unpredictable to be worth further consideration.

Here, we aimed to study behaviourally the specific perceptual processing supporting these features of peripheral vision. In particular, we were interested in whether these phenomena can be explained by a distinct calibration of bottom-up (i.e., sensory information) and top-down (e.g., Stewart et al., 2020) selection of task relevant visual space in the peripheral vision. We selected numerosity as a perceptual property because it is a low-level feature susceptible to gist processing (Park et al., 2016; Fornaciai et al., 2017), and independent to other primary visual properties like objects, colour, shape, or location (Burr and Ross, 2008). It is also easy to manipulate on a numerical continuum for psychophysics purposes (e.g., Valsecchi et al., 2013).

We designed a task using Maximum Likelihood Conjoint Measurement (MLCM; Ho et al., 2008; Knoblauch and Maloney, 2012; Maloney and Knoblauch, 2020), a signal-detection based scaling paradigm which we used to characterises the separate contribution of perceptual attributes to perceived numerosity in the periphery. We presented arrays of dots of varying numbers in the left and right peripheral visual fields and participants had to indicate whether there were more dots on the left or the right side. The peripheral areas on each side were further divided into a centre and a surrounding region, only one of which was the task-relevant target while the other was a task-irrelevant context. We could then quantify how perceived numerosity in the relevant part of the display (bottom-up information) is biased toward the number of dots presented in the irrelevant part (contextual information).

In the case that we found a perceptual bias toward the task-irrelevant signal, a possible account would be based on the imprecision of the top-down connections that span out to a larger region. The feedforward input is then not matched by the correct top-down predictions. In such a leaking model prediction errors around the boundaries could lead to an over or underestimation around the boundaries. This Predictive Coding account would lead to an integrative process in which bottom-up and top-down signals are combined to a distorted perception based on integration of a prediction error around boundaries (e.g., Rao and Ballard, 1999; Friston, 2005). An alternative outcome would be observing a serial process under which only one source of information can be perceived at a time (e.g., limited perceptual capacity; Yiğit-Elliott et al., 2011), and including contextual signals increases the ambiguity of the overall stimuli to the extent that the irrelevant cue is sometimes perceived as the target (e.g., Craig, 1976; Berry and Fristedt, 1985). This model of a serial process could be explained in terms of Predictive Coding the way Jakob Hohwy explains binocular rivalry (Hohwy et al., 2008): the target model explains away the stimuli of the target region, but leaves the stimuli in the irrelevant regions unexplained as a consequence the irrelevant region creates so much prediction error that it sometimes forces the internal model into one that is consistent with the irrelevant information. In order to assess these two accounts, we consider the integration and switching models that make different assumptions about mandatory integration for perceiving the target and contextual stimuli. We aimed to clarify whether the perceptual decisions were made as a weighted average of relevant and irrelevant signals, or were made on the basis of a probability either according to the relevant or irrelevant part of the display, on a trial by trial basis.

Quantifying contextual effects also allowed us to study how the combination of perceptual cues is modulated by both higher-level top-down and bottom-up factors. Firstly, as a bottom-up factor, we varied the duration of presentation intervals to assess how contextual influence is related to the amount of acquired information. Summary representation of visual features in the periphery can be processed within a brief temporal window as short as 50 ms (Chong and Treisman, 2003), and here we aimed to investigate how perceived numbers in the periphery is affected by the strength of bottom-up signals with temporal intervals up to 1 s. Secondly, as a top-down factor, we modulated the area of display participants were attending to. We compared a condition in which the task-relevant targets were the centre areas and the irrelevant context were the surround areas (i.e., contextual filling-in from surround to centre), to a condition in which the task relevant targets were the surround areas and the irrelevant context were the centre areas (i.e., contextual expansion from centre to surround). By varying the spatial arrangement of task-relevant regions, we examine whether the perceptual mechanism of spatial interaction is comparable for filling-in or expanding the context information.

## MATERIALS AND METHODS

### Participants

We recruited 18 volunteers with normal or corrected to normal vision (15 females and 3 males, mean age 25.3 ± 4.22 years) through the Psychology Experiments Participant Pool of the University of Glasgow, and we paid participants 6 pounds per hour. All participants gave informed consent prior to the experiment. We randomly assigned participants to two groups of equal size, each group performing a different task (see Procedure). The study was approved by the ethics committee of the College of Science and Engineering of the University of Glasgow and conducted according to the principles expressed in the Declaration of Helsinki.

## Apparatus and Design

The experiment was controlled by PsychoPy v3.2.3 (Peirce, 2007) on a Windows 10 HP EliteOne 800 All-in-one PC, with a monitor size of 525 × 296 mm, refresh rate of 59 Hz and 1,920 × 1,080 resolution. Participants were placed at a distance of 57.3 cm from the screen with a chin rest so that 1 cm was equivalent to 1 degree of visual angle.

The display had a mid-grey background with the central part of the visual field masked by a 200 × 295 mm black area, in order to restrict stimulations to only the near-peripheral visual field. A red fixation point (5 × 5 mm) was placed in the middle of the central black area. In the left and right peripheral parts (10 degrees each from the fixation point) were two peripheral displays, each divided into two sub-parts, *centre* and *surround*, with a black rectangle shape indicating the border of the centre (see **Figure 1**). The centre and surround areas were always of the same size for each peripheral display. Thus, the four regions of interest in the study were: centre (area inside the rectangle) in the left periphery (CL), centre in the right periphery (CR); surround (area outside of the rectangle) in the left periphery (SL) and surround in the right periphery (SR).

Visual stimuli were circular black and white dots (50 and 50%), 5 mm diameter each applied with a Gaussian blur. The spacing of dots was determined by a uniform random distribution. Note that the dots covered only the peripheral areas of the visual field, and not the central black area (**Figure 1**). In each trial, the number of dots in a given area (CL, CR, SL or SR) could be 30, 60 or 120, in accordance to the Weber's Law (Fechner, 1860) that a multiplicative increase in the physical magnitude of numbers is expected to translate into a linear perceptual increase (Brannon et al., 2001; Ross, 2003; Jordan and Brannon, 2006; Merten and Nieder, 2009; Anobile et al., 2014). This yielded 9 possible combinations of the number of dots for the centre and surround regions on one side of the display (**Table 1**). Hence, for the whole display, there were a total of 36 combinations in which left and right peripheries were non-identical, as non-identical dot numbers on the two sides are required for the psychophysical task (see below). Centre/surround dot combinations were randomly assigned over left and right peripheral displays on each trial. In addition, we varied the duration of visual display by presenting the stimuli for 100, 400, 700 or 1,000 ms. We chose these values to include a range of exposure durations that gradually increase in clarity.

## Procedure

The testing cubicle remained dark throughout the experiments to prevent observers from experiencing changes in luminosity. For each trial, participants were shown the stimuli (for 100, 400, 700 or 1,000 ms) followed by a 200 ms white visual noise (covering only both peripheral areas, excluding rectangles that define the central and surround areas and the mid-screen black area) to control for visual aftereffects. All observers performed a numerosity discrimination task: when the visual

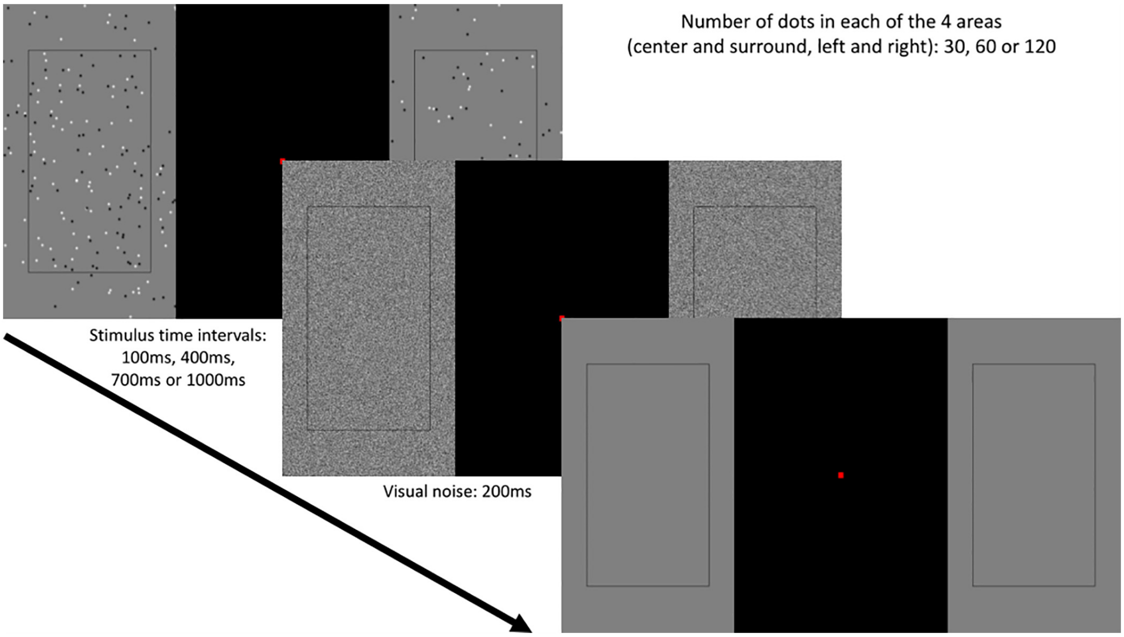


**FIGURE 1 |** Experimental protocol. For each trial, a varying number of dots was presented in left centre, left surround, right centre, right surround, for a varying duration: 100, 400, 700 or 1,000 ms. This was followed by brief visual white noise, after which the observers had to discriminate, for the relevant area, either left or right side presented more dots (Two- Alternative Forced Choice).

**TABLE 1 |** An example of the experimental conditions (number of dots) presented on one side of the display.

| Experimental combinations | Centre | Surround |
|---|---|---|
| | Number of dots | |
| 1 | 120 | 30 |
| 2 | 120 | 60 |
| 3 | 120 | 120 |
| 4 | 60 | 30 |
| 5 | 60 | 60 |
| 6 | 60 | 120 |
| 7 | 30 | 30 |
| 8 | 30 | 60 |
| 9 | 30 | 120 |

noise disappeared, observers had to press one of the two buttons to indicate whether there were more dots on the left or right relevant part of the screen. One of the two groups of participants was instructed that only the centre regions of both sides were relevant (judging CL vs. CR), and the other group was assigned the surround of both sides as relevant regions (judging SL vs. SR). The next trial began after a 300 ms inter-trial interval after each response. The experiment lasted approximately 30 min with a short break halfway through the experiment.

The 36 combinations of the number of dots were presented 8 times at each temporal interval, yielding 36 (combinations) × 4 (time intervals) × 8 (repetitions for each unique trial) = 1,152 trials in total for each participant. The timing factor was a within-subject design, and the task factor was a between-subject design to avoid adaptation effects and confusion between the tasks.

## Analysis

In this section we describe the implementation of Maximum Likelihood Conjoint Measurement, allowing us to use scaling measures to estimate the perceptual bias of judgments and examine the possibility of integrated perceptual information with three decision models. We also simulate the observer's responses with two specific decision rules to determine the mechanism of such perceptual integration.

### Maximum Likelihood Conjoint Measurement

Our protocol and analyses followed the principles of Maximum Likelihood Conjoint Measurement (MLCM; Knoblauch and Maloney, 2012; Maloney and Knoblauch, 2020), a signal-detection based scaling paradigm, under which the contribution of different features to perceptual decisions can be quantified. Although initially designed to study how multiple physical properties (e.g., visual roughness and glossiness) interact in their perception (Ho et al., 2008), MLCM has recently been applied to study how the properties of a background surface affect the perception of a central surface (Hansmann-Roth and Mamassian, 2017; Hansmann-Roth et al., 2018). Here, we manipulate the physical properties of centre and surrounding areas in the periphery, and we examine how irrelevant areas contribute to the perceived numerosity in the relevant area, depending on the task (centre task or surround task).

Assuming we are handling the data of an observer from the centre task group (CL vs. CR), the simplest decision model would be one where the observer compares some internal function of the number of dots in left and right centre areas:

$$\Delta_C = \psi_C(CL) - \psi_C(CR) + \epsilon \quad (1)$$

Where CL and CR are the number of dots in the centre on the left and on the right, respectively. On a given trial, $\psi_C$ is some internal function determining the perceived number of dots in centre on a single side given the actual number, $\epsilon$ is an unbiased and normally distributed decision noise, and $\Delta_C$ is the decision variable whereby the left side (if $\Delta_C > 0$) or the right side (if $\Delta_C < 0$) is chosen by the observer as containing the highest number of dots in the central area. This is called an Independence Model, and such a model assumes that the perceived number of dots in CL is completely independent from SL. However, it is also possible that the number of dots in SL will contribute to the numerical estimate of CL.

The simplest model to take such effects into account is the Additive Model:

$$\Delta_C = [\psi_C(CL) + \psi_S(SL)] - [\psi_C(CR) + \psi_S(SR)] + \epsilon \quad (2)$$

Where SL and SR are the number of dots in surround on the left and on the right, respectively, and $\psi_S$ is some internal function determining the contribution of the number of dots in the surround to the number of dots perceived in the centre on a single side. In the Additive Model, we make the hypothesis that the contribution of CL will not vary when changing the number of dots in SL (and vice versa).

To test this hypothesis, we can introduce interaction effects in the Full Model:

$$\begin{aligned}\Delta_C = {} & [\psi_C(CL) + \psi_S(SL) + \psi_{CS}(CL, SL)] \\ & - [\psi_C(CR) + \psi_S(SR) + \psi_{CS}(CR, SR)] + \epsilon \quad (3)\end{aligned}$$

Where $\psi_{CS}$ is a function determining interaction effects for each combination of the number dots in centre and surround. A possible instance of the full model could be a contrast enhancement model allowing for the centre to appear more numerous in the context of low numerosity and less in case of a surround of higher numerosity.

The three models defined here can be formalized as Generalized Linear Models to estimate $\psi$ functions using maximum likelihood. As the models are nested within each other, the difference of their log-likelihoods is distributed as $\chi 2$ with degree of freedom the difference in the number of parameters (see e.g., Wood, 2015). We can therefore compare them to test our hypotheses using likelihood ratio tests (Maloney and Knoblauch, 2020). Such analysis, applied to different tasks (centre task group or surround task group) and different temporal intervals, will allow us to reconstruct different but comparable contribution values for perceiving CL, CR, SL, and SR.

### Simulated Observers: Integration and Switching

If the independence model is rejected in favour of the additive model in MLCM, one would usually assume that results suggest a form of perceptual integration took place between the cues of interest. In our paradigm, for example, the number of dots perceived on one peripheral display can be interpreted as a weighted sum of the number of dots presented in the centre and surround of that side. However, because observers were instructed to judge only one of the two areas of each side, the perceptual weights of each sub-part contribution will vary depending on the task. For the centre task condition, for instance, more weight should be attributed to the centre and less to surround. This combination rule can be expressed as:

$$\psi_1 = w_C \cdot \psi_C(CL) + (1 - w_C) \cdot \psi_S(SL) \qquad (4)$$

Where $\psi_1$ is the number of dots perceived on a single side of the screen and $w_C \in [0, 1]$ is the weight attributed from the centre. Under this rule and given assumptions from the Signal Detection Theory (Green and Swets, 1966), an experimental combination repeated over many independent trials with constant number of dots in centre and surround should follow a Gaussian distribution (**Figure 2**, top left panel).

As an alternative account, it is also plausible that responses followed a switching pattern: in any given trial, only one source of information is selected with probability $w_C$ of choosing according to centre and $1 - w_C$ of choosing according to surround. This switching rule can be expressed as:

$$\psi_1 = [x < w_C] \cdot \psi_C(CL) + [x > w_C] \cdot \psi_S(SL) \qquad (5)$$

Where $x \in [0, 1]$ is a random uniform variable. With the switching rule, an experimental combination repeated over many trials with constant number of dots in centre and surround should follow a bimodal distribution (**Figure 2**, top right panel).

Note that when weight values are extreme, i.e., $w_C = 0$ and $w_C = 1$, there is no difference between the two rules over many repetitions of the same trial. The average value over many repetitions with a constant $w_C$ is also the same between the rules.

We can then simulate the responses to an MLCM experiment with either integration rule, where the response to each trial is determined by:

$$\Delta_\mathrm{I} = w_\mathrm{C} \,.\, (CL - CR) + (1 - w_\mathrm{C}) .\, (SL - SR) + \epsilon \qquad (6)$$

or switching rule, where the response to each trial is determined by;

$$\Delta_\mathrm{S} = [x < w_\mathrm{C}] .\, (CL - CR) + [x > w_\mathrm{C}] .\, (SL - SR) + \epsilon \qquad (7)$$

Where the $\Delta_I$ and $\Delta_S$ are the decision variables for the additive integration and switching model, respectively, and the notation otherwise follows Eqs. 2, 4, and 5. In particular we define the link functions as $\psi_C(X) = X$ and $\psi_s(X) = X$ for these simulated observers.

This yields contribution scales (**Figure 2**, middle panel) that are similar to typical empirical results when applying MLCM analysis. Most importantly, this method allows us to recover specifically which rule and weight value were implemented by a given simulated observer if we compare, with root-mean squared error, the observer's result with the results of other simulated observers using a representative sample of rules and weight values (**Figure 2**, lower panel, which also includes a random rule under which the observer responds left or right randomly regardless of trial). We will use this method to determine which rule better describes the responses of our human observers under different experimental conditions.

# RESULTS

## Maximum Likelihood Conjoint Measurement

First, we compared MLCM models in terms of complexity. We fitted the independence, additive and full models at each time interval for each participant. This allowed us to compare independence vs. additive and additive vs. full model in each case using likelihood ratio tests. The details of these comparisons are presented in **Supplementary Table 1**. To summarize, while the finding is somewhat noisy at 100 ms time interval due to the difficulty of the task, with longer time intervals we found that the independence model should always be rejected in favour of the additive model (all $p < 0.001$), and that in most cases the additive model should not be rejected in favour of the full model (except one participant from the surround task group at 400 ms: $\chi^2(4) = 14.87, p = 0.005$).

This being the case, we turned our focus to the additive model, of which the average contributions across participants are shown in **Figure 3**. Given the outcome of model comparison, results of the additive model were expected, which showed a qualitative difference between 100 ms and other temporal intervals (i.e., 400, 700 ms, and 1 s). For the 100 ms interval, the contributions of centre and surround were always low across all combinations but not completely flat, which indicated that participants did not respond randomly even with 100 ms. For longer intervals, results showed more contribution of the task-relevant area. Interestingly, there was no reduced contribution of the task-irrelevant area with longer intervals, and these contributions were still significant even at the 1,000 ms interval for both centre and surround tasks as indicated by model comparison.

Moreover, the increase of the perceived number of dots when multiplying the actual number of dots in centre or surround by 2 is, as expected, relatively linear in all cases. The variation between timings and tasks can therefore be interpreted in terms of changes in perceptual weights and/or in perceptual strategy between Integration and Switching. These questions will be addressed in the next section.

In **Supplementary Figure 1**, we propose an alternative representation of the results without relying on MLCM modelling. The proportion of times each combination of centre/surround dot number was chosen for both tasks is represented for each time interval. Grey lines represent "ideal observers" responding either randomly (horizontal line) or always choosing the highest number of dots in the relevant dimension and randomly when the same number is presented on both sides (this occurs in 1/3 of cases). This shows that the number of dots in the relevant area (centre or surround) is the most important factor determining the observers' choices, while the number of dots in the irrelevant area biases this choice. When

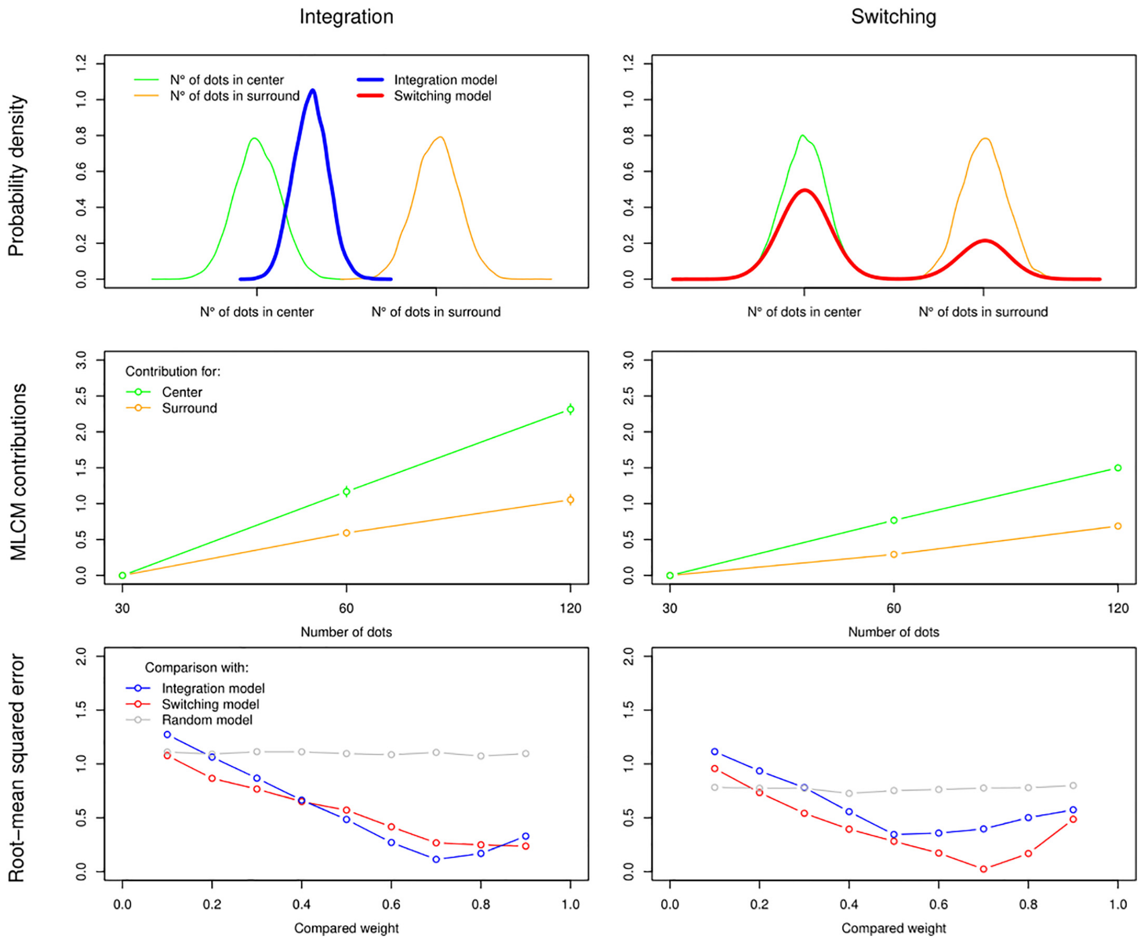


**FIGURE 2 |** Simulating observers using an Integration rule (left column) or a Switching rule (right column) to estimate perceived numerosity in the centre and surround. Top panel: Probability density functions over 1,000 simulations of the same single combination. In this example weights of 0.7 for centre and 0.3 for surround are applied. Density for the perceived number of dots in centre and surround are represented in green and orange, respectively, and the resulting representations under Integration and Switching rules are in blue and red, respectively. Second panel: results of the MLCM analysis (additive model) applied to simulated responses of observers using Integration and Switching rules (in both cases, the simulated trials were the sum of all trials done by 9 of our human participants to make it comparable to empirical results). Bottom panel: comparison of two rules using RMSE between the simulated data (from the mid-panel) and other simulations with varying weight values and combination rules (Integration, Switching, and Random choice for each trial as a control).

the number of dots in the irrelevant area is 30 or 120, the number of dots in the relevant area is underestimated or overestimated, respectively. When the number of dots in the irrelevant area is 60, the estimated number of dots in the relevant area is very close to an ideal observer which responds to the number of dots in the relevant area with maximal accuracy. Compared to this analysis, MLCM allows a straightforward significance test for the effect of the irrelevant area (independent vs. additive model) and for interaction effects (additive vs. full model). It also allows further modelling of the underlying decision processes as proposed in the next section.

## Integration and Switching, Simulated Observers

We compared the empirical data to simulated data to establish the best-fitted weight value and combination rule for each task and at each timing interval. In accordance with our previous results, we found a difference between the 100 ms interval and the other intervals. At 100 ms, neither the Integration nor Switching rule at any weight value performed better than an observer choosing at random to predict the participants' decisions (**Figure 4** top panel: the left tab in both plots). For the remaining longer intervals, we observed a consistent pattern in which the Random Choice Model was the worst-fitted model, the Integration Model was better at predicting the choices of observers judging according to centre, and the Switching Model was better at predicting the choices of observers judging according to surround (**Figure 4** top panel, the right tab in both plots). This difference was significant over 10 independent simulations as assessed by independent $t$-tests: for centre task, Integration ($M = 0.09$, SD = 0.03), Switching ($M = 0.17$, SD = 0.03), $t(18) = -6.85$, $p < 0.001$. For Surround task, Integration ($M = 0.20$, SD = 0.04), Switching ($M = 0.13$, SD = 0.02), $t(18) = 4.87$, $p < 0.001$.

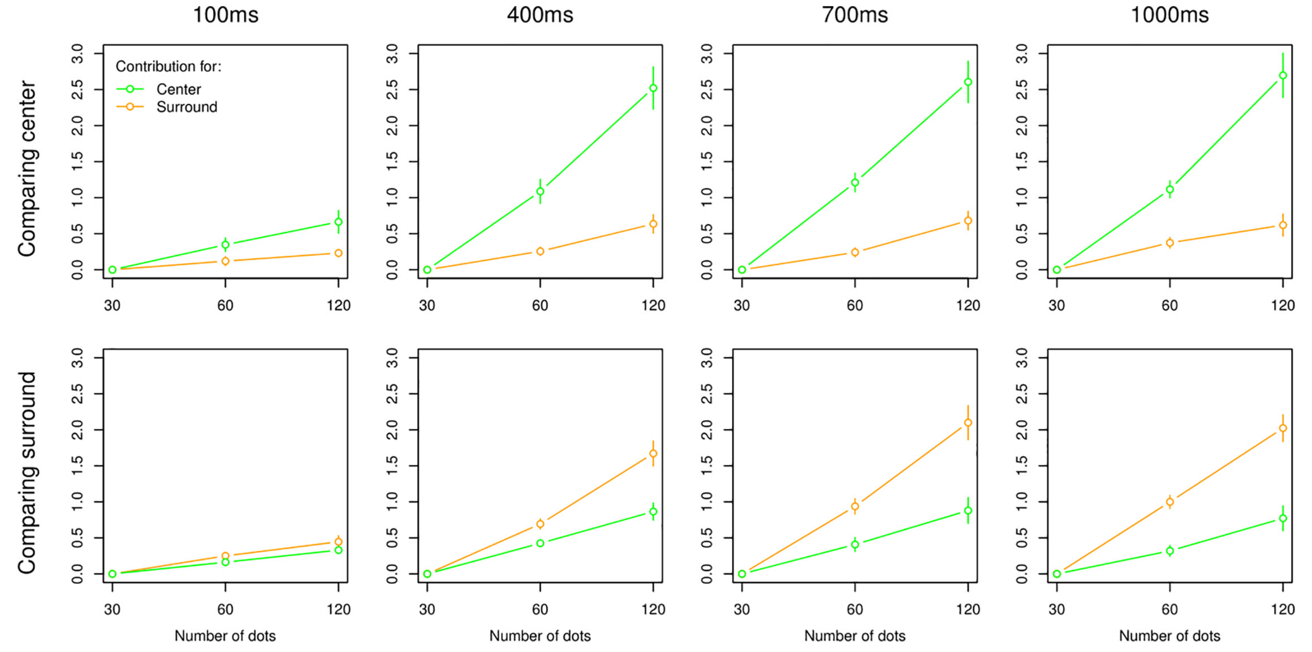


**FIGURE 3 |** Results of MLCM analysis (additive model). Average contributions to the perceived numerosity in centre (green) and surround (orange) as a function of 30, 60, and 120 dots. Plots are organized by task (i.e., top panel: centre task, bottom panel: surround task) and presentation intervals (columns). All error bars are standard error of the mean.

Furthermore, because each rule model had a single free parameter that has been varied to show the range of patterns that can be captured by the model (i.e., **Figure 2**), here with the smallest root-mean square error, we obtained the statistically optimal weight values for each timing interval and for each rule (**Figure 4**, bottom row, see **Supplementary Table 2** for a statistical comparison between RMSE distributions for the best and second-to-best weight values at each time interval, and **Supplementary Figure 2** for an illustration). Our results, again, suggested a consistent and robust behavioural trend that performances of 100 ms intervals were significantly distinctive from other longer intervals, where Integration and Switching Model performed similarly and obtained a weight value around 0.6 for the centre task and 0.4 for the surround task. It is worth noting that intervals above 100 ms held the same weight value for the centre task, but for the surround task there was an increase in optimal weight, specifically in a way that decisions favoured the relevant area with longer intervals until a floor effect.

## DISCUSSION

We investigated the processing of numerical magnitude in peripheral visual displays, in which we found that the perceived numerosity of the target area is biased toward the number of dots presented in irrelevant neighbouring areas. Specifically, our results suggested that numerical magnitudes in the periphery were sampled following either a "weighted integration" or a probability switching' process between the target and irrelevant areas. In other words, contextual cues presented in the surroundings were used for inferences about the numerosity in the centre; whereas contextual cues presented in the centre competed on a trial-by-trial basis to the perceived numerosity in the targeted surround. Thus, we argue that top-down factors, such as directing attention toward different areas in the peripheral visual field, have an impact on how predictions are incorporated into perceptual decisions about numerosity in peripheral vision.

We generally observed that numerosity perception in peripheral displays required a sufficient sampling time. Results from both centre and surround tasks showed that peripheral displays presented for 100 ms seemed highly ambiguous, and consistent with random responses of simulated observers. For presentation intervals longer than 100 ms, our results suggested an involuntary perceptual bias between task relevant and irrelevant parts of a display, supported by the significant advantage of the MLCM additive model (where the model's decisions were based on both relevant and irrelevant areas), compared to the independence model (where the model considered only the relevant area). Interestingly, results showed that the perceptual contribution of target and irrelevant areas remained consistent across all intervals above 100 ms in the centre task, while the contribution of irrelevant central information while judging the surrounding region decreased progressively with longer intervals. By comparing the results of centre and surround task to the simulated observers, we identified an integration process for the centre task and a switching process for the surround task.

Specifically, perceived numerosity of a target region with irrelevant surrounding influences is best described as "contextual

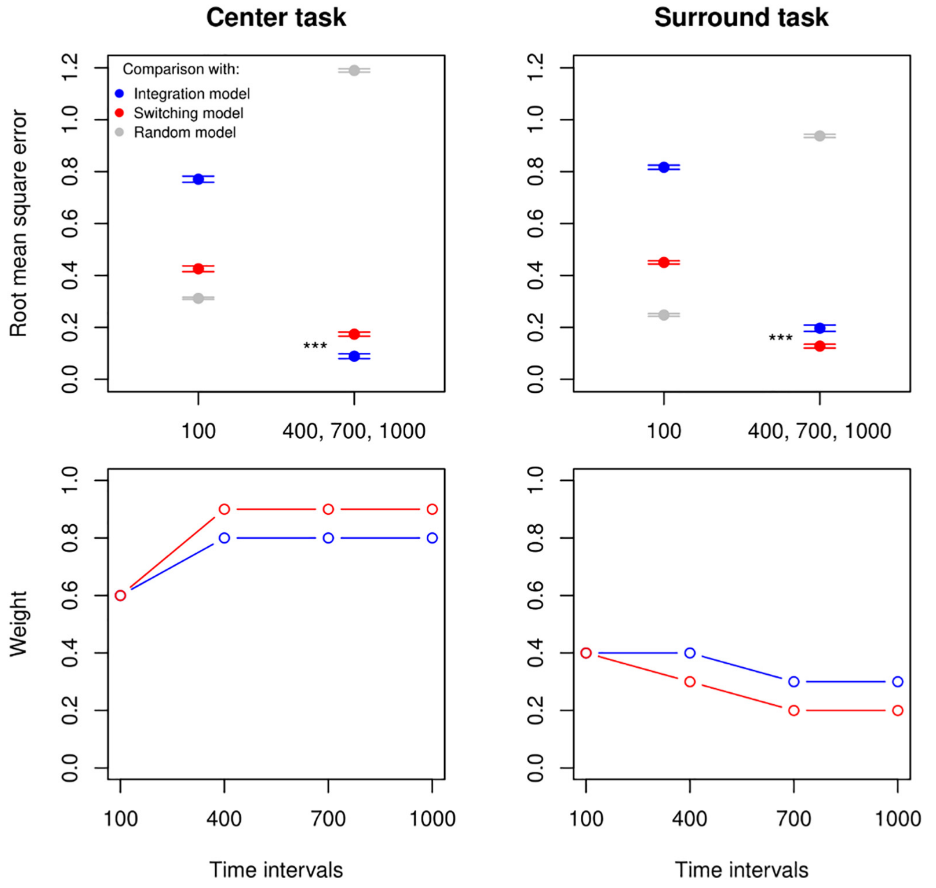


**FIGURE 4 |** Observer simulations and optimal weights. Top panel: a comparison using the root-mean squared error between empirical data from the centre and surround conditions (left and right plot, respectively) and simulated data with varying combination rules (Integration in blue, Switching in red, and Random choice in grey). Ten datasets were simulated for 9 weight values (from 0.1 to 0.9) and for each rule. In both plots the lowest RMSE values across all weights at the 100 ms interval for all combination rules (and their standard errors across the 10 simulations) are shown on the left, while the mean and standard errors of the lowest RMSE values averaged across the remaining intervals are shown on the right. Bottom panel: pattern of weight values yielding the lowest RMSE at each timing and for Integration Model (in blue) and Switching Model (in red) when compared to empirical data (left: centre task, right: surround task). ***$p < 0.001$.

leaking in" effect, i.e., perceived numerosity of the central area is a perceptual combination with a weight of 0.8 attributed to the number of dots physically presented in centre, and 0.2 to the number of dots physically presented in its surround. In other words, it is as if the integration window fits with some insufficient precision on the centre areas, and 20% of the information from the surrounding spills into the decision in a way that the two channels cannot be separated with sufficient spatial precision. These weight values are stable across timing conditions above 100 ms, suggesting a common mechanism of integration for stimuli presented briefly and for longer durations. This low-level spatial integration involves incorporating a small amount of sensory signal from the surround area, and corresponds to an integration process described in the cue combination literature whereby independent noisy sources of information are combined into a weighted average, boosting the precision of perceptual estimates (e.g., Ernst and Banks, 2002; Ernst and Bülthoff, 2004; Hillis et al., 2004; Acerbi et al., 2014). Sensory integration, in fact, has been the only interpretation for additive MLCM models so far.

In contrast, perceptual decisions about surrounding areas with irrelevant central influences are described as an expansion contextual prior. In this case, participants' responses are most consistent with a perceptual switching process, i.e., participants responded inaccurately according to the central region in some instances, while the surround targets are perceived accurately in the remaining majority of trials. Contrary to the weighted combination of perceptual inputs described above, this switching process does not require spatial integration of signals from relevant and irrelevant areas. Our results showed that participants made incorrect switches 30% of the time with 400 ms intervals, but the proportion was reduced to 20% of the trials for 700 and 1,000 ms intervals. To put this more parsimoniously, perhaps two separate streams of information are processed for the surround task, and the perceptual decision is reached by a lateral inhibition process between the two channels. Unlike an integration process, the switching process was therefore sensitive to changes in

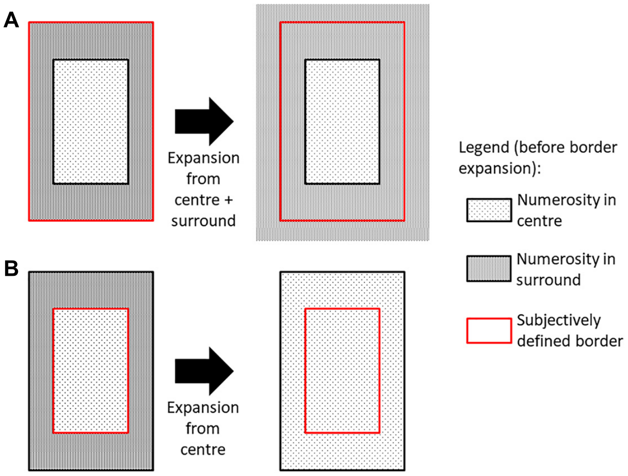


**FIGURE 5 |** One plausible explanation for border expansion responses when participants were asked to judge the number of dots in the surrounding portion of the display. Here bistable perception is assumed to depend on whether the subjectively defined border which is expanded includes both central and surround areas, which doesn't affect numerosity perception in surround **(A)**, or only the central area, which would then be expanded into the surround **(B)**.

interval lengths, although with a floor effect. One may argue that a perceptual switching process is evidence for sequential processing with limited resources (e.g., Landy et al., 2007; Whiteley and Sahani, 2008; Scharff et al., 2011; Yiğit-Elliott et al., 2011). This account implies that participants only had time to process one of the two parts of the display due to capacity limitations (e.g., difficulty of the surround task), where participants mistakenly prioritised assessing the number of dots in centre, instead of surround, for a small proportion of the trials. This account is partially supported by our empirical data where we showed a decreased tendency to make incorrect switches when participants were provided longer temporal intervals with the display. However, the fact that we still observed the switching behaviour at the longest interval makes this explanation less plausible unless we consider the possibility of another, incompressible, source of error such as motor or attentional mistakes.

A potential mechanism that could induce such an attentional error is biased bistable perception of ambiguous stimuli (e.g., Mamassian and Landy, 1998; Meng and Tong, 2004), in which there is a competition for awareness between several mutually exclusive interpretations of the same stimulus. Bistable perception could be influenced by the boundary extension effect, under which our recollection of scenes tends to extend beyond the border of what was actually presented (as reviewed in Hubbard et al., 2010). It shows how overall perception is affected by ambiguity in our displays. While boundary-extension has mostly been studied for natural scenes, line-drawing paradigms also provided evidence for such an effect (Gagnier and Intraub, 2012) and occurs also in the absence of semantic associations (McDunn et al., 2014), which leads to the possibility of similar processing for the current paradigm. If the display boundary is subjectively defined and the definition varies from trial to trial (e.g., condition to saliency of varying density of dots), such ambiguity would affect the perceived numerosity of the display. In particular, if the boundary includes the end of the complete peripheral display (centre *plus* surround, left or right), extending this "overall boundary" only affects the broader surrounding area. As a result, perceived numerosity in the surround would not be affected by this extension (see **Figure 5A**). However, if one considers only the boundary of the centre region, applying boundary extension would have the consequence of generalizing it to the surround area, making the observer perceive the same numerosity in centre and surround (**Figure 5B**), which may be the cause of incorrect switching in our results.

On the whole, we showed that perception of a task-relevant region is biased toward task irrelevant region, in agreement to contextual influences that have previously been found in situations where the resolution of sensory inputs was low and thus signals were combined to strengthen the reliability of perception (e.g., Levi et al., 2002; Kersten and Yuille, 2003), or foveal-to-peripheral extrapolation in which foveal vision is used to estimate the strength of stimulus properties in the periphery where visual resolution is limited (Toscani et al., 2017). However, we established the persistence of contextual influences with numerosity information presented for as long as 1 s, indicating that the accuracy of numerical magnitude judgment is not rectified with more processing time added awareness. We suggest that the influence of spatial context and the perceptual bias of perceived numerosity that we observed depends on a top-down mechanism, in which that signals from the irrelevant areas automatically create a perceptual expectation that participants used to infer about their perceived number of dots, because peripheral vision is more limited in terms of acuity than foveal vision. Furthermore, we found this perceptual bias to be affected by top-down contextual factors: the surround effect while judging a central area is most consistent with spatial integration, while the centre effect when judging a surrounding area is better explained as switching between two information channels which we hypothesise to be linked to ambiguous boundary extension. These findings illustrate the complexity and flexibility of processing in peripheral vision (Stewart et al., 2020), and more broadly, our data are in line with evidence for predictive models of vision where top-down priors are combined with incoming sensory inputs (Rao and Ballard, 1999; Friston, 2002; Edwards et al., 2017; Spratling, 2017; De Lange et al., 2018). Peripheral vision is more limited in terms of visual acuity than foveal vision and might involve lower precision predictions than cortical areas processing foveal representations. Nevertheless, our data suggest that peripheral vision encodes its inputs in a context-dependent manner, even when that context is not necessary for the task. This process could serve to explain away information in the periphery during navigation for example, where we could use contextual clues to filter out predictable features that we do not need to attend to. In the future, studies should observe in what measure our results are generalizable to other features and complex displays, potentially introducing multimodal effects. Investigating the neuronal bases of the current findings will also be necessary toward understanding how the visual system encodes numerical magnitude.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the College of Science and Engineering of the University of Glasgow. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

ML, CA, and LM designed the study together. ML and CA acquired, analysed, and modelled the data. All authors wrote and edited the manuscript.

## ACKNOWLEDGMENTS

We thank the Human Brain Project for support (Awarded to LM; HBP: H2020 FETFLAGSHIP Nos. 785907 and 945539 (Human Brain Project SGA2 and SGA3).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.750417/full#supplementary-material

## REFERENCES

Acerbi, L., Vijayakumar, S., and Wolpert, D. M. (2014). On the origins of suboptimality inhuman probabilistic inference. *PLoS Comput. Biol.* 10:e1003661. doi: 10.1371/journal.pcbi.1003661

Anobile, G., Cicchini, G. M., and Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychol. Sci.* 25, 265–270. doi: 10.1177/0956797613501520

Balas, B., Nakano, L., and Rosenholtz, R. (2009). A summary-statistic representation inperipheral vision explains visual crowding. *J. Vis.* 9:13. doi: 10.1167/9.12.13

Berry, D. A., and Fristedt, B. (1985). *Bandit Problems: Sequential Allocation of Experiments (Monographs on Statistics and Applied Probability)*, Vol. 5. London: Chapman and Hall, 7–7.

Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., and Gibbon, J. (2001). Numerical subtractionin the pigeon: evidence for a linear subjective number scale. *Psychol. Sci.* 12, 238–243. doi: 10.1111/1467-9280.00342

Burr, D., and Ross, J. (2008). A visual sense of number. *Curr. Biol.* 18, 425–428.

Chong, S. C., and Treisman, A. (2003). Representation of statistical properties. *Vis. Res.* 43, 393–404.

Craig, A. (1976). Signal recognition and the probability-matching decision rule. *Percept. Psychophys.* 20, 157–162. doi: 10.3758/BF03198594

De Lange, F. P., Heilbron, M., and Kok, P. (2018). How do expectations shape perception? *Trends Cogn. Sci.* 22, 764–779. doi: 10.1016/j.tics.2018.06.002

Edwards, G., Vetter, P., McGruer, F., Petro, L. S., and Muckli, L. (2017). Predictive feedback to V1 dynamically updates with sensory input. *Sci. Rep.* 7:16538.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in astatistically optimal fashion. *Nature* 415, 429–433. doi: 10.1038/415429a

Ernst, M. O., and Bülthoff, H. H. (2004). Merging the senses into a robust percept. *Trends Cogn. Sci.* 8, 162–169. doi: 10.1016/j.tics.2004.02.002

Fechner, G. T. (1860). *Elemente der Psychophysik*, Vol. 2. Leipzig: Breitkopf und Härtel.

Fornaciai, M., Brannon, E. M., Woldorff, M. G., and Park, J. (2017). Numerosity processing inearly visual cortex. *Neuroimage* 157, 429–438. doi: 10.1016/j.neuroimage.2017.05.069

Friston, K. (2002). Functional integration and inference in the brain. *Prog. Neurobiol.* 68, 113–143. doi: 10.1016/s0301-0082(02)00076-x

Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622

Gagnier, K. M., and Intraub, H. (2012). When less is more: line drawings lead to greater boundary extension than do colour photographs. *Vis. Cogn.* 20, 815–824. doi: 10.1080/13506285.2012.703705

Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*, Vol. 1. New York, NY: Wiley, 1969–1972.

Hansmann-Roth, S., and Mamassian, P. (2017). A glossy simultaneous contrast: conjoint measurements of gloss and lightness. *Iperception* 8:2041669516687770. doi: 10.1177/2041669516687770

Hansmann-Roth, S., Pont, S. C., and Mamassian, P. (2018). Contextual effects in humangloss perception. *Electronic Imaging* 2018, 1–7.

Hillis, J. M., Watt, S. J., Landy, M. S., and Banks, M. S. (2004). Slant from texture and disparity cues: optimal cue combination. *J. Vis.* 4, 967–992.

Ho, Y.-X., Landy, M. S., and Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychol. Sci.* 19, 196–204.

Hohwy, J., Roepstorff, A., and Friston, K. (2008). Predictive coding explains binocular rivalry: an epistemological review. *Cognition* 108, 687–701. doi: 10.1016/j.cognition.2008.05.010

Hubbard, T. L., Hutchison, J. L., and Courtney, J. R. (2010). Boundary extension: findings and theories. *Q. J. Exp. Psychol.* 63, 1467–1494. doi: 10.1080/17470210903511236

Jordan, K. E., and Brannon, E. M. (2006). Weber's Law influences numerical representations in rhesus macaques (*Macaca mulatta*). *Anim. Cogn.* 9, 159–172. doi: 10.1007/s10071-006-0017-8

Kersten, D., and Yuille, A. (2003). Bayesian models of object perception. *Curr. Opin. Neurobiol.* 13, 150–158. doi: 10.1016/s0959-4388(03)00042-4

Knoblauch, K., and Maloney, L. T. (2012). *Modeling Psychophysical Data in R*, Vol. 32. Heidelberg: Science & Business Media.

Landy, M. S., Goutcher, R., Trommershäuser, J., and Mamassian, P. (2007). Visual estimationunder risk. *J. Vis.* 7:4.

Larson, A. M., and Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *J. Vis.* 9:6. doi: 10.1167/9.10.6

Levi, D. M., Hariharan, S., and Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: peripheral crowding is neither size invariant nor simple contrast masking. *J. Vis.* 2:3. doi: 10.1167/2.2.3

Levick, W. R. (1967). Receptive fields and trigger features of ganglion cells in the visual streak of the rabbit's retina. *J. Physiol.* 188, 285–307. doi: 10.1113/jphysiol.1967.sp008140

Maloney, L. T., and Knoblauch, K. (2020). Measuring and modeling visual appearance. *Annu. Rev. Vis. Sci.* 6, 519–537. doi: 10.1146/annurev-vision-030320-041152

Mamassian, P., and Landy, M. S. (1998). Observer biases in the 3D interpretation of linedrawings. *Vis. Res.* 38, 2817–2832. doi: 10.1016/s0042-6989(97)00438-0

McDunn, B. A., Siddiqui, A. P., and Brown, J. M. (2014). Seeking the boundary of boundary extension. *Psychon. Bull. Rev.* 21, 370–375. doi: 10.3758/s13423-013-0494-0

Meng, M., and Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *J. Vis.* 4:539–551. doi: 10.1167/4.7.2

Merten, K., and Nieder, A. (2009). Compressed scaling of abstract numerosity representations in adult humans and monkeys. *J. Cogn. Neurosci.* 21, 333–346. doi: 10.1162/jocn.2008.21032

Morgan, A. T., Petro, L. S., and Muckli, L. (2019). Scene representations conveyed by cortical feedback to early visual cortex can be described by line drawings. *J. Neurosci.* 39, 9410–9423. doi: 10.1523/JNEUROSCI.0852-19.2019

Muckli, L., De Martino, F., Vizioli, L., Petro, L. S., Smith, F. W., Ugurbil, K., et al. (2015). Contextual feedback to superficial layers of V1. *Curr. Biol.* 25, 2690–2695. doi: 10.1016/j.cub.2015.08.057

Otten, M., Pinto, Y., Paffen, C. L., Seth, A. K., and Kanai, R. (2017). The uniformity illusion: central stimuli can determine peripheral perception. *Psychol. Sci.* 28, 56–68. doi: 10.1177/0956797616672270

Park, J., DeWind, N. K., Woldorff, M. G., and Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cereb. Cortex* 26, 748–763.

Peirce, J. W. (2007). PsychoPy–psychophysics software in Python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017

Rao, R. P., and Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2:79. doi: 10.1038/4580

Revina, Y., Petro, L. S., and Muckli, L. (2018). Cortical feedback signals generalise across different spatial frequencies of feedforward inputs. *Neuroimage* 180, 280–290. doi: 10.1016/j.neuroimage.2017.09.047

Ross, J. (2003). Visual discrimination of number without counting. *Perception* 32, 867–870. doi: 10.1068/p5029

Scharff, A., Palmer, J., and Moore, C. M. (2011). Extending the simultaneous-sequential paradigm to measure perceptual capacity for features and words. *J. Exp. Psychol. Hum. Percept. Perform.* 37:813. doi: 10.1037/a0021440

Smith, F. W., and Muckli, L. (2010). Nonstimulated early visual areas carry information about surrounding context. *Proc. Natl. Acad. Sci. U.S.A.* 107, 20099–20103. doi: 10.1073/pnas.1000233107

Spratling, M. W. (2017). A review of predictive coding algorithms. *Brain Cogn.* 112, 92–97. doi: 10.1016/j.bandc.2015.11.003

Steele-Russell, I., Russell, M. I., Castiglioni, J. A., and Graham, J. (2012). Differential retinal origins of separate anatomical channels for pattern and motion vision in rabbit. *Exp. Brain Res.* 222, 99–111. doi: 10.1007/s00221-012-3198-1

Stewart, E. E., Valsecchi, M., and Schütz, A. C. (2020). A review of interactions between peripheral and foveal vision. *J. Vis.* 20:2. doi: 10.1167/jov.20.12.2

Toscani, M., Gegenfurtner, K. R., and Valsecchi, M. (2017). Foveal to peripheral extrapolation of brightness within objects. *J. Vis.* 17:14. doi: 10.1167/17.9.14

Valsecchi, M., Toscani, M., and Gegenfurtner, K. R. (2013). Perceived numerosity is reduced in peripheral vision. *J. Vis.* 13:7. doi: 10.1167/13.13.7

Whiteley, L., and Sahani, M. (2008). Implicit knowledge of visual uncertainty guides decisions with asymmetric outcomes. *J. Vis.* 8:2. doi: 10.1167/8.3.2

Wood, S. N. (2015). *Core Statistics*, Vol. 6. Cambridge: Cambridge University Press.

Yiğit-Elliott, S., Palmer, J., and Moore, C. M. (2011). Distinguishing blocking from attenuation in visual selective attention. *Psychol. Sci.* 22, 771–780. doi: 10.1177/0956797611407927

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 15 November 2021
doi: 10.3389/fnhum.2021.751098

# Resources Underlying Visuo-Spatial Working Memory Enable Veridical Large Numerosity Perception

*Elisa Castaldi*[1,2]*, *Manuela Piazza*[3] *and Evelyn Eger*[4]

[1] *Department of Neuroscience, Psychology, Pharmacology and Child Health, University of Florence, Florence, Italy,*
[2] *Department of Translational Research and New Technologies in Medicine and Surgery, University of Pisa, Pisa, Italy,*
[3] *Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy,* [4] *Cognitive Neuroimaging Unit, INSERM, CEA DRF/JOLIOT, Université Paris-Saclay, NeuroSpin Center, Gif-sur-Yvette, France*

Humans can quickly approximate how many objects are in a visual image, but no clear consensus has been achieved on the cognitive resources underlying this ability. Previous work has lent support to the notion that mechanisms which explicitly represent the locations of multiple objects in the visual scene within a mental map are critical for both visuo-spatial working memory and enumeration (at least for relatively small numbers of items). Regarding the cognitive underpinnings of large numerosity perception, an issue currently subject to much controversy is why numerosity estimates are often non-veridical (i.e., susceptible to biases from non-numerical quantities). Such biases have been found to be particularly pronounced in individuals with developmental dyscalculia (DD), a learning disability affecting the acquisition of arithmetic skills. Motivated by findings showing that DD individuals are also often impaired in visuo-spatial working memory, we hypothesized that resources supporting this type of working memory, which allow for the simultaneous identification of multiple objects, might also be critical for precise and unbiased perception of larger numerosities. We therefore tested whether loading working memory of healthy adult participants during discrimination of large numerosities would lead to increased interference from non-numerical quantities. Participants performed a numerosity discrimination task on multi-item arrays in which numerical and non-numerical stimulus dimensions varied congruently or incongruently relative to each other, either in isolation or in the context of a concurrent visuo-spatial or verbal working memory task. During performance of the visuo-spatial, but not verbal, working memory task, precision in numerosity discrimination decreased, participants' choices became strongly biased by item size, and the strength of this bias correlated with measures of arithmetical skills. Moreover, the interference between numerosity and working memory tasks was bidirectional, with number discrimination impacting visuo-spatial (but not verbal) performance. Overall, these results suggest that representing visual numerosity in a way that is unbiased by non-numerical quantities relies on processes which explicitly segregate/identify the locations of multiple objects that are shared with visuo-spatial (but not verbal) working memory. This shared resource may potentially be impaired in DD, explaining the observed co-occurrence of working memory and numerosity discrimination deficits in this clinical population.

**Keywords: numerosity perception, visuo-spatial working memory, approximate number system, saliency map, developmental dyscalculia, arithmetic**

***Edited by:***
*Michele Fornaciai,*
*International School for Advanced Studies (SISSA), Italy*

***Reviewed by:***
*Tokita Midori,*
*Mejiro University, Japan*
*Darko Odic,*
*University of British Columbia, Canada*

***Correspondence:***
*Elisa Castaldi*
*elisa.castaldi@gmail.com*

***Specialty section:***
*This article was submitted to Cognitive Neuroscience, a section of the journal Frontiers in Human Neuroscience*

***Received:*** *31 July 2021*
***Accepted:*** *21 October 2021*
***Published:*** *15 November 2021*

***Citation:***
*Castaldi E, Piazza M and Eger E (2021) Resources Underlying Visuo-Spatial Working Memory Enable Veridical Large Numerosity Perception. Front. Hum. Neurosci. 15:751098. doi: 10.3389/fnhum.2021.751098*

# INTRODUCTION

Extracting estimates of the number of objects in a visual scene is important to guide many of our daily decisions. Much evidence suggests that numerosity perception is spontaneous and based on a non-verbal capacity which allows for judgments of the approximate number of objects at a glance, commonly termed "number sense" (Dehaene, 1997; Cantlon et al., 2009; Halberda et al., 2012; Anobile et al., 2016). Nevertheless, the precise perceptual and cognitive resources underlying this ability remain controversial.

For small numerosities within the subitizing range, i.e., up to four items, numerosity judgments are typically much more accurate compared to those for larger numbers (Kaufman et al., 1949; Mandler and Shebo, 1982; Revkin et al., 2008). However, the higher precision for small numerosities appears to depend on the availability of "domain general" cognitive resources such as working memory and attention (Burr et al., 2010; Melcher and Piazza, 2011; Piazza et al., 2011). One study found that the enumeration accuracy for small quantities in the subitizing range was affected by a concurrent visuo-spatial working memory task and that, vice versa, accuracy on the visuo-spatial task was lower when there were many compared to few items to enumerate (Piazza et al., 2011). This mutual interference, and the similar capacity limits measured across tasks, were interpreted to suggest that both visuo-spatial working memory and enumeration of small numbers of items may be supported by a basic mechanism of "visual indexing" of multiple objects, that is, a mechanism which allows us to simultaneously attend to multiple objects in parallel and explicitly represent their positions. This mechanism has been hypothesized to correspond to a mental map of the locations of salient objects in the visual scene, also referred to by the term "visual saliency map" (Koch and Ullman, 1987; Itti and Koch, 2000). In such a map, the saliency of individual objects is thought to be determined by either bottom-up (e.g., visual contrast) or top-down (e.g., task relevance) factors, and capacity limits of the map are flexible and determined by competitive interactions. In accordance with predictions from salience map theories, making one particular item more salient was found to reduce the subitizing range as well as memory performance for all other less salient items (Melcher and Piazza, 2011). The degree of involvement of such resources related to visuo-spatial working memory in the perception of larger numerosities remains unclear: although a small effect of performing a concurrent visuo-spatial working memory task was also found to decrease the precision of discrimination of larger numerosities (10–44 dots) by Piazza et al. (2011), the absence of a dependence of this effect on working memory load as well as the lack of another control task made it difficult to specifically attribute this effect to visuo-spatial working memory resources as opposed to more non-specifically enhanced cognitive load during dual task performance.

One issue which recently has given rise to much controversy on the cognitive underpinnings of "number sense" is the fact that performance in numerosity discrimination tasks can often non-veridically reflect the discrete number of items and instead be influenced by non-numerical properties of the sets (such as total luminance, area, density, and so on). When making non-numerical dimensions uninformative for the numerosity judgment, numerosity can still be discriminated, however, with lower accuracy when non-numerical dimensions vary incongruently with numerosity (e.g., Hurewitz et al., 2006; Nys and Content, 2012; Szûcs et al., 2013b; Salti et al., 2016). Moreover, non-numerical dimensions can bias behavioral choices in numerosity discrimination tasks, leading to consistent over or under estimation of numerosity. In adults these effects are typically subtle and mostly arise when the numerical ratios compared are rather small (Tokita and Ishiguchi, 2010; Nys and Content, 2012; DeWind et al., 2015) and become more evident when the variation in non-numerical dimensions is perceptually more salient than the numerical one (Hurewitz et al., 2006; Tokita and Ishiguchi, 2010; Gebuis and Reynvoet, 2011, 2012a,b). Recent studies have proposed that during development and/or arithmetical learning, children progressively learn to "focus on number" and to discard the influence of non-numerical quantities (Starr et al., 2017; Piazza et al., 2018). Interestingly, individuals with developmental dyscalculia (DD), a specific learning disability that prevents them from learning numerical and arithmetical skills (American Psychiatric Association, 2013), have not only been found to show decreased numerosity precision (Piazza et al., 2010; Mazzocco et al., 2011; Mejias et al., 2012; Anobile et al., 2018; Decarli et al., 2020) but be particularly impaired when non-numerical quantities provide incongruent information which tends to strongly bias their judgments (Szûcs et al., 2013a; Bugden and Ansari, 2016; Castaldi et al., 2018; Piazza et al., 2018). Some authors have attributed such findings to deficits in executive functions and more specifically, problems in inhibiting responses to task-irrelevant dimensions of the stimuli (Gilmore et al., 2013; Szûcs et al., 2013a; Bugden and Ansari, 2016). Nevertheless, our own findings in adults with DD showed that enhanced interference from unattended quantities was present only during numerosity comparisons, but not when subjects had to compare an orthogonal dimension (average item size) of the same stimuli (Castaldi et al., 2018), arguing against an impairment in general inhibitory skills as the source of the underlying problem.

More generally, it has been observed that difficulties in DD individuals span beyond the specific domain of numerical cognition: both DD children and adults often present working memory, attention and cognitive control deficits (Ashkenazi et al., 2013; Szûcs et al., 2013a; Menon, 2016; Castaldi et al., 2018, 2020b; Mammarella et al., 2018; Decarli et al., 2020, for reviews see: Fias et al., 2013; Iuculano, 2016; Castaldi et al., 2020a). With respect to working memory, a recent metanalysis showed that visuo-spatial working memory deficits characterize the "pure" DD subtype with respect to profiles with associated reading deficit (comorbid dyslexic dyscalculic disability), which are instead frequently associated with weak verbal working memory (Szûcs, 2016). Interestingly, Bugden and Ansari (2016) found that differences in numerosity precision and error rate during trials with incongruent non-numerical properties correlated with visuo-spatial working memory performance in DD children.

In the context of the reviewed findings on the observed co-existence of enhanced susceptibility to bias from non-numerical quantities and visuo-spatial working memory impairments in dyscalculia, together with the earlier mentioned evidence for potential shared resources between visuo-spatial working memory and enumeration at least of small numbers of items, we speculated that resources involved in visuo-spatial working memory might also be crucial for representing larger numerosities without bias from non-numerical quantities. Specifically, mechanisms as assumed by theories related to salience maps that explicitly represent the locations of multiple items and thereby allow for object segregation (as opposed to mechanisms that encode the visual scene in a mere gist-like, undifferentiated fashion) might be required to extract an unbiased representation of discrete numbers of items and such a representation might be relevant for arithmetical learning. While the co-occurrence of visuo-spatial working memory impairments and enhanced susceptibility to bias from non-numerical quantities in DD could still be explained by coexisting but functionally unrelated cognitive phenomena, we reasoned that to support a causal role of visuo-spatial working memory resources in veridical representation of large numerosities, manipulating the availability of working memory resources in neurotypical adults during performance of a numerosity task should interfere with numerosity judgments and lead to increasing perceptual biases from non-numerical stimulus dimensions.

In the current study, we therefore adapted the paradigm previously used to study numerosity perception and interference from non-numerical quantities in DD adults and controls by Castaldi et al. (2018), so that the numerosity discrimination task was performed either in isolation or in the context of a concurrent working memory task, and evaluated both numerosity precision and interference from the unattended size dimension (perceptual biases). Going beyond previous demonstrations of effects of working memory load on enumeration accuracy in the context of small sets of items, we further explicitly tested the specificity of the observed interference to the type of working memory resources: Given that visuo-spatial, but not verbal, working memory requires representing spatial locations of multiple items in parallel, we predicted that loading specifically visuo-spatial (but not verbal) working memory should give rise to imprecise and biased numerosity judgments. Moreover, if the systems supporting visuo-spatial working memory and numerosity perception share common resources, we expect a bidirectional interference between numerosity discrimination and working memory performance. Finally, if the shared resource contributing to visuo-spatial working memory and to veridical perception of numerosity is relevant also for more abstract arithmetical abilities, inter-individual differences in measures of arithmetical abilities should be predicted by the numerosity biases measured while participants' visuo-spatial, but not verbal, working memory was loaded.

## MATERIALS AND METHODS

Twelve adults with normal or corrected to normal vision (age 24 ± 3, 6 females) were included in the study. Prior to the study, written informed consent was obtained from all participants in accordance with the Declaration of Helsinki, and the study was approved by the research ethics committee of University Paris-Saclay. Prior to the study participants were asked whether they have ever encountered problems in learning math (or other school achievements, such as reading or writing), to qualitatively evaluate whether they might present learning disabilities. None of the participants reported having ever had such difficulties.

Participants were tested with seven conditions in which they performed different tasks: one baseline numerosity discrimination task, two single working memory tasks, two single and two dual tasks where numerosity discrimination was probed either in isolation or together with working memory, during presentation of identical stimuli (detailed below).

Participants sat in a dimly lit room at approximately 60 cm from a 15-inch Laptop (HP) with LCD monitor running at 60 Hz and with 1600 × 900 resolution. Visual stimuli were viewed binocularly and were generated under Matlab using PsychToolbox routines (Brainard, 1997).

### Baseline Numerosity Discrimination Task

The aim of this first condition was to measure participants' numerosity discrimination performance at baseline, while being presented with no visual stimuli other than the ones for which a numerosity judgment was required. Stimuli were the same as the ones used in the numerosity task by Castaldi et al. (2018). Participants were presented with two heterogenous arrays of dots, half black and half white, displayed on a gray background so that luminance was not a cue for number. The arrays were simultaneously presented at the two sides of a central fixation point at 6 visual degrees (°) of eccentricity along the horizontal meridian. Individual dots were constrained to fall within a virtual circle of 5.8° or 7.6° diameter, to not overlap with the fixation point and to be at least 0.25° apart from each other. The test arrays contained 5, 6, 8, 12, 17, and 20 dots (ratios 0.5, 0.6, 0.8, 1.2, 1.7, 2 with respect to the reference of 10 dots). The arrays had either small (0.25°) or big (0.5°) average item diameter. Test stimuli were compared against a reference stimulus of 10 dots with 0.35° average item diameter with the same total field area as the test. Test and reference stimuli were presented either to the right or to the left of the central fixation point. The two arrays of dots were presented for 200 ms, followed by two questions. The first question asked the participant to report which of the two stimuli appeared more numerous (the question "number?" appeared onscreen). Participants were instructed to press either the left or the right arrow to provide the response. Then, a second question "same or different?" appeared onscreen, but in the current task participants were instructed to ignore it and to press the spacebar to move on to the next trial. Participants performed 12 practice trials, after which the experiment started. No feedback was provided during the practice trials, nor during the following experimental runs. Each participant performed three sessions. Each one of the 6 comparison ratios was presented 48 times: 2 average item size (small and big), 2 possible total field areas, 2 possible spatial positions with respect to the reference (left-right) repeated 2 times in each one of the 3 sessions. A total of 288 trials were collected and used for the analysis.

## Single Working Memory Tasks

Participants then performed two single working memory tasks that aimed at measuring the number of elements that could be held in both verbal and visuo-spatial working memory. The number of elements selected in these single working memory tasks was then used in the dual tasks described in the next section. The two tasks had the same structure but required participants to hold in memory a different type of information (either visuo-spatial or verbal). Participants viewed a first set of elements that had to be held in memory (they were either colored squares or letters), followed by a display showing two arrays of dots (that had to be ignored) and then by a second set of elements, to be compared with the first one. Next, two questions appeared onscreen: the first question "number?" had to be ignored and bypassed by pressing the spacebar, then the second question "same or different?" appeared onscreen and participants were asked to judge whether the second set of elements displayed and the one held in memory were identical or not by pressing the letter "e" or "x" on the keyboard, respectively.

In the visuo-spatial working memory task (**Figure 1A**), the stimuli consisted in arrays of either two or four squares (0.4° side) of different colors (selected randomly between red, green, blue, yellow, and cyan). The number of squares presented (either two or four) was varied to modulate the working memory load (either low or high load, respectively). The arrays of squares were presented in the center of the screen for 500 ms. On every trial the spatial location of the squares was randomly selected within a virtual circle of 3° diameter, not overlapping with the fixation point. In half of the trials the second array of squares was identical to the first one, while in the other half the color of one square was changed. In the verbal working memory task (**Figure 1B**), participants were presented with sequences of either two or four letters (corresponding to low or high load, respectively). The letters were randomly selected between A, B, C, D, E, and F and were presented just above the fixation point. Each letter stayed onscreen for 500 ms and was immediately replaced by the following one. In half of the trials the second sequence of letters was identical to the first one, while in the other half one letter was replaced by another one not yet presented. Participants performed 96 trials for each task, after which data were analyzed to check that performance was approximately comparable across tasks for both the low and high load conditions (corresponding to the comparison of two or four elements). If the proportion of correct responses in the verbal working memory task largely differed from the one obtained in the visuo-spatial working memory task, then the number of letters displayed was decreased or increased by one element and the participant was tested again with the verbal working memory task. The difference between the two tasks was minimized by selecting, for each participant, the number of letters that allowed equating accuracy across the two tasks. This resulted in selecting 2 letters for the low load condition (except for two participants for which 3 letters were selected) and 4 letters for the high load condition (except for one participant for which 3 letters were selected and three participants for which 5 letters were selected). With this selection the proportion of correct responses between the two tasks did not differ by more than 0.1 (in both directions).

## Single Numerosity and Dual Tasks

After the baseline numerosity and single working memory tasks, participants performed two single numerosity and two dual task experiments in randomized order (**Figures 1A,B**). The structure of the tasks was the same as for the previously described single working memory tasks: participants were presented with the first set of elements (either squares or letters), then with two arrays of dots and finally with the second set of elements. Following the presentation of these stimuli, the two questions "number?" and "same or different?" appeared onscreen. In the two single tasks, participants had to ignore the set of stimuli (either squares or letters) presented both before and after the arrays of dots and respond to the question "number?" by pressing the left or right arrow to indicate which array contained more dots. The question "same or different?" had then to be ignored by pressing the spacebar. These two single numerosity tasks differed from

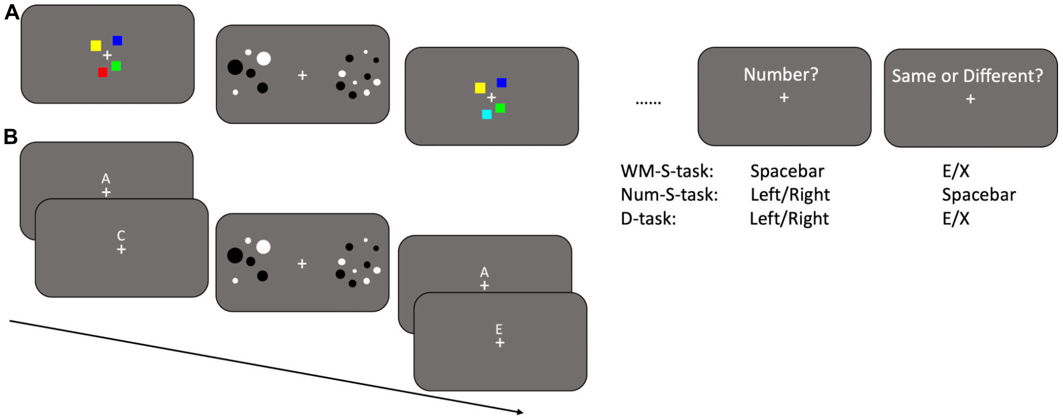


**FIGURE 1 |** Paradigm and stimuli. Representation of the stimuli used in the visuo-spatial single and dual tasks **(A)** and verbal single and dual tasks **(B)**. Possible responses in the single working memory tasks (WM-S-task), single numerosity (Num-S-task) and dual (D-task) tasks are shown on the right below the relative questions.

the baseline numerosity task by the presence of visual stimuli (squares and letters) presented before and after the arrays of dots.

In the two dual tasks, participants were instructed to perform the primary task, that is determining which array of dots was more numerous and then to perform the secondary task in which they had to say whether the first and the second sets of stimuli (either squares or letters) presented before and after the arrays of dots were identical or not. The response to the primary task had to be provided after the question "number?" by pressing the left or right arrow, while the response to the secondary task had to be provided after the question "same or different?" by pressing the letter "e" or "x" on the keyboard.

For each task (two single and two double tasks), participants performed three sessions, with the same number of trials as the ones detailed in the baseline numerosity discrimination task. Half of the trials tested the low load condition, and the other half the high load condition in the secondary task.

## Arithmetic Test

Participants were also tested with an arithmetic test taken from the Italian battery for developmental dyscalculia (Biancardi and Nicoletti, 2004). Participants were asked to respond as quickly and as accurately as possible to a set of single-digit arithmetical operations. The operations were orally presented by the experimenter, who also started the time recording with a chronometer as soon as the question was completely formulated. Time recording was stopped when participants spell out the result. The single-digit operations included 16 multiplications, 6 additions and 6 subtractions. Reaction time (RT) and response accuracy were used to calculate the inverse efficiency score (IES) as the mean RT divided by the proportion of correct responses.

## Analysis

For the working memory tasks, we calculated the proportion of correct responses after splitting the data for the two load conditions (low and high load) and we compared them by means of repeated measures ANOVAs and $t$-tests.

To quantify the precision of numerosity judgments, we plotted the percentage of test trials with "greater than reference" responses against the log-transformed difference between test and reference and fitted it with a cumulative Gaussian function using Psignifit toolbox (Schütt et al., 2016) available at https://github.com/wichmann-lab/psignifit. The point of subjective equality (PSE) was estimated at the 50% point, while the just noticeable difference (JND) was estimated as the difference between the 50 and 75% points.

Next, we estimated the perceptual bias to quantify the influence of the unattended dimension (average size) on the numerosity judgments. To this aim, we fitted participants' responses after splitting the dataset for the different magnitudes (small or big) of the unattended size dimension. This means that the "unattended small" and the "unattended big" trials had a small (0.25°) or big (0.5°) average item diameter, respectively. A shift of the psychometric curve away from 0 would indicate a bias from the unattended dimension, meaning that the arrays' average item size induced over- or underestimation of numerosity. For each participant, we fitted the data after splitting for the magnitude of the unattended dimension and calculated the difference (small-big) between the two PSE estimates (signed bias).

Previous studies found diverging results regarding the direction of the numerosity bias induced by item size, sometimes reporting overestimation (Hurewitz et al., 2006; Nys and Content, 2012) and sometimes underestimation (Ginsburg and Nicholls, 1988; Tokita and Ishiguchi, 2010; Gebuis and Reynvoet, 2012b) of number with big item size. Moreover, even within the same study, the direction of the bias is not always the same in all participants (DeWind et al., 2015; Castaldi et al., 2018). Following Castaldi et al. (2018), we additionally calculated the unsigned bias as a measure of the degree of the interference (irrespective of its direction) of the unattended dimension on the numerosity judgments, by taking the absolute value of the signed bias.

The precision of numerosity judgments and biases measured in the different tasks were compared by means of repeated measures ANOVAs and *post hoc* tests. One sample $t$-tests were used to evaluate whether signed biases were significantly different from 0. However, strong but opposite sign effects at the individual participant level could cancel each other out, leading to absence of average bias. We tested whether this was the case by performing individual participant analysis on signed biases. Psignifit toolbox allows to compute Bayesian confidence intervals (credible intervals) based on the posterior marginal densities of the psychometric curve's parameters. From individual participants' posterior distributions for unattended small and big's PSEs, we obtained the 95% credible interval of the difference and the probability p corresponding to 1- the confidence level for which the credible interval would include 0. If $p < 0.05$, 0 was outside the 95% credible interval and the given participant's bias was considered reliably different from 0.

Finally, on the data collected in the current experiment, we performed correlation analyses based on Pearson correlation to evaluate the relation between the bias and the participants' arithmetical performance defined as the IES measured with the arithmetic test.

To evaluate the reliability of the current results, we additionally performed Bayesian statistical analysis using JASP (JASP Team, 2020). Hypotheses were tested two-sided using a default prior distribution. For Bayesian ANOVA, models were ordered by their predictive performance relative to the best model. Inclusion Bayes factors resulting from the analysis of effects across "all matched models" are reported for main effects and interaction terms (van den Bergh et al., 2019). Bayes factors are reported in logarithmic base 10 units (LogBF) and their absolute values should be interpreted as providing anecdotal (0–0.5), substantial (0.5–1), strong (1–1.5), or very strong (>1.5) evidence, in favor of the alternative hypothesis if positive, or the null hypothesis if negative.

# RESULTS

## Comparison Between Baseline and Single Numerosity Tasks

In the baseline and in the two single tasks, participants performed a numerosity discrimination task, while ignoring other visual stimuli that were presented. We compared Weber fractions and biases induced by the unattended size dimension measured during numerosity discrimination and found that these variables did not differ across conditions, suggesting that the mere presence of visual stimuli in the two single tasks had no impact on numerosity judgments and these tasks could thus be considered as baseline conditions (see **Supplementary Analyses**).

## Comparison Between Single Numerosity and Dual Tasks

In the two dual tasks, participants viewed the same images that were shown in the single tasks and were instructed to perform both a numerosity discrimination task (primary task) and a working memory task (secondary task). In the working memory task, participants were asked to judge whether two sets of items presented before and after the dot arrays were the same or not. Items were either a set of squares (visuo-spatial working memory task) or a sequence of letters (verbal working memory task). The working memory tasks had two difficulty levels, requiring participants to hold in memory fewer or more items (low vs. high load conditions).

### Precision of Numerosity Judgments

We evaluated participants' precision in the numerosity discrimination task, as indexed by the Weber fraction (Wf). Performing a secondary task increased participants' Wfs, irrespective of the working memory load, and most strongly when participants performed the visuo-spatial compared to verbal working memory task (**Figure 2**). Specifically, Wfs measured in the visuo-spatial dual task were on average larger (low load: 0.18 ± 0.05; high load: 0.19 ± 0.06), compared to the ones measured during the verbal dual task (low load: 0.17 ± 0.05; high load: 0.15 ± 0.03), and both single tasks (visuo-spatial low load: 0.14 ± 0.04; visuo-spatial high load: 0.14 ± 0.04; verbal low load: 0.16 ± 0.03; verbal high load: 0.15 ± 0.03). We entered Wfs in a three-way repeated measures ANOVA with condition (2 levels: single vs. dual task), working memory type (2 levels: visuo-spatial vs. verbal working memory) and load (low vs. high load) as factors. There was a significant interaction between condition and working memory type [$F(1, 11) = 5.2$; $p = 0.04$, LogBF = 0.8], while the triple interaction between condition, working memory type and load [$F(1, 11) = 1.04$; $p = 0.33$, LogBF = –0.3], as well as the interaction between load and the other two factors [condition × load: $F(1, 11) = 0.09$; $p = 0.76$, LogBF = –0.5; working memory type × load: $F(1, 11) = 1.6$; $p = 0.23$, LogBF = –0.1] were not significant. *Post hoc* tests showed that Wfs measured in the visuo-spatial dual task were significantly larger with respect to those measured in the visuo-spatial single task [$t(11) = -3.9$, $p = 0.004$, LogBF = 0.8]. On the contrary, Wfs measured in the verbal dual task were not significantly larger with respect to the verbal single task [$t(11) = -0.4$, $p > 0.99$, LogBF = –0.4]. The Wf differences between the two single and the two dual tasks, respectively, were not significant [single tasks: $t(11) = 0.99$, $p > 0.99$, LogBF = –0.3; dual tasks: $t(11) = -2.2$, $p = 0.23$, LogBF = 0.3].

Overall, compared to the single tasks, participants' precision when estimating numerosity was affected by the concurrent dual task, especially when visuo-spatial and not verbal stimuli

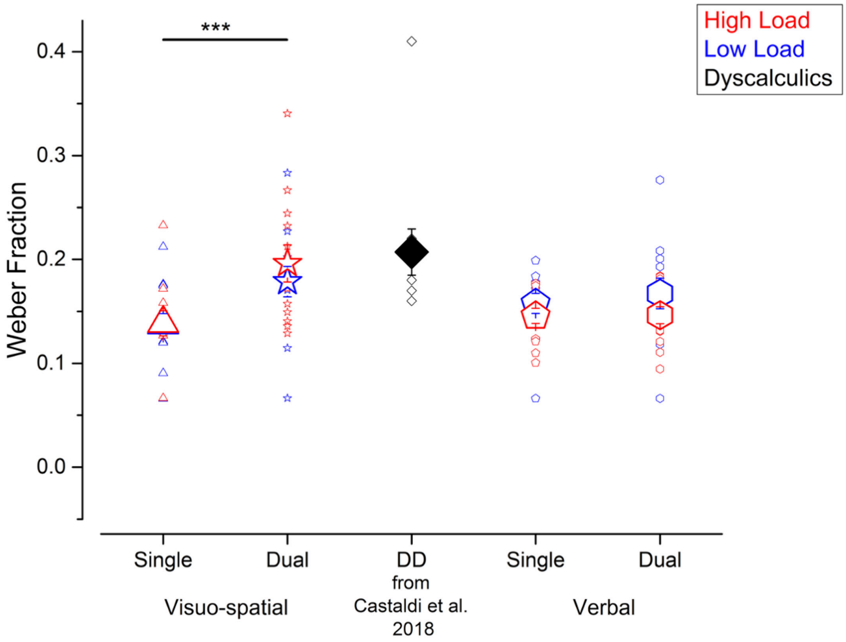


**FIGURE 2 |** Participants' precision in the numerosity discrimination tasks. Wfs measured in the different experiments performed in the current study (blue and red symbols) and for the DD participants tested in Castaldi et al. (2018) (black symbols, for qualitative comparison). Blue and red symbols identify the Wfs measured when fewer or more elements were presented before and after the arrays of dots (corresponding to the low and high load conditions in the working memory task). Large and small symbols indicate the average ± SEM and individual Wfs, respectively. ***$p < 0.005$.

had to be held in memory. On average, numerosity precision in the visuo-spatial dual task approached the one measured in a group of adults with DD measured in a previous study (Castaldi et al., 2018).

## Interference From the Unattended Size Dimension

Next, to test whether the unattended size dimension interfered with participants' judgments, we evaluated the biases defined as the signed difference of the PSEs for psychometric curves fitted using trials with small or big average item size. **Figure 3** shows that average signed biases did not seem to be affected by condition and were close to zero for both single (visuo-spatial low load: –0.04 ± 0.14; visuo-spatial high load: –0.009 ± 0.16; verbal low load: –0.02 ± 0.13; verbal high load: –0.04 ± 0.11) and dual (visuo-spatial low load: –0.08 ± 0.3; visuo-spatial high load: –0.004 ± 0.32; verbal low load: –0.05 ± 0.14; verbal high load: –0.04 ± 0.15) tasks. A three-way repeated measures ANOVA on signed biases with condition, working memory type and load as factors showed a significant interaction between working memory type and load [$F(1, 11) = 5.3$; $p = 0.04$, LogBF = –0.4], but *post hoc* tests were all not significant. Other interactions and main effects were not significant [condition: $F(1, 11) = 0.18$; $p = 0.68$, LogBF = –0.7; working memory type: $F(1, 11) = 0.005$, LogBF = –0.7; $p = 0.95$; load: $F(1, 11) = 2.1$; $p = 0.17$, LogBF = –0.5; condition × working memory type: $F(1, 11) = 9{*}10^{-5}$; $p = 0.99$, LogBF = –0.5; condition × load: $F(1, 11) = 1.09$; $p = 0.32$, LogBF = –0.4; condition x working memory type × load: $F(1, 11) = 3{*}10^{-5}$; $p = 0.99$, LogBF = –0.5]. One sample *t*-tests against zero were not significant (all ps > 0.1), consistent with biases measured in all conditions being close to zero on average across participants. However, the absence of a group average bias can be potentially due to strong but opposite effects at the single participant level which cancel each other out. Analysis at the individual participant level (using Bayesian statistics, see methods for details) showed that this was indeed the case for the visuo-spatial dual task: in this condition the signed bias was reliably different from 0 in six and eight participants for the low and high load trials, respectively. Within the participants who showed a bias reliably different from 0, four out of six participants and five out of eight participants for the low and high load conditions, respectively, tended to overestimate numerosity when the unattended size dimension was small and to underestimate numerosity when the unattended size dimension was big, while the remaining participants showed the opposite effect. In the other conditions, the signed biases were reliably different from 0 only in very few participants (3 participants for both loads of the visuo-spatial single task, 1 participant for both loads of the verbal single task and 2 and 1 participants for the low and high load conditions of the verbal dual task). Overall, these results suggest that although the average signed bias is close to 0 for all conditions, reliable effects of either positive or negative direction were observed in individual participants in the case of the visuo-spatial dual task.

These results shown by the current participants without DD in the visuo-spatial dual task condition resemble the ones previously obtained in a group of DD participants (Castaldi et al., 2018): the average signed bias in the DD group was not significantly different from zero, but analysis at the individual participant level

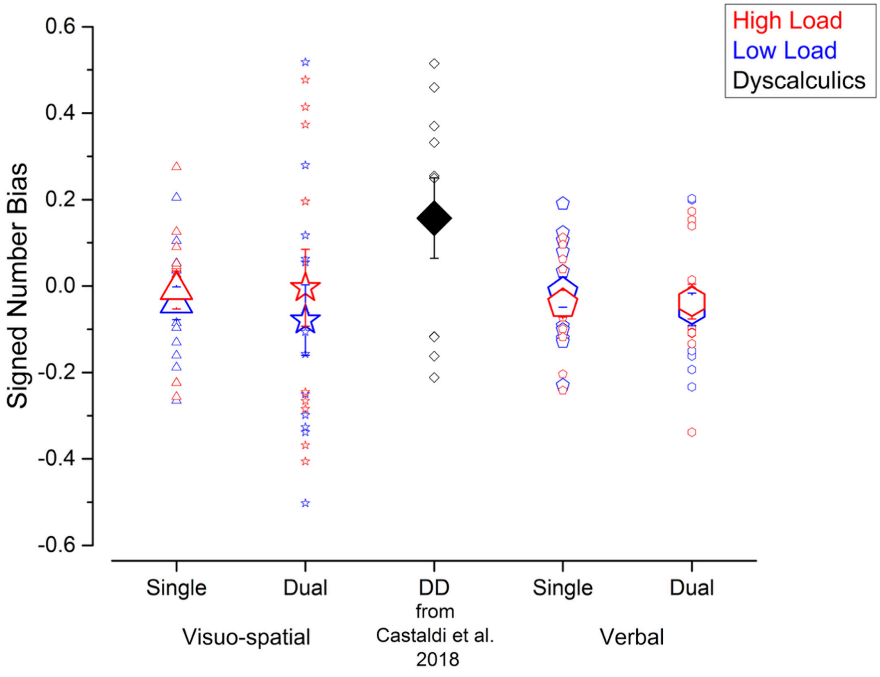


**FIGURE 3 |** Interference from the unattended dimension in the numerosity discrimination tasks – signed bias. Signed biases measured in the different experiments performed in the current study (blue and red symbols) and in a group of DD participants (black symbol, for qualitative comparison) tested by Castaldi et al. (2018). Symbol size and color correspond to the average and individual biases and to the different loads in the secondary task.

showed that the signed biases were reliably different from 0 in all DD participants.

To compare the strength of the biases irrespective of their direction, we next evaluated the unsigned biases defined as the absolute difference of the PSEs for psychometric curves fitted using trials with small or big average item size. Unsigned biases were affected by the secondary task, in a similar way as Wfs (**Figure 4**). Unsigned biases measured in the visuo-spatial dual task were on average higher (low load: 0.25 ± 0.16; high load: 0.28 ± 0.13) than the ones measured during the verbal dual task (low load: 0.12 ± 0.07; high load: 0.12 ± 0.09) and both the single task conditions (visuo-spatial low load: 0.11 ± 0.07; visuo-spatial high load: 0.12 ± 0.09; verbal low load: 0.10 ± 0.06; verbal high load: 0.09 ± 0.07). The three-way repeated measures ANOVA on unsigned biases with condition, working memory type and load as factors revealed a significant interaction between condition and working memory type [$F(1, 11) = 39.5$; $p < 0.001$, LogBF = 1.4], while the triple interaction between condition, working memory type and load [$F(1, 11) = 0.07$; $p = 0.79$, LogBF = –0.5], and the interaction between load and the other two factors [condition × load: $F(1, 11) = 0.21$; $p = 0.66$, LogBF = –0.5; working memory type x load: $F(1, 11) = 0.5$; $p = 0.49$, LogBF = –0.5] were not significant. *Post hoc* tests showed that the unattended size dimension biased participants' judgments significantly more during the visuo-spatial dual task than the visuo-spatial single task [$t(11) = -6.8$, $p < 0.001$, LogBF = 1.1]. On the contrary, unsigned biases were not significantly stronger when participants were involved in the verbal dual task with respect to the verbal single task [$t(11) = -0.9$, $p > 0.99$, LogBF = –0.5]. Unsigned biases were also significantly stronger when participants performed the visuo-spatial dual task compared to the verbal single [$t(11) = 4.6$, $p = 0.001$, LogBF = 0.7] and verbal dual tasks [$t(11) = 4.5$, $p = 0.003$, LogBF = 0.9]. Unsigned biases measured in the two single tasks did not differ from each other [$t(11) = -0.6$, $p > 0.99$, LogBF = –0.5).

Overall, compared to the other conditions measured here, the unattended size dimension biased participants' judgment most strongly during the visuo-spatial dual task, in which case the degree of bias approached the one previously observed in a group of adults with DD in a single task (Castaldi et al., 2018).

## Comparison Between Single and Dual Working Memory Tasks

Participants performed two single working memory tasks (WM-S-task) to select the number of elements subsequently used that matched the verbal and visuo-spatial working memory load (**Figure 5** light and dark gray bars). In both of these single working memory tasks, the proportion of correct responses was overall very high and comparable (visuo-spatial working memory task low load: 0.97 ± 0.03, high load: 0.94 ± 0.04; verbal working memory task low load: 0.98 ± 0.02, high load: 0.94 ± 0.05), confirming that task difficulty was successfully matched across the two systems (see **Supplementary Material**). Working memory performance in the dual task conditions was still relatively high (**Figure 5**, hatched bars), yet lower than that measured in the single working memory tasks (visuo-spatial

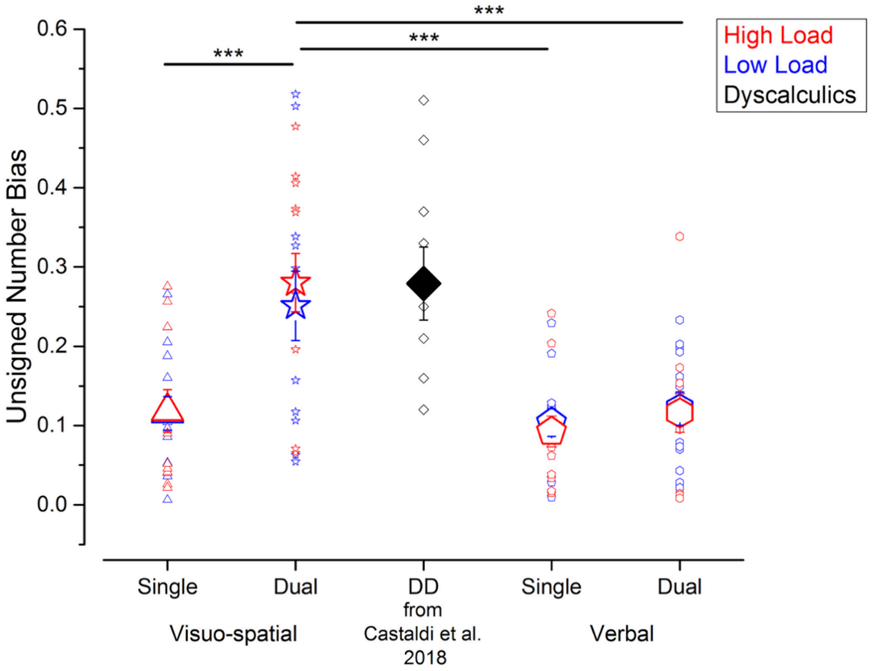


**FIGURE 4 |** Interference from the unattended dimension in the numerosity discrimination tasks – unsigned bias. Unsigned biases measured in the different experiments performed in the current study (blue and red symbols) and in a group of DD participants (black symbol, for qualitative comparison) tested by Castaldi et al. (2018). Symbol size and color correspond to the average and individual biases and to the different loads in the secondary task. ***$p < 0.005$.

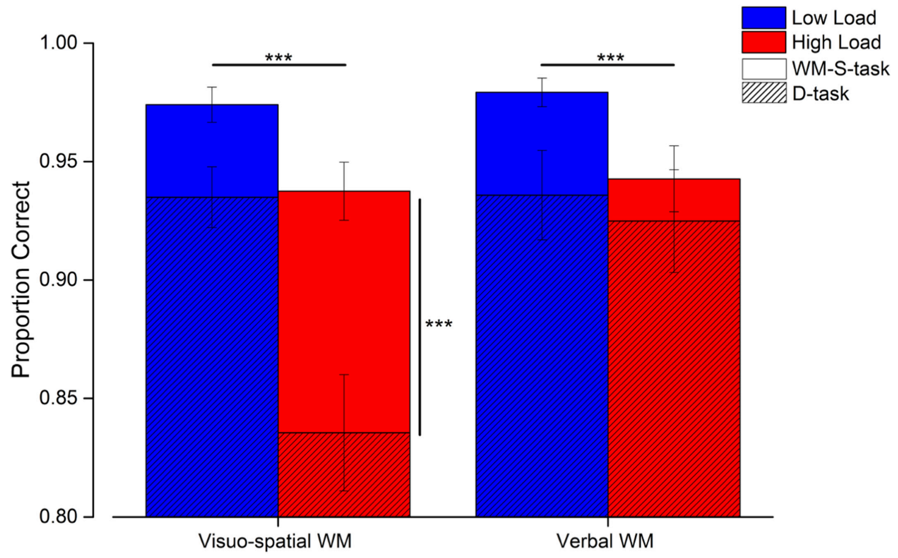


**FIGURE 5 |** Results of the working memory tasks. Average performance in the working memory tasks in the single (solid bars) and dual tasks (hatched bars) for both the low (blue) and high (red) load trials. Error bars are SEM. ***$p < 0.005$.

working memory task low load: 0.94 ± 0.05, high load: 0.84 ± 0.09; verbal working memory task low load: 0.94 ± 0.07, high load: 0.92 ± 0.08).

The proportion of correct responses made in the working memory tasks was entered in a three-way repeated measures ANOVA with condition (2 levels: single and dual task), working memory type (2 levels: visuo-spatial and verbal working memory task) and load (2 levels: low and high) as factors. The triple interaction between working memory type, load and condition was significant [$F(1, 11) = 21.7$; $p < 0.001$, LogBF = 0.8]. Subsequent *post hoc* tests showed that the proportion of correct responses decreased during the high (but not low) load trials of the visuo-spatial dual task with respect to the corresponding single task trials [high load single vs. dual task: $t(11) = 6.2$, $p < 0.001$, LogBF = 2.2; low load single vs. dual task: $t(11) = 2.4$, $p = 0.69$, LogBF = 1.1; **Figure 5**, vertical significant bar]. On the other hand, performing the verbal dual task did not significantly decrease the proportion of correct responses with respect to the verbal working memory single task, neither for the low nor for the high load trials [low load single vs. dual task: $t(11) = 2.6$, $p = 0.38$, LogBF = 0.3; high load single vs. dual task: $t(11) = 1.08$, $p > 0.99$, LogBF = –0.3]. In the dual tasks, the proportion of correct responses in the high load condition of the visuo-spatial working memory task was significantly lower with respect to the high load condition of the verbal working memory task [$t(11) = 6.2$, $p < 0.001$, LogBF = 1.6]. This difference was not significant for the low load trials [$t(11) = 0.06$, $p > 0.99$, LogBF = –0.5].

In sum, these results suggest that performing a concurrent numerosity task interferes with performance in a visuo-spatial working memory task, especially when load levels are relatively high, but this was not the case with a verbal working memory task.

## Correlation Analyses

Finally, we performed exploratory correlation analyses to test whether the tendency to show enhanced interference from non-numerical dimensions under concurrent working memory load was related to arithmetical abilities (**Figure 6**). We observed a significant positive correlation between the size of unsigned biases during the numerosity discrimination task and IE score for calculation, indicating that participants with better arithmetical abilities were those whose numerosity judgments were less biased by the unattended dimension ($r = 0.7$, $p = 0.02$, LogBF = 0.6, **Figure 6A**). Interestingly, arithmetic abilities were predicted by the size of unsigned biases only when numerosity discrimination was performed during the visuo-spatial, and not during the verbal, dual task: the correlation between unsigned bias during the verbal dual task and IE score for calculation was not significant ($r = –0.08$, $p = 0.8$, LogBF = –0.5, **Figure 6B**).

# DISCUSSION

In the current study we investigated whether resources required for visuo-spatial working memory that explicitly encode the locations of multiple items in the visual scene might also be crucial to precisely and veridically perceive larger numerosities without bias from non-numerical quantities. We therefore measured not only numerosity precision, but also perceptual biases (interference) during a dual task design, and directly tested which specific type of working memory (visuo-spatial vs. verbal) showed bidirectional interference with numerosity processing. We further explored whether these shared resources might also be relevant for arithmetical abilities.

In line with our hypothesis of a shared resource, we found that participants' threshold and the interference from the

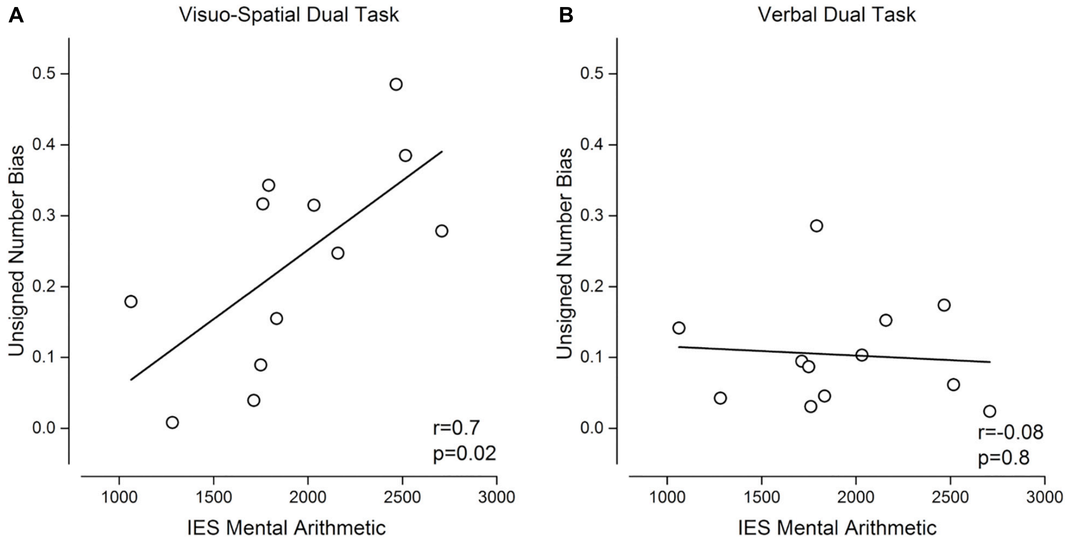


**FIGURE 6 |** Relation between unsigned bias and mental arithmetic skills. Pearson correlation between unsigned bias when discriminating numerosity during the visuo-spatial **(A)** or verbal **(B)** dual tasks and inverse efficiency score (IES, see section "Arithmetic Test") in mental arithmetic.

unattended dimension during a numerosity discrimination task increased when participants performed a concurrent visuo-spatial (but not verbal) working memory task. This interference was bidirectional: performing a numerosity discrimination task also degraded performance during the visuo-spatial working memory task. Finally, we found that the degree of interference from the unattended dimension on numerosity judgments during the concurrent visuo-spatial (but not verbal) task predicted interindividual differences in arithmetical skills.

Our results fit well with a recent study that also suggested a role of visuo-spatial working memory capacity in the extent to which participants' numerosity judgments rely on non-numerical dimensions (Lee and Cho, 2019). In that study, participants were assigned to low or high working memory groups based on their memory span (measured with dedicated tasks) and asked to numerically compare arrays of dots (12–40 dots) in which the non-numerical dimensions varied either congruently or incongruently with numerosity. Numerosity judgments of participants included in the low visuo-spatial working memory group were more influenced by non-numerical dimensions (size, total surface area and density) compared to those of participants in the high visuo-spatial working memory group, while the same was not observed when splitting the groups based on their verbal working memory span. While this plausibly suggested that the susceptibility to non-numerical interference during numerosity judgments depends on the capacity of the participants' visuo-spatial but not verbal working memory, uncontrolled domain general abilities other than working memory may have accidentally characterized the two subgroups tested in that study. Manipulating the engagement of visuo-spatial and verbal working memory resources in the same participants, as done in the current study, is therefore necessary to establish their relation with numerosity processing more unambiguously.

The present results also extend evidence from previous studies suggesting that a mechanism of visual indexing of multiple objects supports both visuo-spatial working memory and enumeration (Melcher and Piazza, 2011; Piazza et al., 2011; Knops et al., 2014) by showing that this mechanism might also operate at higher numerosities. The supposed underlying mechanisms of a saliency map has been simulated in computational studies using networks consisting of interconnected nodes which exhibit recurrent self-excitation and lateral inhibition (Roggeman et al., 2010; Knops et al., 2014; Sengupta et al., 2014). Each node corresponds to a neural population encoding an object location or feature and interacting with the other nodes through lateral inhibition. High levels of lateral inhibition lead to low noise levels and precise representations, but also to a small capacity of the map. On the contrary, lower levels of lateral inhibition lead to higher noise levels, coarser representations, and higher capacity limits. Thus, capacity limits are not fixed, but can vary depending on the representational precision required by the task, which can top-down modulate the level of lateral inhibition in the saliency map (Roggeman et al., 2010; Melcher and Piazza, 2011; Sengupta et al., 2014). The predictions of these models have found support in neurophysiological and fMRI studies (Bisley, 2003; Roggeman et al., 2010; Knops et al., 2014). The lateral intraparietal cortex (LIP) of macaque monkeys was found to represent the attended locations in the visual fields (Bisley, 2003) and its homolog region in humans showed signatures of saliency map models, e.g., by modulating voxels' response profiles depending on the representational precision required by the specific task at hand (Roggeman et al., 2010; Knops et al., 2014). Knops et al. (2014) showed participants a variable number of oriented Gabor gratings and asked them to either remember and compare their orientation or to enumerate them. They found that the average response profiles and the pattern of activation of the same set of voxels in the posterior parietal cortex (PPC) changed as a function of the task. While Knops et al. (2014) mainly investigated small numerosities, other studies have suggested that the capacity of the saliency maps need not be limited to few items within the subitizing range but can extend even to higher numerosities depending on task requirements and the level of top-down attention directed to individual items. Roggeman et al. (2010) found that when participant were asked to perform a numerosity estimation task on numbers above the subitizing range, the activity in IPS regions increased up to eight items and then slightly decreased (as for salience map models with medium inhibition settings), whereas when participants were performing a less demanding pattern detection task, the activity in IPS regions showed a V-shape, decreasing from four to eight or sixteen items and then increasing again (up to 64), as also found for a model with low inhibition settings. Finally, Sengupta et al. (2014) confirmed that changing the level of inhibition between nodes allowed the same network architecture to account for number discrimination in both the subitizing and estimation ranges. Altogether, these studies suggest that the PPC may host a mechanism which might be conceived as a saliency map contributing to both enumeration and visual working memory. This mechanism might allow to form a representation of the locations of a number of items, which would be coarser/less precise for larger numbers, but sufficient for extraction of approximate numerosity. In relation with the findings of the current study, we suggest that this mechanism led participants to localize and segregate items quite accurately during the baseline and single tasks. However, during the dual task this system might have been saturated by the need to precisely represent the stimuli of the visuo-spatial working memory task, leading participants' numerosity judgments to rely more on some kind of coarser, undifferentiated summary statistics representation of the visual arrays. For example, reliance on total energy or surface area might have led participants to overestimate numerosities with big dot sizes. On the contrary, reliance on the relative amount of energy in high and low spatial frequencies of the image, as predicted by a model linking numerosity to texture density processing (Dakin et al., 2011; Tibber et al., 2012), might have led them to overestimate numerosities with small dot sizes.

If the extraction of veridical numerosity estimates relies on indexing spatial locations of objects by a mechanism potentially involving recurrent processing, then it can be predicted that displaying stimuli for a longer time on screen should allow participants to discriminate numerosities more accurately and to provide less biased numerosity judgments. This has partially been observed: Inglis and Gilmore (2013) found that numerosity precision increased with the exposure to the stimulus display, and that this effect could not be explained by differences in the onset to decision latencies (and presumably not even by the adoption of counting strategies given that the effect was observed also for latencies below 1 s). Future studies should test whether longer presentation time also reduces interference from non-numerical dimensions on numerosity judgments. Further work may also manipulate the saliency of individual items and test in how far estimates of larger numerosities in such situations are indeed well explained by saliency map models or would require still somewhat different mechanisms.

The results of the current study clearly indicate that it is not any kind of working memory load irrespective of domain, but more specifically the visuo-spatial component which shares resources with numerosity judgments. The importance of visuo-spatial rather than verbal or auditory resources for a precise numerosity representation has been observed also during other tasks and for other cognitive functions than memory. For example, in a number line task, in which participants had to spatially map the relative location of arrays of different numerosities onto a line defining a numerical interval, participants' responses changed from being linearly distributed to logarithmic-like if they had to perform a concomitant visuo-spatial, but not auditory, task (Anobile et al., 2012a,b). Depriving visuo-spatial attention by means of attentional blink or dual tasks paradigms also affected the precision of numerosity estimation (Vetter et al., 2008; Burr et al., 2010; Anobile et al., 2012b; Pomè et al., 2019), while this was much less observed when the distractor task required directing attention to auditory stimuli (Pomè et al., 2019). Interestingly, although the strongest detrimental effects of attentional deprivation on numerosity estimation have been reported for low numerosities in the subitizing range, some smaller but consistent effects have been observed also for higher numerosities: Depriving visuo-spatial attention increased the degree of underestimation (Vetter et al., 2008; Burr et al., 2010; Anobile et al., 2012b) and decreased numerosity estimation precision also for numerosities beyond the subitizing range (Vetter et al., 2008; Pomè et al., 2019). Splitting visuo-spatial attention during numerosity adaptation by simultaneously presenting a numerically neutral adapter alongside with the real one led to underestimation of the real adaptor and to a consequent reduction of the adaptation effect (Grasso et al., 2021b,a).

The fact that deprivation of both visuospatial attention and working memory resources can affect aspects of numerosity perception fits well with a supposed functional overlap in the mechanisms of spatial working memory and spatial selective attention which may both be based on the same spatial saliency map (Awh and Jonides, 2001; Deco and Rolls, 2005). The current results extend the existing literature showing that engaging visuo-spatial working memory resources does not merely make numerosity estimates more noisy overall which could have been one possibility, but also increases the perceptual biases.

The fact that loading visuo-spatial working memory in neurotypical participants qualitatively simulated previous findings obtained in DD adults compared to controls and the finding that in the current study the bias correlated with interindividual differences in arithmetic abilities, make us speculate that the common resource that supports both visuo-spatial working memory and numerosity extraction may also play an important role in arithmetical learning, and be potentially impaired in DD. The limited visuo-spatial working memory capacity, the lower precision, and the enhanced reliance on non-numerical dimensions during numerosity discrimination tasks often observed in DD individuals, which are often separately emphasized by alternative and competing explanatory accounts of this disorder, could thus be interdependent phenomena and reflect a weakness of the same system. The fact that this system specifically supports visuo-spatial but not verbal working memory, is in line with the previously reported correlation between numerosity impairments and visuo-spatial, but not verbal, working memory performance in DD children (Bugden and Ansari, 2016). It is also in line with the selective impairment of visuo-spatial working memory characterizing the "pure" DD subtype (i.e., without associated reading problems, Szûcs, 2016).

The parietal regions exhibiting properties of a saliency map (Knops et al., 2014) are also modulated by attention to (high) numerosities as opposed to other non-numerical dimensions (Castaldi et al., 2019) and the pattern of activity read out from these regions correlates with numerosity precision (Lasne et al., 2018). Areas which are likely overlapping or nearby are recruited during visuo-spatial working memory and arithmetic tasks (Zago et al., 2008; Castaldi et al., 2020c; Matejko and Ansari, 2021) and present functional abnormalities in DD individuals during both magnitude discrimination and visuo-spatial working memory tasks (Price et al., 2007; Rotzer et al., 2009). It is thus in theory possible that the neural substrate of the common resource supporting visual working memory and numerosity extraction in parietal cortex is impaired in DD. Nevertheless, the fact that the present study in neurotypical adults yielded qualitatively similar findings to those previously observed in dyscalculics is not necessarily evidence for a shared cause. Future behavioral and imaging studies in dyscalculia may further test this possibility.

## CONCLUSION

In conclusion, the current study shows that estimating large numerosities veridically relies on resources that are also fundamental for visuo-spatial but not verbal working memory, which may relate to explicitly encoding the locations of multiple objects in the visual scene. Loading visuo-spatial working memory may saturate this system and lead participants' numerosity estimates to rely more on a coarse, gist-like representation of the visual input which is susceptible to the influence of non-numerical dimensions. Although speculative, it is possible to hypothesize that the difficulties experienced by DD individuals with both numerosity perception and working memory may result from the impairment of the same resources which would explain why low numerosity discrimination precision, enhanced reliance on non-numerical dimensions during numerosity judgments and impaired visuo-spatial working memory often co-occur.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of University Paris-Saclay. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

EC collected and analyzed the data. All authors contributed to the study concept, experimental design, interpretation of results, manuscript preparation, and approved the final version of the manuscript.

## FUNDING

This research was funded by the European Union's Horizon, 2020 Research and Innovation Program under the Marie Skłodowska-Curie (grant agreement no. 885672—DYSC-EYE-7T) "The neural substrate of numerical cognition in dyscalculia revealed by eye tracking and ultra-high field 7T functional magnetic imaging" awarded to EC and by INSERM, CEA, and the French National Research Agency (ANR) through (grant no. ANR-14-CE13-0020-01) awarded to EE.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.751098/full#supplementary-material

## REFERENCES

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.

Anobile, G., Cicchini, G. M., and Burr, D. (2012a). Linear mapping of numbers onto space requires attention. *Cognition* 122, 454–459. doi: 10.1016/j.cognition.2011.11.006

Anobile, G., Turi, M., Cicchini, G. M., and Burr, D. (2012b). The effects of cross-sensory attentional demand on subitizing and on mapping number onto space. *Vis. Res.* 74, 102–109. doi: 10.1016/j.visres.2012.06.005

Anobile, G., Cicchini, G. M., and Burr, D. (2016). Number as a primary perceptual attribute: a review. *Perception* 45, 5–31. doi: 10.1177/0301006615602599

Anobile, G., Cicchini, G. M., Gasperini, F., and Burr, D. (2018). Typical numerosity adaptation despite selectively impaired number acuity in dyscalculia. *Neuropsychologia* 120, 43–49. doi: 10.1016/j.neuropsychologia.2018.10.006

Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W. S., Swigart, A. G., and Menon, V. (2013). Visuo–spatial working memory is an important source of domain-general vulnerability in the development of arithmetic cognition. *Neuropsychologia* 51, 2305–2317. doi: 10.1016/j.neuropsychologia.2013.06.031

Awh, E., and Jonides, J. (2001). Overlapping mechanisms of attention and spatial working memory. *Trends Cogn. Sci.* 5, 119–126. doi: 10.1016/S1364-6613(00)01593-X

Biancardi, A., and Nicoletti, C. (2004). *Batteria per la Discalculia Evolutiva (BDE) [Battery for Developmental Dyscalculia (BDE)]*. Turin: Omega.

Bisley, J. W. (2003). Neuronal activity in the lateral intraparietal area and spatial attention. *Science* 299, 81–86. doi: 10.1126/science.1077395

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Bugden, S., and Ansari, D. (2016). Probing the nature of deficits in the 'Approximate Number System' in children with persistent Developmental Dyscalculia. *Dev. Sci.* 19, 817–833. doi: 10.1111/desc.12324

Burr, D., Turi, M., and Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *J. Vis.* 10, 20–20. doi: 10.1167/10.6.20

Cantlon, J. F., Platt, M. L., and Brannon, E. M. (2009). Beyond the number domain. *Trends Cogn. Sci.* 13, 83–91. doi: 10.1016/j.tics.2008.11.007

Castaldi, E., Mirassou, A., Dehaene, S., Piazza, M., and Eger, E. (2018). Asymmetrical interference between number and item size perception provides evidence for a domain specific impairment in dyscalculia. *PLoS One* 13:e0209256. doi: 10.1371/journal.pone.0209256

Castaldi, E., Piazza, M., Dehaene, S., Vignaud, A., and Eger, E. (2019). Attentional amplification of neural codes for number independent of other quantities along the dorsal visual stream. *eLife* 8:e45160. doi: 10.7554/eLife.45160

Castaldi, E., Turi, M., Gassama, S., Piazza, M., and Eger, E. (2020b). Excessive visual crowding effects in developmental dyscalculia. *J. Vis.* 20:7. doi: 10.1167/jov.20.8.7

Castaldi, E., Piazza, M., and Iuculano, T. (2020a). "Learning disabilities: developmental dyscalculia," in *Handbook of Clinical Neurology*, eds A. Gallagher, C. Bulteau, D. Cohen, and J. L. Michaud (Amsterdam: Elsevier), 61–75. doi: 10.1016/B978-0-444-64148-9.00005-3

Castaldi, E., Vignaud, A., and Eger, E. (2020c). Mapping subcomponents of numerical cognition in relation to functional and anatomical landmarks of human parietal cortex. *Neuroimage* 221, 117210. doi: 10.1016/j.neuroimage.2020.117210

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., and Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19552–19557. doi: 10.1073/pnas.1113195108

Decarli, G., Paris, E., Tencati, C., Nardelli, C., Vescovi, M., Surian, L., et al. (2020). Impaired large numerosity estimation and intact subitizing in developmental dyscalculia. *PLoS One* 15:e0244578. doi: 10.1371/journal.pone.0244578

Deco, G., and Rolls, E. T. (2005). Attention, short-term memory, and action selection: a unifying theory. *Prog. Neurobiol.* 76, 236–256. doi: 10.1016/j.pneurobio.2005.08.004

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. New York, NY: Oxford University Press.

DeWind, N. K., Adams, G. K., Platt, M. L., and Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition* 142, 247–265. doi: 10.1016/j.cognition.2015.05.016

Fias, W., Menon, V., and Szucs, D. (2013). Multiple components of developmental dyscalculia. *Trends Neurosci. Educ.* 2, 43–47.

Gebuis, T., and Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behav. Res. Methods* 43, 981–986. doi: 10.3758/s13428-011-0097-5

Gebuis, T., and Reynvoet, B. (2012a). The interplay between nonsymbolic number and its continuous visual properties. *J. Exp. Psychol. Gen.* 141, 642–648. doi: 10.1037/a0026218

Gebuis, T., and Reynvoet, B. (2012b). The role of visual information in numerosity estimation. *PLoS One* 7:e37426. doi: 10.1371/journal.pone.0037426

Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., et al. (2013). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One* 8:e67374. doi: 10.1371/journal.pone.0067374

Ginsburg, N., and Nicholls, A. (1988). Perceived numerosity as a function of item size. *Percept. Mot. Skills* 67, 656–658.

Grasso, P. A., Anobile, G., and Arrighi, R. (2021a). Numerosity adaptation partly depends on the allocation of implicit numerosity-contingent visuo-spatial attention. *J. Vis.* 21:12. doi: 10.1167/jov.21.1.12

Grasso, P. A., Anobile, G., Caponi, C., and Arrighi, R. (2021b). Implicit visuospatial attention shapes numerosity adaptation and perception. *J. Vis.* 21:26. doi: 10.1167/jov.21.8.26

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., and Germine, L. (2012). Number sense across the lifespan as revealed by a massive internet-based sample. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11116–11120. doi: 10.1073/pnas.1200196109

Hurewitz, F., Gelman, R., and Schnitzer, B. (2006). Sometimes area counts more than number. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19599–19604.

Inglis, M., and Gilmore, C. (2013). Sampling from the mental number line: how are approximate number system representations formed? *Cognition* 129, 63–69. doi: 10.1016/j.cognition.2013.06.003

Itti, L., and Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* 40, 1489–1506. doi: 10.1016/S0042-6989(99)00163-7

Iuculano, T. (2016). Neurocognitive accounts of developmental dyscalculia and its remediation. *Prog. Brain Res.* 227, 305–333.

JASP Team (2020). *JASP (Version 0.12.2)[Computer software]*.

Kaufman, E. L., Lord, M. W., Reese, T. W., and Volkmann, J. (1949). The discrimination of visual number. *Am. J. Psychol.* 62, 498–525. doi: 10.2307/1418556

Knops, A., Piazza, M., Sengupta, R., Eger, E., and Melcher, D. (2014). A shared, flexible neural map architecture reflects capacity limits in both visual short-term memory and enumeration. *J. Neurosci.* 34, 9857–9866. doi: 10.1523/JNEUROSCI.2758-13.2014

Koch, C., and Ullman, S. (1987). "Shifts in selective visual attention: towards the underlying neural circuitry," in *Matters of Intelligence*, ed. L. M. Vaina (Dordrecht: Springer Netherlands), 115–141. doi: 10.1007/978-94-009-3833-5_5

Lasne, G., Piazza, M., Dehaene, S., Kleinschmidt, A., and Eger, E. (2018). Discriminability of numerosity-evoked fMRI activity patterns in human intra-parietal cortex reflects behavioral numerical acuity. *Cortex* 114, 90–101. doi: 10.1016/j.cortex.2018.03.008

Lee, K., and Cho, S. (2019). Visuo-spatial (but not verbal) executive working memory capacity modulates susceptibility to non-numerical visual magnitudes during numerosity comparison. *PLoS One* 14:e0214270. doi: 10.1371/journal.pone.0214270

Mammarella, I. C., Caviola, S., Giofrè, D., and Szûcs, D. (2018). The underlying structure of visuospatial working memory in children with mathematical learning disability. *Br. J. Dev. Psychol.* 36, 220–235. doi: 10.1111/bjdp.12202

Mandler, G., and Shebo, B. J. (1982). Subitizing: an analysis of its component processes. *J. Exp. Psychol. Gen.* 111, 1–22.

Matejko, A. A., and Ansari, D. (2021). Shared neural circuits for visuospatial working memory and arithmetic in children and adults. *J. Cogn. Neurosci.* 33, 1003–1019. doi: 10.1162/jocn_a_01695

Mazzocco, M. M. M., Feigenson, L., and Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (Dyscalculia): impaired numerical acuity contributes to MLD. *Child Dev.* 82, 1224–1237. doi: 10.1111/j.1467-8624.2011.01608.x

Mejias, S., Grégoire, J., and Noël, M.-P. (2012). Numerical estimation in adults with and without developmental dyscalculia. *Learn. Individ. Dif.* 22, 164–170. doi: 10.1016/j.lindif.2011.09.013

Melcher, D., and Piazza, M. (2011). The role of attentional priority and saliency in determining capacity limits in enumeration and visual working memory. *PLoS One* 6:e29296. doi: 10.1371/journal.pone.0029296

Menon, V. (2016). Working memory in children's math learning and its disruption in dyscalculia. *Curr. Opin. Behav. Sci.* 10, 125–132. doi: 10.1016/j.cobeha.2016.05.014

Nys, J., and Content, A. (2012). Judgement of discrete and continuous quantity in adults: number counts! *Q. J. Exp. Psychol.* 65, 675–690. doi: 10.1080/17470218.2011.619661

Piazza, M., De Feo, V., Panzeri, S., and Dehaene, S. (2018). Learning to focus on number. *Cognition* 181, 35–45. doi: 10.1016/j.cognition.2018.07.011

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition* 116, 33–41. doi: 10.1016/j.cognition.2010.03.012

Piazza, M., Fumarola, A., Chinello, A., and Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition* 121, 147–153. doi: 10.1016/j.cognition.2011.05.007

Pomè, A., Anobile, G., Cicchini, G. M., Scabia, A., and Burr, D. C. (2019). Higher attentional costs for numerosity estimation at high densities.

*Atten. Percept. Psychophys.* 81, 2604–2611. doi: 10.3758/s13414-019-01831-3

Price, G. R., Holloway, I., Rashal, P., Vesterinen, M., and Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Curr. Biol.* 17, R1042–R1043.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., and Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychol. Sci.* 19, 607–614. doi: 10.1111/j.1467-9280.2008.02130.x

Roggeman, C., Fias, W., and Verguts, T. (2010). Salience maps in parietal cortex: imaging and computational modeling. *Neuroimage* 52, 1005–1014. doi: 10.1016/j.neuroimage.2010.01.060

Rotzer, S., Loenneker, T., Kucian, K., Martin, E., Klaver, P., and von Aster, M. (2009). Dysfunctional neural network of spatial working memory contributes to developmental dyscalculia. *Neuropsychologia* 47, 2859–2865. doi: 10.1016/j.neuropsychologia.2009.06.009

Salti, M., Katzin, N., Katzin, D., Leibovich, T., and Henik, A. (2016). One tamed at a time: a new approach for controlling continuous magnitudes in numerical comparison tasks. *Behav. Res. Methods* 49, 1120–1127. doi: 10.3758/s13428-016-0772-7

Schütt, H. H., Harmeling, S., Macke, J. H., and Wichmann, F. A. (2016). Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vis. Res.* 122, 105–123. doi: 10.1016/j.visres.2016.02.002

Sengupta, R., Surampudi, B. R., and Melcher, D. (2014). A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. *Brain Res.* 1582, 114–124. doi: 10.1016/j.brainres.2014.03.014

Starr, A., DeWind, N. K., and Brannon, E. M. (2017). The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition* 168, 222–233. doi: 10.1016/j.cognition.2017.07.004

Szûcs, D. (2016). Subtypes and comorbidity in mathematical learning disabilities. *Prog. Brain Res.* 227, 277–304.

Szûcs, D., Nobes, A., Devine, A., Gabriel, F. C., and Gebuis, T. (2013b). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Front. Psychol.* 4:444. doi: 10.3389/fpsyg.2013.00444

Szûcs, D., Devine, A., Soltesz, F., Nobes, A., and Gabriel, F. (2013a). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex* 49, 2674–2688. doi: 10.1016/j.cortex.2013.06.007

Tibber, M. S., Greenwood, J. A., and Dakin, S. C. (2012). Number and density discrimination rely on a common metric: similar psychophysical effects of size, contrast, and divided attention. *J. Vis.* 12, 8–8. doi: 10.1167/12.6.8

Tokita, M., and Ishiguchi, A. (2010). How might the discrepancy in the effects of perceptual variables on numerosity judgment be reconciled? *Attent. Percept. Psychophys.* 72, 1839–1853. doi: 10.3758/APP.72.7.1839

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., van Kesteren, E., Derks, K., et al. (2019). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *PsyArXiv* [Preprint]. doi: 10.31234/osf.io/spreb

Vetter, P., Butterworth, B., and Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: evidence against a pre-attentive subitizing mechanism. *PLoS One* 3:e3269. doi: 10.1371/journal.pone.0003269

Zago, L., Petit, L., Turbelin, M. R., Andersson, F., Vigneau, M., and Tzourio-Mazoyer, N. (2008). How verbal and spatial manipulation networks contribute to calculation: an fMRI study. *Neuropsychologia* 46, 2403–2414.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

BRIEF RESEARCH REPORT
published: 23 November 2021
doi: 10.3389/fnhum.2021.762344

# Numerosity Comparison, Estimation and Proportion Estimation Abilities May Predict Numeracy and Cognitive Reflection in Adults

*Midori Tokita[1]* and Sumire Hirota[2]*

*[1] Faculty of Health Sciences, Mejiro University, Saitama, Japan, [2] Graduate School of Environmental and Information Studies, Tokyo City University, Yokohama, Japan*

This study explores whether and how different tasks associated with approximate number system (ANS) ability are related to numeracy and cognitive reflection in adults. We conducted an online experiment using a sample of 300 Japanese adults aged 20–39. Participants were given three ANS tasks (numerosity comparison, numerosity estimation, and proportion estimation) as well as Rasch-based numeracy scale and cognitive reflection test, and we tested the correlation among the measures of these tasks. We explored the hypothesis that the typical measures used to gauge ANS ability, numerosity comparison and numerosity estimation may mediate different cognitive mechanisms in adults. We also introduced a task measuring proportion estimation, added because such estimation requires numerosity perception and the ability to map symbolic numerals. Our findings suggest that there is a weak, but significant correlation among the three ANS-related tasks. Moreover, there is a significant relationship between each of these measures and the numeracy and CRT score, suggesting that the ANS-related ability may be associated with higher cognitive abilities such as numeracy and cognitive reflection. In addition, we found that performances on the numerosity and proportion estimation are more clearly related to CRT score than the numerosity comparison task.

**Keywords: numerosity comparison, numerosity estimation, proportion estimation, approximate number system, numeracy, cognitive reflection**

***Edited by:***
*Joonkoo Park,*
*University of Massachusetts Amherst,*
*United States*

***Reviewed by:***
*Giovanni Anobile,*
*University of Florence, Italy*
*José Manuel Reales,*
*National University of Distance*
*Education (UNED), Spain*

****Correspondence:***
*Midori Tokita*
*tokita@mejiro.ac.jp*

***Specialty section:***
*This article was submitted to*
*Cognitive Neuroscience,*
*a section of the journal*
*Frontiers in Human Neuroscience*

***Received:*** *21 August 2021*
***Accepted:*** *18 October 2021*
***Published:*** *23 November 2021*

***Citation:***
*Tokita M and Hirota S (2021)*
*Numerosity Comparison, Estimation*
*and Proportion Estimation Abilities*
*May Predict Numeracy and Cognitive*
*Reflection in Adults.*
*Front. Hum. Neurosci. 15:762344.*
*doi: 10.3389/fnhum.2021.762344*

## INTRODUCTION

The ability to extract the approximate numerical values of objects/events is crucial for surviving in the natural world as well as in modern society, which is full of numerical information in everyday life. Several studies on behavioral, neurophysiological, and brain imaging have demonstrated a dedicated mechanism for this numerical ability, denoted as the approximate number system (ANS), in humans as well as many other species (Nieder and Dehaene, 2009). A number of studies have shown that ANS plays a crucial role not only in the perception of numerosity, but also in understanding symbolic numerals, arithmetic, and mathematics (Malone et al., 2019; Prather, 2019; Sobkow et al., 2020).

Over the past decade, there has been great interest in the role ANS ability plays in higher cognitive processes, such as numerical ability, cognitive reflection, and decision-making. Some studies have demonstrated that high ANS ability is a predictor of good decision-making (Mueller and Brand, 2018; Mueller et al., 2018). The question of whether and how ANS ability relates to higher cognitive functions such as arithmetic, numerical computation, and decision-making is currently receiving much attention, but studies have not yielded consistent results, especially in adults. In developmental studies, it has been shown that ANS ability is closely related to early arithmetic skills, and that ANS ability predicts mathematical achievement in children (Halberda et al., 2008; Mazzocco et al., 2011; Hyde et al., 2014; Libertus et al., 2016; Wang et al., 2016; Cai et al., 2018; Libertus, 2019; Malone et al., 2019, 2021). In line with these arguments, several correlational studies have suggested that ANS ability may be foundational to the acquisition of formal math abilities (Mazzocco et al., 2011; Wang et al., 2016; Malone et al., 2019), and a deficit in the mechanism for representing and processing numerosity has proven to be one of the causes of low ability in symbolic numerical tasks (Butterworth, 2017). These studies suggest that both numerosity comparison and verbal numerical estimation, as well as ANS-number word mapping, may be important for acquisition of math abilities (Libertus et al., 2016). However, evidence from adult studies is not as clear cut as that found in developmental studies; many studies have pointed out that the relationship is inconsistently observed in adults (Yeo et al., 2019; Yeo and Price, 2021). For example, there is evidence that ANS acuity indirectly reflects only certain domains of math achievement in adults (Inglis et al., 2011; Patalano et al., 2015; Jang and Cho, 2016, 2018).Scholars have pointed out that a possible reason for the mixed results might be that different tasks have been used to measure both ANS ability and mathematics achievement (Lindskog et al., 2013; Prather, 2019). The ANS-related tasks used in each study for the purpose of assessing ANS ability have been inconsistent. Typically, there are three types of tasks used to assess ANS ability: numerosity comparison, numerosity estimation, and mental number-line mapping. For example, some studies used the numerosity comparison task in which the participants were presented with a pair of dot arrays, and asked to determine which array contained the larger number of dots (Mueller and Brand, 2018), while others used numerosity estimation tasks in which a number of elements were presented on a screen, and participants had to estimate the approximate numerical values (Yeo et al., 2019). Some studies argue that the acuity of symbolic-number mapping (a measure of approximate numeracy) is a robust predicator of numeracy and decision-making (Sobkow et al., 2019, 2020). To clarify the difference between each of the ANS-related tasks and their relation to arithmetic ability, Guillaume et al. (2016) compared two numerical tasks: numerical comparison and numerical estimation. Their results, which found no relation between the performance of these tasks, demonstrated that numerical comparison and estimation may mediate different cognitive mechanisms (Guillaume et al., 2016). They also tested the relationship between each numerosity task and arithmetic competence, and suggested that the performance of numerical comparison does not provide a pure measure of ANS ability. This evidence calls into question the relevance of correlating this measure with numerical ability, such as arithmetic competence, and underscores the importance of gaining a clear understanding of what each task assesses.

The purpose of the present study was to examine whether and how different ANS tasks relate to cognitive ability, which is highly related to numerical ability in adults in the general population. Specifically, we conducted an online experiment with adults, using three types of ANS-related tasks to assess ANS ability (i.e., numerosity comparison, numerosity estimation, and proportion estimation), and two types of cognitive ability tasks that might be related to ANS ability: numeracy and cognitive reflection tests (CRT). We tested the correlation among the tasks, and examined the relationship between the three types of ANS-related tasks. In addition to the two conventional ANS-related tasks, numerosity comparison and numerosity estimation, we introduced a proportion estimation task. Proportion estimation falls within the framework of perceived numerosity and probability judgments (Varey et al., 1990; Hollands and Dyre, 2000; Slusser and Barth, 2017). Although proportion estimation has not been used to assess ANS ability, we consider it to be an effective measure for assessing ANS ability as estimating proportion requires numerosity perception and the ability to map symbolic numerals. To assess numeracy, we used the Rasch-based numeracy scale (Weller et al., 2013). To assess cognitive reflection, we used CRT (Frederick, 2005; Toplak et al., 2014).

The concept of numeracy is typically defined as the ability to understand and process numerical information (Reyna et al., 2009). This includes computational skills such as multiplying, proportional reasoning, metacognitive monitoring, and understanding the gist of relative magnitude. Some research suggest that individual differences in numeracy may have important consequences for decision-making. CRT is also an extensively investigated measure of individual differences in rationality. This test was originally developed within the dual-process framework (Epstein et al., 1996; Evans and Stanovich, 2013), and captures whether people are able to inhibit their first incorrect response and follow it up with an intuitive and correct response. This score is also positively correlated with superior decision-making in a variety of decision tasks (Sinayev and Peters, 2015; Juanchich et al., 2016). As CRT items consist of mathematical tasks, it is suggested that the test largely captures not only cognitive reflection, but also other aspects related to numerical ability (Liberali et al., 2012; Campitelli and Gerrans, 2014; Patalano et al., 2015, 2020).

Based on the findings of previous studies, we made three predictions: First, no relation would be observed between the performance of numerosity comparison and numerosity estimation. Second, the measures of numerosity comparison and numerosity estimation would independently relate to the numeracy and CRT scores. Third, the proportion estimation measure would relate to the numeracy and CRT scores because both, abilities of numerosity comparison and proportion estimation, were required in performing the task.

# METHODS

## Participants

A total of 300 (150 female, 150 male) adults aged 20–39 years participated in the experiment through a Web inquiry company

(Cross Marketing Inc.). This age group was chosen as cognitive functions such as spatial visualization, reasoning, and memory and speed are reported to be considerably stable in this age group (Salthouse, 2010). All were native Japanese speakers and residents. Participants were required to use a laptop computer to be eligible to take part in the experiment. There was no regulation on the presentation time for each question and stimulus, and participants could take up the tasks at their own pace. Each numerosity task included these instructions: "There is no need to count dots one by one. Please answer based on your quick impression." There were no practice trials, and there was no feedback given on the correctness of the choices for any of the tasks.

## Materials and Procedure

All participants performed three ANS-related tasks (numerosity comparison, numerosity estimation, and proportion estimation), Rasch-based numeracy scale, and CRT. Each task is described in the following sections. The questions in numeracy and CRT task are listed in the **Supplementary Materials**.

### Rasch-Based Numeracy Scale

The Japanese version of the Rasch-based numeracy scale developed by Weller et al. (2013) was used (Hirota, 2019). The scale consists of eight questions on mathematical expressions and calculation of ratios, and two questions from Frederick's original CRT (Frederick, 2005). The scale has been used in a wide range of populations, and its usefulness and advantages have been tested (Weller et al., 2013; Peterson and Cheng, 2020). In the present study, participants were asked to answer the questions and record their answers using the numeric keypad on a computer. For each participant, we counted the number of correct answers, and computed the rate of correct response (number of correct responses out of eight) as the numeracy score.

### Cognitive Reflection Tests

Participants were asked to answer five questions composed of one from Frederick's original CRT and four from Toplak's additional CRT (Toplak et al., 2014). The Japanese version of these tests was used (Harada et al., 2018). The number of correct answers was used as the CRT score. Possible scores ranged from 0 to 5 points, with higher scores indicating higher cognitive reflection.

### Numerosity Comparison Task

As shown in **Figure 1A**, two sets of arrays, a standard array and a comparison array, were presented on the screen. The stimuli consisted of black dots on a light gray background. The diameter of the dots varied from array to array. The average dot size was controlled so that the total area of the dots was not a reliable cue for numerosity. Four standard numbers of stimuli (60, 90, 135, and 202) were used. These were within the range of numerical values presented in the numerosity estimation (40–451) and proportion estimation (15–302) tasks. The ratio of the comparison to the standard values was 0.85–0.9; thus, the sets of arrays were 60/51, 90/77, 135/115, 202/172, 60/54, 90/81, 135/122, and 202/182. These ratios were chosen to ensure the validity of the performance (Lindskog et al., 2013). The presentation order of the trials was randomized within a block. The participants were asked to indicate which array had more dots by clicking on the button below each array. At the beginning of the task, the participants were instructed to judge by the number of dots, and not by other properties of the arrays such as area and density. Each participant performed one trial for each stimuli pair, completing eight trials in total. The correct rate (CR) of each participant was calculated and used as the performance measure.

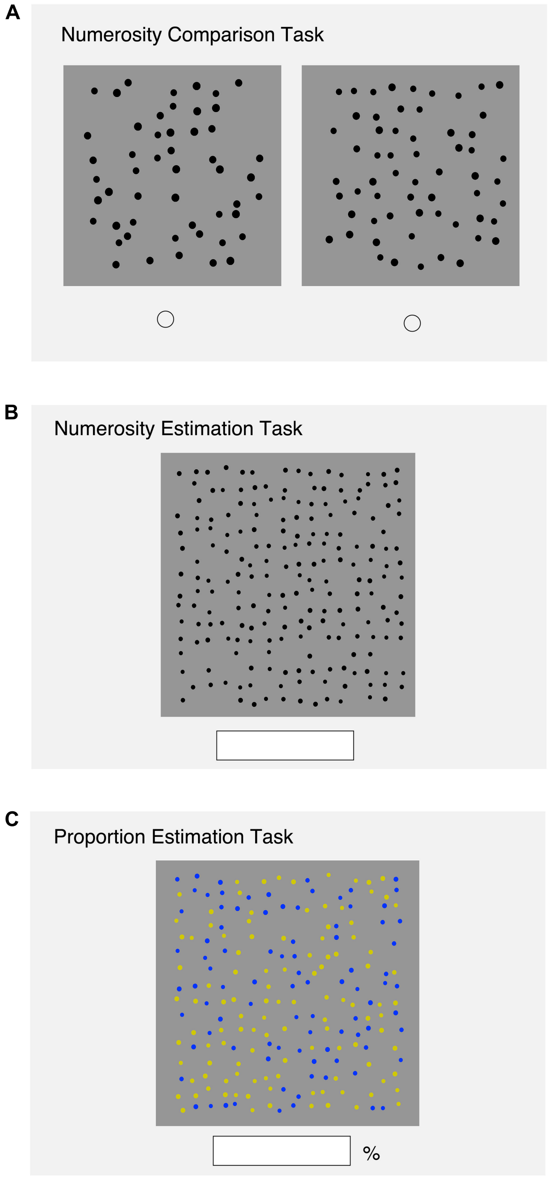


**FIGURE 1 | (A)** An illustration of the numerosity comparison task. **(B)** An illustration of the numerosity estimation task. **(C)** An illustration of the Proportion estimation task.

### Numerosity Estimation Task

As shown in **Figure 1B**, the participants saw a set of dot arrays presented on a gray background. Eight sets of dots, 27, 40, 60, 90, 135, 202, 302, and 451, which were logarithmically spaced, were presented in random order. The diameter of the dots varied within and between arrays, and the sizes of the invisible grid also varied so that the occupancy ratio of the dots to grid number ranged from 0.70 to 0.80. Neither the total area of the dots nor the spatial configuration could be a cue to numerosity. Each participant performed one trial in each set; thus, eight trials were performed in total. Participants were instructed to estimate the number of dots, and record their estimates using the numeric keypad as accurately as possible. We computed each participant's estimation accuracy by calculating the mean absolute error rate (AER) for each stimulus set, and used this value as the performance measure. The slope of the linear regression of the data points for each participant was calculated to assess the bias in numerosity estimation.

### Proportion Estimation Task

As shown in **Figure 1C**, a stimulus array was randomly presented at the center of the array. The stimuli consisted of blue and yellow dots on a light gray background. The diameter of the dots in the array varied, ranging from 0.9 to 1.1 times the average size, and the positions of stimuli also varied so that the total area of the dots and the spatial configuration would not be a possible cue to numerosity. There were 10 proportions for each set of dots relative to the total number of dots: approximately 0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, and 0.95. Two stimulus set sizes, 202 and 302 total dots, were applied. Specifically, the number of blue and yellow dots that appeared were 11/191, 32/170, 52/150, 72/130, 92/110, 112/90, 132/70, 152/50, 172/30, and 192/10 in set size 202, and 15/287, 45/257, 75/227, 105/197, 135/167, 166/136, 196/106, 226/76, 256/46, and 286/16 in set size 302. Thus, there were 20 conditions in total. Participants were instructed to estimate the percentage of blue dots relative to total dots, and record their estimates using the numeric keypad on a computer as accurately as possible. For each participant, we computed the estimation accuracy by calculating the mean AER in each stimulus set and used the value as the performance measure. The slope of the linear regression of data points for each participant was calculated to assess the bias in proportion estimation.

# RESULTS

## Criteria for Data Exclusion

Data were collected from 300 participants through a Web inquiry company. However, data of questionable reliability were removed according to the following criteria: we excluded untrustworthy responses, such as pressing the button on the same side for all trials in the numerosity comparison task, or entering the same number or a patterned number sequence in the numerosity, proportion estimation, numeracy, and CRT tasks. We also excluded responses that indicated the respondent did not understand the problem (for example, in the proportion estimation task, 191 blue dots out of a total of 202 dots is equivalent to approximately "5%"; however, some participants answered "95%"). Responses that appeared to be typing errors were also excluded, specifically responses that were greater than or equal to ten times higher than the correct answer or/and less than or equal to ten times smaller. We visually scrutinized participants' responses to assess for any outlying estimates that might have been missed by the trimming procedure described above. Data of 50 participants from all tasks were excluded because we performed within-subject correlation analysis. Then, we calculated the measures of each task; data three standard deviations above or below average were considered outliers and excluded from the analysis. As a result, only the responses of 238 participants were included in the analysis.

## Results of Each Task

### Rasch-Based Numeracy Scale and Cognitive Reflection Tests

The mean numeracy score of all participants was 56.0% (SD = 25.25). This result is consistent with Weller's original study, which showed 53.3% (SD = 29.5). The mean CRT task score of all participants was 2.63 (SD = 1.67; *n* = 238) out of 5, equivalent to 51.7% (SD = 33.34). This result is also consistent with previous studies (Harada et al., 2018). The value of Cronback's Alpha for numeracy and CRT was α = 0.75 and α = 0.70, respectively.

### Numerosity Comparison Task

The results of the numerosity comparison task are shown in **Figure 2A**. The mean correct rates for the ratios of 0.85 and 0.90 were 0.92 (SD = 0.16) and 0.77 (SD = 0.22), respectively. The mean total correct rate was 0.85 (SD = 0.16).

### Numerosity Estimation Task

**Figure 2B** shows the average estimated value and variance of each numerosity. The estimated value for each participant's response was considerably smaller than the actual value. The slope of the regression function for 190 out of the 238 participants was below 1, suggesting that the majority of participants underestimated the objective number of dots. In addition, the coefficient of variance ($CV = \frac{SD\ of\ estimation\ across\ participants}{mean\ of\ estimation\ across\ participants}$) for each numerosity across participants suggests that the variability of estimation increases with the numerosity value. These results are consistent with those of previous studies (Krueger, 1982; Guillaume et al., 2016).

### Proportion Estimation Task

**Figure 2C** shows the mean estimated proportion and variance of each proportion. No significant difference was observed in the estimation accuracy between the two set sizes (202 and 302). The slopes for the two sets were below 1.0 (186 out of 238 participants in set size 202, and 190 out of 238 participants in set size 302), demonstrating that the estimated proportions were overestimated when the proportion was smaller, and underestimated when the proportion was larger. These results are consistent with those of previous studies (Varey et al., 1990; Hollands and Dyre, 2000).

The value for Cronbach's Alpha for a numerosity comparison, numerosity estimation, and proportion estimation was α = 0.43, α = 0.64, and α = 0.58, respectively. It should be noted that the reliability of numerosity tasks were considerably low.

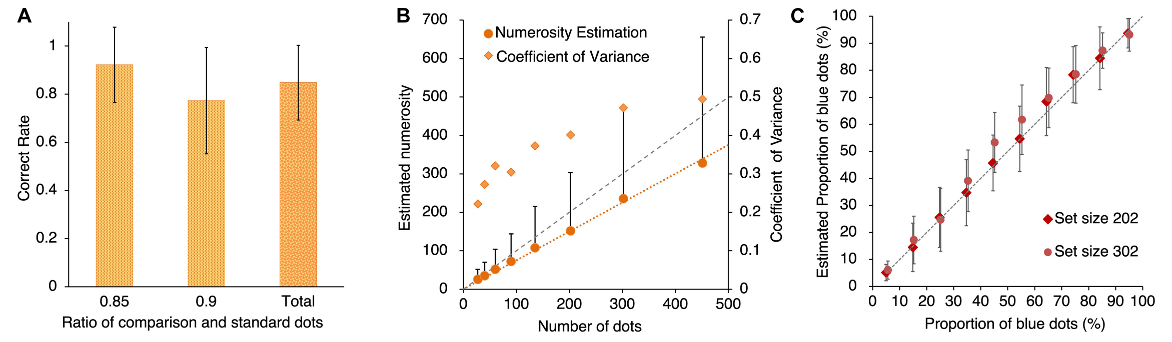


**FIGURE 2 | (A)** Mean correct rate of the numerosity comparison task. Error bars shows the standard deviation. **(B)** Mean estimated numerosity and coefficient of variance (CV) as a function of the number of actual dots in the numerosity estimation task. Error bars shows the standard deviation. **(C)** Mean of estimated proportion of blue dots as a function of actual proportion of blue dots. Error bars shows the standard deviation.

## Results of Correlation Analysis

The Pearson correlation coefficient matrix between the performance measures of the three numerosity tasks, CR from the numerosity comparison task, AERs from the numerosity estimation and proportion estimation tasks, numeracy, and CRT scores, was computed. The mean and standard deviations of all tasks and Pearson's correlation coefficients between all measures are shown in **Table 1**.

### Relationships Between Numerosity Measures

**Figures 3A–C** show how the numerosity measures were related to one another. The correlations between the performance of numerosity comparison and numerosity estimation reached a significant level ($r = -0.17$, $p < 0.01$), suggesting a relationship between numerosity comparison and estimation. The results were contrary to our prediction, which was based on previous studies (Guillaume et al., 2016; Prather, 2019). The correlations between the CR in numerosity comparison, AER in numerosity estimation, and AER in proportion estimation reached the significance level with AER in the numerosity estimation ($r = 0.30$, $p < 0.001$) and proportion estimation ($r = 0.20$, $p < 0.01$) tasks.

### Relationship Between ANS-Related Measures and Numeracy-Related Scores

As expected, the correlation between the numeracy scale and CRT was significant ($r = 0.71$, $p < 0.001$). The relationships between numeracy scores and CR from the numerosity comparison task and AERs from the numerosity estimation and proportion estimation tasks—as well as their respective relationships with the CRT scores—are shown in **Figures 3D–F**. First, a significant correlation between CR in numerosity comparison and numeracy score ($r = 0.24$, $p < 0.001$), between AER in numerosity estimation and numeracy score ($r = -0.25$, $p < 0.001$), and between AER in proportion estimation and numeracy score ($r = -0.33$, $p < 0.001$) was observed. Participants with higher numerosity comparison measures performed significantly better in the numeracy task than those with lower numerosity comparison measures. Participants with lower AERs in the numerosity estimation and proportion estimation tasks performed significantly better in the numeracy task than those with higher AERs.

A significant correlation between CR from numerosity comparison and CRT score ($r = 0.19$, $p < 0.001$), between AER in numerosity estimation and CRT ($r = -0.37$, $p < 0.01$), and between AER in proportion estimation and CRT ($r = -0.38$, $p < 0.001$) was observed. Participants with higher numerosity comparison measures performed significantly better in CRT relative to those with lower numerosity comparison measures: participants with lower AER in numerosity estimation and proportion estimation performed significantly better in CRT than those with higher AER. In addition, the results showed that the correlation coefficient between CRT score and AER in numerosity estimation [$p < 0.01$, $t(235) = 2.72$] and AER in proportion estimation [$p < 0.05$, $t(235) = 2.58$] was larger than that between CR from numerosity comparison and CRT.

To sum up, the results suggest that the three ANS-related tasks are associated with both numeracy and CRT scores. In particular, it has been suggested that the performance in numerosity and proportion estimation could be a significant predictor of CRT.

## DISCUSSION

The purpose of the present study was to examine whether and how abilities in different ANS-related tasks related to higher cognitive ability associated with numeracy and cognitive reflection in adults. In assessing ANS-related ability, numerosity comparison, numerosity estimation, and proportion estimation tasks were performed. To assess numeracy and cognitive reflection, the Rasch-based numeracy scale and CRT were administered, respectively.

There are three major findings of this study. First, the performance of three ANS-related tasks correlated with each other to suggest that the common numerical ability mediates to carry out these tasks. Contrary to our prediction, there was correlation between performance in the numerosity comparison and numerosity estimation tasks. The findings were in contrast with those of Guillaume et al. (2016) and Prather (2019), who found that performance in the comparison task did not

**TABLE 1 |** Bivariate correlations between numerosity discrimination, estimation, proportion estimation, numeracy and CRT.

| | *Mean* | *SD* | Numerosity comparison (CR) | Numerosity estimation (AER) | Proportion estimation (AER) | Numeracy score (CR) | CRT score |
|---|---|---|---|---|---|---|---|
| Numerosity Comparison (CR) | 0.85 | 0.16 | – | −0.17* | −0.20* | 0.24** | 0.19* |
| Numerosity Estimation (AER) | 0.36 | 0.14 | | – | 0.30** | −0.25** | −0.37** |
| Proportion Estimation (AER) | 7.89 | 2.28 | | | – | −0.33** | −0.38** |
| Numeracy score (CR) | 0.56 | 0.25 | | | | – | 0.71** |
| CRT score | 2.63 | 1.67 | | | | | – |

**p < 0.01 and **p < 0.001.*

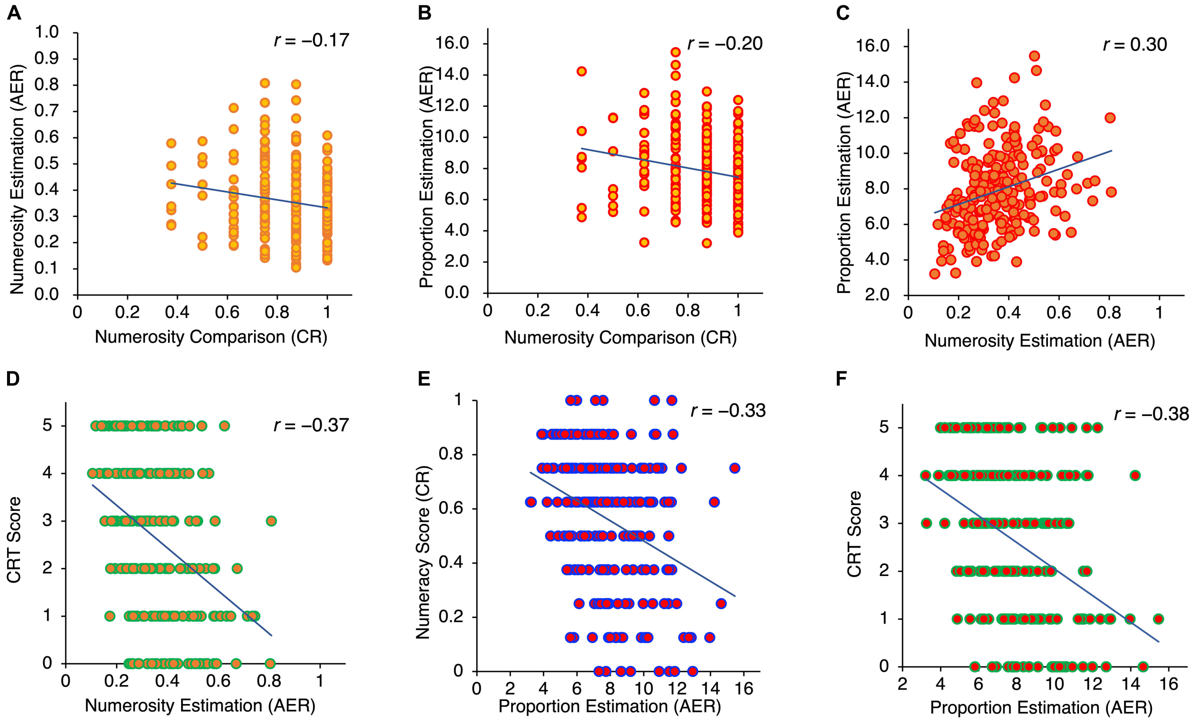


**FIGURE 3 |** Scatterplots depicting the relation between **(A)** numerosity comparison (CR) and numerosity estimation (AER), **(B)** numerosity comparison (CR) and proportion estimation AER), **(C)** numerosity estimation (AER) and proportion estimation (AER), **(D)** numerosity estimation (AER) and CRT Score, **(E)** proportion estimation (ASE) and numeracy score, and **(F)** proportion estimation (AER) and CRT score.

correlate with that in the estimation task, and claimed that numerosity comparison and estimation may mediate different cognitive mechanisms. The findings revealed that there were inconsistencies among the measures assessing ANS ability, as some previous studies had speculated. Why is the relationship between the numerosity comparison and numerosity estimation tasks inconsistent in adult studies? A possible explanation for the absence of a correlation between numerosity comparison and estimation in previous studies could be the sample size. Data from 71 participants were corrected in Guillaume et al. (2016) and data from 30 participants were corrected in Prather (2019). Why has a consistent correlation between numerosity comparison and estimation been observed in developmental studies? It could be attributed to the stimuli number of the estimation task: in development studies, the number of stimuli in numerosity estimation is considerably small, ranging from 5 to 20. In contrast, in adults, it is higher (10–400 or more). With high numerosity, the strategy and cognitive resources for mapping symbolic numerals to perceived numerosity may have large individual differences, as Yeo et al. (2019) claimed. Moreover, although the performance of numerosity comparison is highly associated with that of numerosity estimation in developmental studies, the relationship may change during the process of development.

Second, there was a significant relation between the performance in the numerosity comparison task and Rasch-based numeracy and CRT scores, as well as the performance in numerosity estimation and numeracy and CRT scores. The results support the claim that numerosity processing ability is significantly related to numeracy and cognitive reflection in adults, although the correlation is weak to moderate. The question arises as to how ANS ability relates to cognitive reflection. There are two possible reasons. First, the questions in the cognitive deliberation test used in this experiment included numerical processing skills, such as understanding proportions and calculations. Second, the skill of attending all the items in array, and grasping the approximate numerosity, may share common cognitive ability with cognitive reflection.

Third, it has been proved that the proportion estimation task introduced in this study relates to the performance on numerosity estimation, and the proportion estimation task is more clearly related to CRT score than the numerosity comparison task. The results suggest that proportion estimation could be an effective predicator of numeracy and CRT. As the proportion estimation task requires multiple ANS-related abilities, such as extraction of approximate numerosity and mapping of numerals to perceived proportion, its relation to numeracy and CRT is predicted as well as a single ANS ability.

This study has some limitations that should be considered in future research. First, the experiment was conducted on a large number of 20–39-year-old adults selected online from the general population, and each participant answered the question individually. Therefore, the size of stimuli, presentation time, each participant's response time, environment, and degree of commitment in performing the tasks were not controlled by the experimenter as this would have involved being in a laboratory. This might cause large individual differences within and between tasks, and raises questions regarding the reliability of the data. It has been suggested that people tend to use economical strategies and minimum cognitive loads in Web experiments, especially when there is no feedback or reward attached. As such, the participants were more likely to have performed the tasks using the least efforts. Second, the reliability indicated by Cronbach's alpha showed that the reliability of the three numerosity-related tasks was low and that there were differences among the tasks. As the correlation between unreliable items will result in lower values than the actual correlation coefficient, there is a possibility that the correlation between these values may be higher than the results of present studies. More reliable measures to test the numerosity ability for many general participants need to be invented and used for further research. Third, the measures used in proportion estimation and numerosity estimation could be elaborated further. Although previous studies applied these measures, the correct ratio in numerosity comparison and AER in the estimation task may represent a different aspect of numerical ability. To elaborate on the difference between the numerosity comparison and estimation processes, it is necessary to re-conduct the experiment in a laboratory under strictly controlled conditions, and with the appropriate measures for each performance. In addition, it should be noted that although the correlation analysis proves that a relationship exists between ANS-related tasks and numeracy and CRT, it does not prove that higher ANS ability contributes to higher numeracy and cognitive reflection; there is a possibility that numeracy and CRT may influence performance in ANS-related tasks. For example, an understanding of probability and proportion may affect performance in numerosity estimation and/or proportion estimation. The cause-effect relationship between ANS-related ability and numeracy should be examined more concretely. A further study on how each ANS-related ability relates to numeracy and cognitive reflection in adults, and how each ability develops and interacts with one another, should be examined.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, under reasonable request.

## ETHICS STATEMENT

Ethical approval was not provided for this study on human participants because the web survey company used in this study was approved by the participants in advance for ethical issues.

## AUTHOR CONTRIBUTIONS

MT and SH contributed to conception and design of the study. SH prepared Japanese version of the numeracy test. MT performed the statistical analysis and wrote the first draft of the manuscript. Both authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This work was partly supported by the Japan Society for the Promotion of Science (KAKENHI Grant Number 17K04501).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.762344/full#supplementary-material

## REFERENCES

Butterworth, B. (2017). The implications for education of an innate numerosity-processing mechanism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 373:20170118. doi: 10.1098/rstb.2017.0118

Cai, D., Zhang, L., Li, Y., Wei, W., and Georgiou, G. K. (2018). The role of approximate number system in different mathematics skills across grades. *Front. Psychol.* 9:1733. doi: 10.3389/fpsyg.2018.01733

Campitelli, G., and Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Mem. Cogn.* 42, 434–447. doi: 10.3758/s13421-013-0367-9

Epstein, S., Pacini, R., Denes-Raj, V., and Heier, H. (1996). Individual differences in intuitive-experiential and analytical-rational thinking styles. *J. Pers. Soc. Psychol.* 71, 390–405. doi: 10.1037//0022-3514.71.2.390

Evans, J. S. B. T., and Stanovich, K. E. (2013). Dual-Process theories of higher cognition:advancing the debate. *Perspect. Psychol. Sci.* 8, 223–241.

Frederick, S. (2005). Cognitive reflection and decision making. *J. Econ. Perspect.* 19, 25–42. doi: 10.1257/089533005775196732

Guillaume, M., Gevers, W., and Content, A. (2016). Assessing the approximate number system: no relation between numerical comparison and estimation tasks. *Psychol. Res.* 80, 248–258. doi: 10.1007/s00426-015-0657-x

Halberda, J., Mazzocco, M. M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–668. doi: 10.1038/nature07246

Harada, Y., Harada, E. T., and Suto, S. (2018). Examination of Cognitive Reflection Test (CRT): homogeneity of questions, order effects, and effects of instruction. *Tsukuba Psychol. Res.* 56, 27–34.

Hirota, S. (2019). Effects of respondents' objective numeracies and various kinds of alternative questions on response patterns in web-survey. *Philosophy(Mita tetsugakukai)* 142, 221–243. (in Japanese)

Hollands, J. G., and Dyre, B. P. (2000). Bias in proportion judgments: the cyclical power model. *Psychol. Rev.* 107, 500–524. doi: 10.1037/0033-295x.107.3.500

Hyde, D. C., Khanum, S., and Spelke, E. S. (2014). Brief non-symbolic, approximate number practice enhances subsequent exact symbolic arithmetic in children. *Cognition* 131, 92–107. doi: 10.1016/j.cognition.2013.12.007

Inglis, M., Attridge, N., Batchelor, S., and Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: but only in children. *Psychon. Bull. Rev.* 18, 1222–1229. doi: 10.3758/s13423-011-0154-1

Jang, S., and Cho, S. (2016). The acuity for numerosity (but not continuous magnitude) discrimination correlates with quantitative problem solving but not routinized arithmetic. *Curr. Psychol.* 35, 44–56. doi: 10.1007/s12144-015-9354-6

Jang, S., and Cho, S. (2018). The mediating role of number-to-magnitude mapping precision in the relationship between approximate number sense and math achievement depends on the domain of mathematics and age. *Learn. Individ. Dif.* 64, 113–124. doi: 10.1016/j.lindif.2018.05.005

Juanchich, M., Dewberry, C., Sirota, M., and Narendran, S. (2016). Cognitive reflection predicts real-life decision outcomes, but not over and above personality and decision-making styles. *J. Behav. Decis. Making* 29, 52–59. doi: 10.1002/bdm.1875

Krueger, L. E. (1982). Single judgments of numerosity. *Percept. Psychophys.* 31, 175–182.

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., and Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *J. Behav. Decis. Making* 25, 361–381. doi: 10.1002/bdm.752

Libertus, M. E. (2019). "Chapter 4 – Understanding the link between the approximate number system and math abilities," in *Cognitive Foundations for Improving Mathematical Learning*, Vol. 5, eds D. C. Geary, D. B. Berch, and K. Mann Koepke (Cambridge, MA: Academic Press), 91–106.

Libertus, M. E., Odic, D., Feigenson, L., and Halberda, J. (2016). The precision of mapping between number words and the approximate number system predicts children's formal math abilities. *J. Exp. Child Psychol.* 150, 207–226. doi: 10.1016/j.jecp.2016.06.003

Lindskog, M., Winman, A., Juslin, P., and Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: the evaluation of an adaptive test procedure. *Front. Psychol.* 4:510. doi: 10.3389/fpsyg.2013.00510

Malone, S. A., Pritchard, V. E., and Hulme, C. (2021). Separable effects of the approximate number system, symbolic number knowledge, and number ordering ability on early arithmetic development. *J. Exp. Child Psychol.* 208:105120. doi: 10.1016/j.jecp.2021.105120

Malone, S. A., Pritchard, V. E., Heron-Delaney, M., Burgoyne, K., Lervag, A., and Hulme, C. (2019). The relationship between numerosity discrimination and arithmetic skill reflects the approximate number system and cannot be explained by inhibitory control. *J. Exp. Child Psychol.* 184, 220–231. doi: 10.1016/j.jecp.2019.02.009

Mazzocco, M. l. M. M., Feigenson, L., and Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS One* 6:e23749. doi: 10.1371/journal.pone.0023749

Mueller, S. M., and Brand, M. (2018). Approximate number processing skills contribute to decision making under objective risk: interactions with executive functions and objective numeracy. *Front. Psychol.* 9:1202. doi: 10.3389/fpsyg.2018.01202

Mueller, S. M., Schiebener, J., Delazer, M., and Brand, M. (2018). Risk approximation in decision making: approximative numeric abilities predict advantageous decisions under objective risk. *Cogn. Process.* 19, 297–315. doi: 10.1007/s10339-018-0854-9

Nieder, A., and Dehaene, S. (2009). Representation of number in the brain. *Annu. Rev. Neurosci.* 32, 185–208. doi: 10.1146/annurev.neuro.051508.135550

Patalano, A. L., Saltiel, J. R., Machlin, L., and Barth, H. (2015). The role of numeracy and approximate number system acuity in predicting value and probability distortion. *Psychon. Bull. Rev.* 22, 1820–1829. doi: 10.3758/s13423-015-0849-9

Patalano, A. L., Zax, A., Williams, K., Mathias, L., Cordes, S., and Barth, H. (2020). Intuitive symbolic magnitude judgments and decision making under risk in adults. *Cogn. Psychol.* 118:101273. doi: 10.1016/j.cogpsych.2020.101273

Peterson, N., and Cheng, J. (2020). Decision experience in hyperchoice: the role of numeracy and age differences. *Curr. Psychol.* 1–13. doi: 10.1007/s12144-020-01041-3

Prather, R. (2019). Individual differences in numerical comparison is independent of numerical precision. *J. Numer. Cogn.* 5, 220–240. doi: 10.5964/jnc.v5i2.164

Reyna, V. F., Nelson, W. L., Han, P. K., and Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychol. Bull.* 135, 943–973. doi: 10.1037/a0017327

Salthouse, T. A. (2010). Selective review of cognitive aging. *J. Int. Neuropsychol. Soc.* 16, 754–760. doi: 10.1017/s1355617710000706

Sinayev, A., and Peters, E. (2015). Cognitive reflection vs. calculation in decision making. *Front. Psychol.* 6:532. doi: 10.3389/fpsyg.2015.00532

Slusser, E., and Barth, H. (2017). Intuitive proportion judgment in number-line estimation: converging evidence from multiple tasks. *J. Exp. Child Psychol.* 162, 181–198. doi: 10.1016/j.jecp.2017.04.010

Sobkow, A., Fulawka, K., Tomczak, P., Zjawiony, P., and Traczyk, J. (2019). Does mental number line training work? The effects of cognitive training on real-life mathematics, numeracy, and decision making. *J. Exp. Psychol. Appl.* 25, 372–385. doi: 10.1037/xap0000207

Sobkow, A., Olszewska, A., and Traczyk, J. (2020). Multiple numeric competencies predict decision outcomes beyond fluid intelligence and cognitive reflection. *Intelligence* 80:101452. doi: 10.1016/j.intell.2020.101452

Toplak, M. E., West, R. F., and Stanovich, K. E. (2014). Assessing miserly information processing: an expansion of the cognitive reflection test. *Think. Reason.* 20, 147–168. doi: 10.1080/13546783.2013.844729

Varey, C. A., Mellers, B. A., and Birnbaum, M. H. (1990). Judgments of proportions. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 613–625. doi: 10.1037/0096-1523.16.3.613

Wang, L., Sun, Y., and Zhou, X. (2016). Relation between approximate number system acuity and mathematical achievement: the influence of fluency. *Front. Psychol.* 7:1966. doi: 10.3389/fpsyg.2016.01966

Weller, J. A., Dieckmann, N. F., Tusler, M., Mertz, C. K., Burns, W. J., and Peters, E. (2013). Development and testing of an abbreviated numeracy scale: a rasch analysis approach. *J. Behav. Decis. Making* 26, 198–212. doi: 10.1002/bdm.1751

Yeo, D. J., and Price, G. R. (2021). Probing the mechanisms underlying numerosity-to-numeral mappings and their relation to math competence. *Psychol. Res.* 85, 1248–1271. doi: 10.1007/s00426-020-01299-z

Yeo, D. J., Wilkey, E. D., and Price, G. R. (2019). Malleability of mappings between Arabic numerals and approximate quantities: factors underlying individual differences and the relation to math. *Acta Psychol. (Amst.)* 198:102877. doi: 10.1016/j.actpsy.2019.102877

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 04 January 2022
doi: 10.3389/fnhum.2021.750582

# Electrophysiological Signatures of Numerosity Encoding in a Delayed Match-to-Sample Task

**Wanlu Fu[1]†, Serena Dolfi[2]†, Gisella Decarli[1], Chiara Spironelli[1,3] and Marco Zorzi[1,4]***

[1] Department of General Psychology, University of Padova, Padua, Italy, [2] Department of Developmental Psychology and Socialisation, University of Padova, Padua, Italy, [3] Padova Neuroscience Center (PNC), University of Padova, Padua, Italy, [4] IRCCS San Camillo Hospital, Venice, Italy

The number of elements in a small set of items is appraised in a fast and exact manner, a phenomenon called subitizing. In contrast, humans provide imprecise responses when comparing larger numerosities, with decreasing precision as the number of elements increases. Estimation is thought to rely on a dedicated system for the approximate representation of numerosity. While previous behavioral and neuroimaging studies associate subitizing to a domain-general system related to object tracking and identification, the nature of small numerosity processing is still debated. We investigated the neural processing of numerosity across subitizing and estimation ranges by examining electrophysiological activity during the memory retention period in a delayed numerical match-to-sample task. We also assessed potential differences in the neural signature of numerical magnitude in a fully non-symbolic or cross-format comparison. In line with behavioral performance, we observed modulation of parietal-occipital neural activity as a function of numerosity that differed in two ranges, with distinctive neural signatures of small numerosities showing clear similarities with those observed in visuospatial working memory tasks. We also found differences in neural activity related to numerical information in anticipation of single vs. cross-format comparison, suggesting a top-down modulation of numerical processing. Finally, behavioral results revealed enhanced performance in the mixed-format conditions and a significant correlation between task performance and symbolic mathematical skills. Overall, we provide evidence for distinct mechanisms related to small and large numerosity and differences in numerical encoding based on task demands.

**Keywords: approximate number system, subitizing, symbolic numbers, mathematics, numerosity encoding, event-related potentials**

***Edited by:***
*Michele Fornaciai,*
*International School for Advanced Studies (SISSA), Italy*

***Reviewed by:***
*Chiara Tagliabue,*
*University of Trento, Italy*
*Jacob M. Paul,*
*The University of Melbourne, Australia*

****Correspondence:***
*Marco Zorzi*
*marco.zorzi@unipd.it*

*†These authors have contributed equally to this work and share first authorship*

***Specialty section:***
*This article was submitted to Cognitive Neuroscience, a section of the journal Frontiers in Human Neuroscience*

***Received:*** *30 July 2021*
***Accepted:*** *29 November 2021*
***Published:*** *04 January 2022*

***Citation:***
*Fu W, Dolfi S, Decarli G, Spironelli C and Zorzi M (2022) Electrophysiological Signatures of Numerosity Encoding in a Delayed Match-to-Sample Task. Front. Hum. Neurosci. 15:750582. doi: 10.3389/fnhum.2021.750582*

## INTRODUCTION

Humans possess the ability to rapidly assess the number of items in a set (numerosity) without the necessity to count the objects. However, the speed and precision of these numerical judgments show different patterns depending on the number of elements to be estimated. Individuals provide fast and errorless responses in the case of sets composed of a few items (up to four), a phenomenon named subitizing (Kaufman and Lord, 1949). Instead, estimates of larger sets tend to be imprecise, with variability increasing proportionally to the number of objects (Jevons, 1871).

This dichotomy led to the hypothesis that subitizing and estimation are subserved, at least partially, by different mechanisms (see Piazza, 2010 for review). Numerosity estimation is thought to rely on the Approximate Number System (ANS; Feigenson et al., 2004), a preverbal mechanism characterized by the noisy encoding of numerical information (Feigenson et al., 2004). The ANS is often modeled as a mental number line where numerical magnitudes are coded as Gaussian distributions showing an increase in overlap as numerosity increases (either due to scalar variability or compressive scaling; Gallistel and Gelman, 1992; Dehaene, 2003), thereby accounting for the imprecise estimation of large sets and ratio-dependent performance in comparing different numerosities (in accordance with Weber's law). The same pattern of performance is also shown by state-of-the-art computational models of numerosity perception based on deep neural networks (Stoianov and Zorzi, 2012; Zorzi and Testolin, 2018; Nasr et al., 2019; Testolin et al., 2020). However, it has been proposed that numerosity mechanisms may work only at low to moderate densities, where items can be segregated. At higher densities, where objects become crowded, texture-like mechanisms may operate (Anobile et al., 2014, 2015). Moreover, performance in the subitizing range violates Weber's law, so that accuracy and reaction times in numerical comparison or same-different tasks remain stable across the entire range (Choo and Franconeri, 2014). Accordingly, the subitizing phenomenon has been related to a domain-general system for object identification and localization in space, named Object Tracking System (OTS) (Trick and Pylyshyn, 1994).

Nonetheless, the nature of subitizing remains contentious. An alternative view proposes that subitizing effects are a by-product of the scalar variability of the numerical representation, which would equally predict a high level of precision for smaller numerosities (Gallistel and Gelman, 1992). However, participants are faster and more precise in comparing pairs of numerosities in the subitizing range with respect to pairs in the estimation range with a similar ratio (Revkin et al., 2008). Nevertheless, two recent computational studies have revealed that distinct patterns of behavioral performance in subitizing and estimation range could potentially emerge from a single flexible system (Sengupta et al., 2017; Cheyette and Piantadosi, 2020), reinvigorating the debate.

In support of the idea of separate systems, different developmental trajectories have been described for subitizing and estimation abilities. The limited capacity of OTS improves during the first year of life, from a range of 1–2 items, up to the adult-like average limit of four objects (Coubart et al., 2014). In contrast, the precision of ANS seems to increase more steadily over the entire lifespan (Halberda et al., 2012). In addition, considerable individual differences exist both in the limit of objects that can be subitized and in the precision of large numerosity discrimination (Halberda et al., 2008; Piazza et al., 2011). However, individual subitizing limits seem not to be correlated with estimation precision (Piazza et al., 2011). Notably, converging evidence suggests a connection between individual differences in estimation precision (also known as number acuity) and more advanced mathematical skills, at least in developmental populations (Halberda et al., 2008; Mazzocco et al., 2011b; Starr et al., 2013; Chen and Li, 2014; van Marle et al., 2014). A widely accepted interpretation of this link is that the ANS may play a scaffolding role in the acquisition of symbolic numerical knowledge (Piazza et al., 2010; but see Leibovich and Ansari, 2016). In this view, during the acquisition of counting, individuals would create a mapping of symbolic numerals (Arabic digits or number words) onto the preexisting analog representations of numerical magnitudes (Gallistel and Gelman, 1992). Indeed, size and distance effects have also been reported in the case of comparison of symbolic numerals (e.g., Mussolin et al., 2010). It must, however, be noted that some authors have argued in favor of fully distinct processing for symbolic numbers, claiming that psychophysical similarities with ANS are limited to tasks where numerosities and numerals are interleaved, and observing a cognitive cost when symbolic and non-symbolic information needs to be integrated (Lyons et al., 2012; Sasanguie et al., 2017; Marinova et al., 2020). Indeed, although mathematical competence has been reliably related to both non-symbolic and symbolic magnitude processing (De Smedt et al., 2013), a stronger association with symbolic comparison abilities has been reported, especially in adults (Castronovo and Göbel, 2012; Schneider et al., 2017).

In contrast, individual variability in subitizing capacity has not been reliably associated with arithmetic skills (Anobile et al., 2019). Moreover, individuals with specific difficulties in mathematics (developmental dyscalculia) show impaired numerosity estimation (Piazza et al., 2010; Mazzocco et al., 2011a) but intact subitizing capacity (Decarli et al., 2020; but see Schleifer and Landerl, 2011). Conversely, impaired subitizing (but not estimation) has been observed in individuals with Down syndrome (Sella et al., 2013), who are also known to suffer from visuospatial working memory deficits. Accordingly, subitizing has been linked to domain-general visuospatial processing mechanisms. In particular, subitizing requires attentional resources and is disrupted by dual tasks with high attentional demands (Piazza et al., 2011), whereas numerosity estimation is carried out by a pre-attentive mechanism and is minimally affected by attentional load (Burr et al., 2010). Finally, individual subitizing limits have been related to visual working memory capacity (Piazza et al., 2011) and can be improved by cognitive training involving visuospatial abilities (Green and Bavelier, 2003).

Neuroimaging studies have consistently associated numerical processing with frontoparietal cortical circuits (Piazza et al., 2006). More specifically, fMRI activity in the bilateral intraparietal sulcus (IPS) elicited by deviant stimuli in a number adaptation paradigm has shown ratio dependency, a signature effect of ANS (Piazza et al., 2004). Moreover, this area has been related to magnitude processing during numerical judgments (e.g., Eger et al., 2015) and approximate computation (e.g., Bugden et al., 2019), although recent studies revealed that separate subregions near the IPS could be differentially engaged during different tasks (Castaldi et al., 2020). Moreover, in line with the idea of a partially shared semantic representation, neural responses to Arabic digits and number words were individuated in areas associated with magnitude processing, although additional lateralized circuits seem to be implicated in symbolic representation (Eger et al., 2003; Piazza and Eger, 2016; Sokolowski et al., 2017; but see Bulthé et al., 2014). Several studies report that bilateral IPS activation is sensitive to numerical magnitude changes in response to both symbolic and non-symbolic stimuli or cross-format presentation, with stronger effects on the left IPS in the case of symbolic stimuli (Piazza et al., 2007; Notebaert et al., 2011). Moreover, TMS studies showed that performance in non-symbolic comparison could be disrupted by bilateral parietal stimulation, while only left stimulation on similar sites was sufficient to impair digit discrimination (Andres et al., 2005; Cappelletti et al., 2007). In sum, these results relate the left hemisphere with the processing of exact numerical information and more refined coding of numerical magnitude, possibly because of connections with frontal circuits involved in language processing (Ansari, 2007). This idea is also consistent with a progressive left shift in the lateralization of number-related activity during development, which suggests an increasing differentiation of the symbolic representation from the magnitude system as formal mathematical concepts are learned (Emerson and Cantlon, 2015).

Only a few neuroimaging studies have examined both subitizing and estimation within the same experimental paradigm. Notably, in an fNIRS study, Cutini et al. (2014) found dissimilarities in the hemodynamic response of IPS to small and large numerosities, revealing a non-linear increase with numerosity in response amplitude. Also, a specific implication of the temporoparietal junction (TPJ) during small compared to large number discriminations supports the idea that a separate system could intervene in the processing of small numerosities (Ansari et al., 2007).

Importantly, previous electrophysiological studies based on the event-related potentials (ERPs) have produced mixed results. Although modulation of activity in posterior parietal sites has been reported both in response to large and small sets of objects, the use of different paradigms, procedures, and stimulus formats is the likely cause of discrepancies in timing and polarity of numerosity-related effects across studies. A positive component around 200 ms after stimulus onset has been found to increase in amplitude for small distances or ratios in symbolic and non-symbolic comparison tasks (Dehaene, 1996; Temple and Posner, 1998; Turconi et al., 2004) or passive viewing (Hyde and Spelke, 2009; Liu et al., 2018). While these effects are often interpreted as signatures of approximate magnitude processing, it must be noted that many of these studies mixed sets or numerals from the estimation and subitizing range. Moreover, modulation by numerical ratio with opposite polarity was found in a similar time window by Rubinsten et al. (2013). Using a non-symbolic match-to-sample task, instead, other authors found a distance effect in later negative deflections (300–500 ms), with a larger amplitude for close compared to far distances (Paulsen and Neville, 2008; Paulsen et al., 2010), while others failed to find distance effects even for early ERP components (van Hoogmoed and Kroesbergen, 2018).

In contrast, an earlier modulation (around 150 ms) has been found in response to small non-symbolic numbers. In this case, the amplitude is more reliably reported to increase as a function of the absolute magnitude for small, but not for large, numerosities (Libertus et al., 2007; Hyde and Spelke, 2009, 2012; Fornaciai and Park, 2017) or for mismatch compared to match conditions (Liu et al., 2018). However, some studies also report modulation by distance in this time window (Temple and Posner, 1998; Merkley et al., 2016). Moreover, Park et al. (2016) recently demonstrated that monotonic modulation of parietal activity by numerosity can also be appreciated for larger numerosities in positive peaks around 220 ms. Another line of research focusing on the link between small numerosity processing and attentional functions revealed that neural signatures of object individuation (e.g., N2pc) are modulated by target numerosity up to a fixed limit and correlated with individual subitizing span (Ester et al., 2012; Mazza et al., 2013).

In sum, previous ERP studies on number processing have examined magnitude effects during passive viewing or distance/ratio effects elicited by the comparison of pairs of stimuli, but the results are inconclusive regarding the distinction between small and large numerosities, and their putative relationship with different neurocognitive systems (OTS vs. ANS). In contrast, research on object individuation and/or visual working memory (WM) offers a potential alternative perspective for the investigation of numerosity encoding. More specifically, ERP signatures of memory retention (e.g., contralateral delay activity, CDA) during spatial or object working memory tasks have been reliably shown to be modulated by set size (McCollough et al., 2007; Drew et al., 2012). For example, during the retention period of a memory display in a change detection task, a negative slow wave in parietal sites increased in amplitude, as a function of the number of objects to be remembered, but only by up to 4–6 items (Feldmann-Wüstefeld, 2021). Recently, using an enumeration paradigm with a delay period between stimulus presentation and response prompt, Pagano et al. (2014) observed a similar modulation of CDA by the number of items to be enumerated, thereby showing involvement of working memory processes in subitizing during an explicitly numerical task. A similar approach has been previously used in numerical research in comparative studies investigating monkeys' neuronal response to numerosity during the delay period of a match-to-sample task. These studies revealed the presence of neurons maximally activated by specific numerosities in the prefrontal and posterior parietal cortices (Nieder and Miller, 2004; Nieder et al., 2006; Tudusciuc and Nieder, 2007; Viswanathan and Nieder, 2013). Taken together, these findings suggest that examining neural activity during encoding and maintenance of numerical information could help elucidate possible differential processing of small and large numerosities.

In this study, we exploited a delayed numerosity match-to-sample task (Sella et al., 2013) in which participants had to report if the numerosity of a dot array (sample stimulus) matched (or mismatched) the numerosity of a subsequent test stimulus presented after a 1-s blank-display delay period. More specifically, we hypothesized that EEG activity between sample stimulus offset and second test stimulus onset (i.e., the memory period) would be influenced by the number of items in the array, as typically observed in single-cell neurophysiological studies (e.g., Nieder and Miller, 2004). Importantly, we investigated neural activity during the delay period to avoid confounds from comparison processes and response-related activity. The memory-related activity is also less likely to be affected by changes in task-irrelevant perceptual attributes of the sample stimulus, such as individual dot size, cumulative area, etc., since variability linked to visual processing is usually found in early ERP components (e.g., P1, N270) and generally within 300 ms from stimulus onset (Luck, 2014; Soltész and Szucs, 2014; Park et al., 2016). Crucially, to investigate potential differences in the encoding of small and large numerosities, the number of items in the arrays spanned from subitizing to estimation range. Moreover, at an exploratory level, we investigated whether matching sample numerosity with an Arabic digit (i.e., a cross-format match-to-sample condition) would lead to a more distinctive neural signature of numerical magnitude. From behavioral evidence reporting differences in the comparison of within-format or multi-format numerical information (Marinova et al., 2020) and neuroscientific support in favor of format-dependent neural representation of numerosity (Eger et al., 2009), we hypothesized that the cross-format presentation, because of the symbolic nature of the test stimulus, could induce more precise encoding of the sample numerosity, irrespective of range. This "dots-to-digit" condition was presented as a separate block, but the sample stimulus remained identical to that of the "dots-to-dots" condition. Therefore, our focus remained on ERPs elicited by the non-symbolic sample stimulus to examine differences in physically equal stimuli varying only in psychological conditions (i.e., the format of the test stimulus), according to Hillyard's principle. We also asked if signature patterns of subitizing and estimation would be present at the behavioral level and whether the performance would differ in the two conditions. Finally, we offer our contribution to the ongoing debate on the relationship between basic numerical abilities and broad mathematical skills by looking at the correlation between behavioral performance in the match-to-sample task and more advanced arithmetic abilities assessed during the study.

# MATERIALS AND METHODS

## Participants

Thirty participants took part in the study, after giving written informed consent. Twelve of them received a monetary reward for participating. All the participants had a normal or corrected-to-normal vision. Data from two of the participants were discarded because of poor understanding of task instructions and low performance in the easiest task conditions (see below). The final sample, thus, consisted of twenty-eight participants (18 women, age range: 18–29 years). The sample size was set to be larger than the average of previous ERP studies (i.e., 20 participants) investigating symbolic or non-symbolic number comparison that reported reliable waveform modulations across numerical conditions (Libertus et al., 2007; Heine et al., 2013; Pagano et al., 2014; Fornaciai and Park, 2017). The research procedures were approved by the Psychological Science Ethics Committee of the University of Padova.

## Tasks and Stimuli

### Match-to-Sample Task

All the participants performed a computerized delayed numerosity match-to-sample task (Sella et al., 2013) divided into two blocks corresponding to different task conditions (see **Figure 1**). In the first block (dots-to-dots condition), each trial was composed of two sequentially presented images of dot arrays, and the participants had to report if the two images contained the same number of dots. In the second block (dots-to-digit condition), the second stimulus was an Arabic digit, and the participants were asked to report if the number of dots of the first stimulus matched the number indicated by the digit. Each trial started with a fixation cross appearing in the center of the screen for 400 ms, followed by a blank screen lasting for 150 ms. Then, the first stimulus (sample, always an array of dots) was displayed centrally for 300 ms, followed by another blank interval of 1 s (delay period). The second test stimulus was presented at the center of the screen for a maximum of 2 s, after which a blank screen was presented until response. Participants pressed the left key of the mouse to indicate a matching pair or the right key to indicate an unequal number of dots in the two images or a mismatching digit. The next trial started as soon as a response was provided. Each block consisted of 130 trials, with a short break after 40 and 80 trials. The participants performed six practice trials before the dots-to-dots block and four practice trials before the dots-to-digit block. All the practice trials were identical to the respective test block. The order of the two blocks was kept constant: the aim was to initially engage the participants in a fully non-symbolic condition before the cross-format condition, as the latter might trigger symbolic (e.g., verbal) coding of the sample numerosity in order to match it with the upcoming digit.

The numerosity of the stimuli ranged from 1 to 8. In mismatching trials, the test stimulus could differ from the sample by one less or one more, with equal probability. For example, a sample numerosity of 2 could be followed by a 2 (match) or a 1 or a 3 (mismatch). Each sample numerosity was presented 10 times in a matching pair and 10 times in a non-matching pair for each block, with the exception of numerosities 1 and 8, which were only compared with 2 and 7, respectively, always in a mismatching pair. Therefore, the six match pairings were 2 vs. 2, 3 vs. 3, 4 vs. 4, 5 vs. 5, 6 vs. 6, and 7 vs. 7, while the 7 mismatching pairings were 1 vs. 2, 2 vs. 3, 3 vs. 4, 4 vs. 5, 5 vs. 6, 6 vs. 7, and 7 vs. 8 (including the opposite combinations). On each trial, a pair was randomly selected from the entire list of possible pairings in order to minimize participants' expectations on the upcoming stimulus.

Dot arrays were created online during the experiment. The size and spatial arrangement of each dot in a grid were selected randomly on each trial. In the dots-to-dots condition, an opposite contrast polarity was used to present sample and test arrays, respectively, with white and black dots on a gray background. In the dots-to-digit condition, sample dot arrays were presented in white, while Arabic digits were presented in black as bold

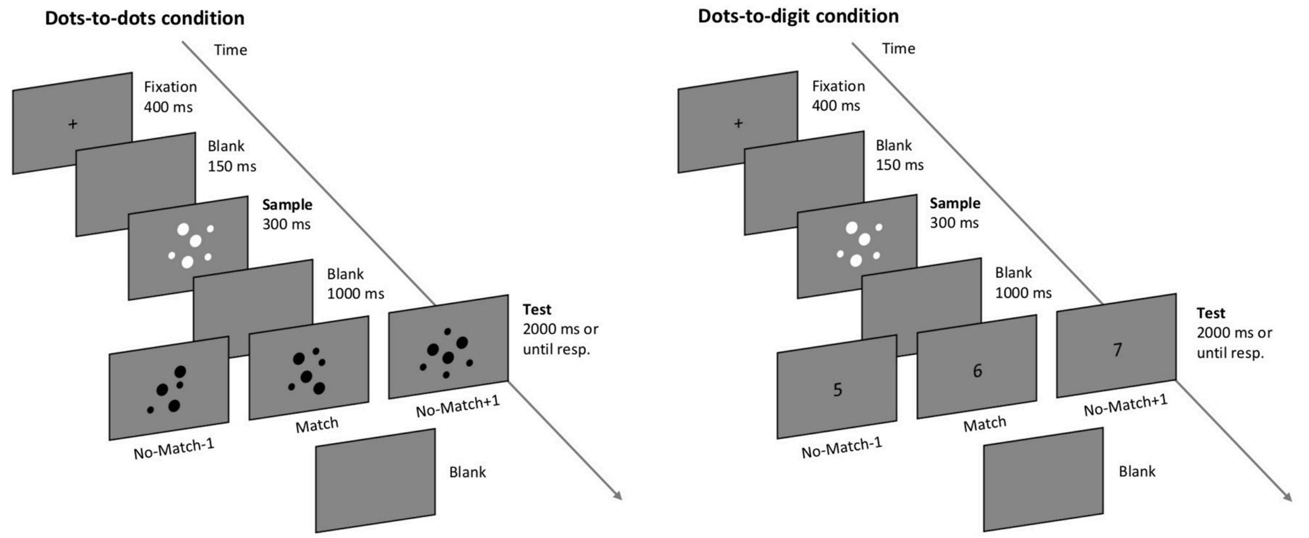


**FIGURE 1 |** Depiction of the match-to-sample paradigm. The participants were asked to decide if the numerosity of the second test stimulus matched the numerosity of a sample dot array. In the dots-to-dots condition, the test stimulus consisted of a dot-array, while in the dots-to-digit condition the test stimulus was an Arabic digit.

Courier New text, size 30. The individual area of the dots and contrast polarity were varied to minimize the influence of visual characteristics on participants' performance. The experimental task was presented with E-Prime 2 (Psychology Software Tools, 92 Pittsburgh, PA) on a 19′ monitor screen with a resolution of 1,024 × 768, running at 60 Hz.

### Calculation Skills

All the participants were also administered two calculation tests taken from a standardized battery for the assessment of literacy and numeracy skills in adults (LSC-SUA) (Montesano et al., 2020). This battery is used in Italy to evaluate learning disabilities in university students and adults. Participants performed two subtests: Mental Calculation and Approximate Calculation.

The Mental Calculation (MC) test consists of ten orally presented arithmetic operations (three additions, two subtractions, three multiplications, and two divisions) to be solved as rapidly and accurately as possible. In each computation, one of the numbers was a two- or three-digit number. The participants were asked to verbally report the result for each operation within 30 s. The experimenter recorded the number of correct responses and the time taken for each calculation.

The Approximate Calculation (AC) test consists of sixteen difficult arithmetic operations, presented to participants as written multiple-choice questions. The participants were asked to indicate for each operation the correct answer among three alternatives, trying to answer as many questions as they could within 1 min. The participants were explicitly instructed to avoid precise calculation to prioritize the speed of response.

## Procedure

The experiment was conducted in a quiet room, in a single session lasting approximately 1 h. Each participant first performed the paper-and-pencil test involving calculation skills. Then, after a brief resting period, they performed the numerical task while electroencephalography (EEG) was being recorded. The participants also completed a reading test at the beginning of the session and two additional computerized tasks tapping into phonological and visuospatial skills at the end of the session. These additional measures are not relevant to this study and will not be considered here.

## Electrophysiological Recording

An elastic cap (actiCAP; Brain Products, Gilching, Germany) with 64 pre-amplified electrodes mounted according to the International 10–20 system (Oostenveld and Praamstra, 2001) was used. Data were stored using the Brain Vision Recorder (Brain Products GmbH, Germany) system. The sampling rate was set at 1,000 Hz, and impedance was kept below 5 kΩ. All cortical electrodes were online referred to FCz electrodes.

## Data Analysis

### Analysis of Behavioral Performance

In the delayed match-to-sample task, we excluded outlier trials where the response was recorded before 200 ms (anticipation) or later than 2 s (maximum image display time). With this procedure, we discarded a total of 122 trials in the entire sample (∼1.6%). The analysis focused on test trials with sample or test numerosity from 1 to 7, excluding the maximum numerosity presented (8) to avoid guessing-end effects (Simon et al., 1998). We also inspected individual performance to ensure that the participants were correctly engaged in the task, which resulted in the exclusion of two participants. One was excluded because of the accuracy level in the easiest condition (1 vs. 2 or 2 vs. 1) that did not differ from chance according to a binomial test. The other was excluded because of an unusual pattern of responses, with lower accuracy in the small numerosity range than in larger numerosities. The data analysis focused on the proportion of correct responses and mean response times (RTs) in correct trials.

To analyze the effect of numerosity on performance, we first computed mean accuracy and RTs for each participant, condition, and sample numerosity. A preliminary ANOVA with numerosity and condition as within subject levels was performed to assess the overall effect of sample numerosity and its interaction with the condition. To investigate potential differences in performance in response to small and large numerosities, we first estimated the subitizing threshold at group level, separately for each condition, by fitting a piecewise linear model to describe group mean error rates as a function of sample numerosity (Pagano et al., 2014). The inflection point of the bilinear model in the dots-to-dots condition was 3.55 ($R^2_{\text{adj}} = 0.9$), while the breakpoint estimated on dots-to-digit trials was 5.33 ($R^2_{\text{adj}} = 0.97$). Note that similar estimates were derived when group subitizing thresholds were estimated with the method of Leibovich-Raveh et al. (2018), which returned thresholds of 4.07 and 5.18 for the dots-to-dots and dots-to-digit conditions, respectively. Based on the bilinear thresholds, mean accuracy and mean RTs were then separately computed for each participant and condition, across trials in small (pre-inflection) and large (post-inflection) numerical ranges. Differences between conditions and numerical ranges were then investigated by repeated measures ANOVAs.

To better investigate the pattern of responses to large and small numerosities, we also applied a bilinear fit on individual mean error rates as a function of numerosity, estimating individual breakpoints and pre-and post- inflection slopes separately for each participant and condition. During this procedure, one participant was excluded because of a lack of variability in their response. Mean model fit across conditions was $R^2_{\text{adj}} = 0.57$. Slopes across ranges and conditions were compared by means of repeated measure ANOVAs, while breakpoints in the two conditions were compared by a paired-sample *t*-test. A comparison between pre- and post-inflection slopes was also carried out on response times, fitting for each participant and condition a segmented model with a fixed breakpoint equal to the corresponding inflection point based on the error rate function. In all the analyses, Greenhouse-Geisser correction was applied in case of violation of the sphericity assumption, and *post hoc* tests were corrected for multiple comparisons with the Bonferroni method.

A correlation analysis was also performed to investigate the relationship between performance in the match-to-sample task and formal arithmetic skills. Standardized scores were computed for both subtests according to the Italian normative data of the LSC-SUA. Performance in the match-to-sample task was summarized for each condition using balanced integration scores (BISs), computed for each participant as the difference between the proportion of correct responses and mean correct response times, both standardized across conditions and participants (Liesefeld and Janczyk, 2019; Vandierendonck, 2021). Pearson correlations, controlled for FDR at alpha equal to 0.05, were then computed between BISs in the task and correctness scores from MC and AC, and the timing score from the MC test.

The data were analyzed with R (package SiZer), MATLAB (R2020a), and JASP (ver. 0.12.1 2020).

## EEG Analysis

After data collection, all cortical electrodes were re-referenced off-line to the mean activity of the whole scalp by the average reference procedure. Signal analysis was then carried out using the Brainstorm toolbox (Tadel et al., 2011). After a 0.1–45 Hz band-pass filter, eye movement artifact components (i.e., blinking, vertical, and horizontal movements) were corrected by applying Independent Component Analysis (ICA) transformation to the EEG signal. Since, as mentioned in the procedure, we analyzed the cortical activity elicited by sample stimuli, and numerosities 1 and 8 had less trial numbers than numerosities 2–7, the following ERP analysis excluded numerosities 1 and 8. Signals for numerosities 2–7 were then segmented into 1,050-ms epochs, ranging between 200 ms before the onset of sample stimuli and 850 ms after stimulus presentation. A baseline correction (−200 to 0 ms) was applied to all the epochs. After a second 0.5–30 Hz band-pass filtering, epochs with amplitude exceeding ± 75 μV were rejected. With this artifact rejection procedure, around 2.11 trials (10.55%) were rejected in each minimum experimental cell for each participant. Another baseline correction (−200 to 0 ms) was performed right before the grand average. Then, the grand-mean average was computed within the same condition across all the participants to compare ERP components among the different conditions. The mean number of trials for each numerosity in each condition was 17.89/20.

Previous studies have found both parietal and temporo-occipital sites related with small and large numerosity modulations (Libertus et al., 2007; Hyde and Spelke, 2009; Liu et al., 2018). A similar scalp distribution is also commonly reported for an activity related to the maintenance of visual information in working memory (Pinal et al., 2014; Feldmann-Wüstefeld, 2021), especially in delayed match-to-sample tasks (McCollough et al., 2007; Ikkai et al., 2010; Pagano et al., 2014). Based on previous studies and visual inspection of the electrophysiological scalp topography, we focused our analyses on parietal-occipital regions. In order to describe the relatively integral brain activation for ERP components, electrodes P7, P5, P3, PO7, and PO3 on the left hemisphere, and electrodes P8, P6, P4, PO8, and PO4 on the right hemisphere were considered separately as the left and right regions of interest (ROIs).

After inspecting the ERP waveform evoked by the sample stimulus onset, we focused on the time window between sample stimulus offset and test stimulus onset in order to investigate the representation of numerosity during the delay period, when no visual stimulus appeared on the computer screen. However, complementary results regarding sample stimulus encoding during the first 300 ms can be found in **Supplementary Figure 1**. Waveforms for the 300–850-ms time window were extracted from the original epochs. In addition, to avoid possible long-lasting effects due to online processing of the sample numerosity, a new baseline correction was computed using the last 100 ms before stimulus offset (from 200 to 300 ms after sample stimulus onset). Such baseline correction has been commonly performed in match-to-sample paradigms (Barriga-Paulino et al., 2014, 2015; Pelegrina et al., 2020) in order to reduce the impact of stimulus encoding and highlight information maintenance during the delay period. A 50-ms long negative peak component was found after the stimulus offset at around 250 ms, followed by a 50-ms positive component. Hereafter, we refer to these components as D-N250 and D-P300 (with the letter D highlighting that these components were observed during the delay period). Global field power waveforms across all the conditions are provided in **Supplementary Figure 2**. The mean amplitude and peak latency of the D-N250 and D-P300 components, as well as the mean amplitude for a later time window between 320 and 550 ms in the two ROIs, were exported from Brainstorm for a preliminary three-way (numerosity, condition, and Hemisphere) repeated-measures ANOVA. Based on the preliminary results, the analysis was then conducted separately for each hemisphere.

In parallel with the behavioral investigation, we also evaluated the inflection point in neural components by fitting the bilinear model on individual mean amplitude as a function of the sample numerosity, separately for each ERP time window and condition. The mean bilinear model fit across time windows, condition, and hemisphere was $R^2_{\text{adj}} = 0.29$. We then investigated potential differences between individual breaking points across the two conditions and compared the steepness of the slopes before and after the individual inflection point. Finally, we computed Spearman correlations to investigate a potential correspondence between behavioral and neural inflection point estimates.

# RESULTS

## Behavioral Results

### Performance in MTS Task

As a preliminary analysis of accuracy, we conducted a repeated measures ANOVA with condition (dots-to-dots and dots-to-digit) and sample numerosity (1–7) as within subject effects (see **Figure 2A**). This analysis revealed an overall higher accuracy in the dots-to-digit compared to dots-to-dots condition [$F(1, 27) = 101.84$, $p < 0.001$, $\eta^2_p = 0.79$], a significant effect of numerosity [$F(2.36, 63.65) = 61.17$, $p < 0.001$, $\eta^2_p = 0.69$], and a significant interaction [$F(3.17, 85.48) = 18$, $p < 0.001$, $\eta^2_p = 0.4$]. To better investigate performance in the two ranges, we then performed a repeated measures ANOVA with condition and range (based on group threshold) as within-subject effects, which confirmed the main effect of condition [$F(1, 27) = 27.28$, $p < 0.001$, $\eta^2_p = 0.5$] and revealed a significant effect of numerical range [$F(1, 27) = 88.68$, $p < 0.001$, $\eta^2_p = 0.77$], and a significant interaction between condition and range [$F(1, 27) = 14.82$, $p < 0.001$, $\eta^2_p = 0.35$]. The *post-hoc* tests showed an overall higher accuracy in small compared to large numerical range in both dots-to-dots [M(SD): $M_{\text{small}} = 0.97$ (0.02), $M_{\text{large}} = 0.79$ (0.09), $p < 0.001$, $d = 1.87$] and dots-to-digit [$M_{\text{small}} = 0.98$ (0.01), $M_{\text{large}} = 0.87$ (0.1), $p < 0.001$, $d = 1.05$] conditions. The difference between conditions, instead, emerged only in the large numerosity range ($p < 0.001$, $d = 1.2$).

We then considered individual thresholds, as the breakpoint individuated, by fitting a segmented model to each participant's mean error rates as a function of numerosity. The comparison of the inflection points between conditions [$t(26) = -3.41$, $p = 0.002$, $d = -0.66$] revealed an overall smaller breakpoint for dots-to-dots [$M = 3.96$ (1.38)] compared to dots-to-digit [$M = 5.1$ (1.04)]. We then compared slopes of the pre-inflection and post-inflection segments of the bilinear model for the two conditions by two-way repeated measures ANOVA. The comparison of slopes indicated a difference between ranges [$F(1, 26) = 39.59$, $p < 0.001$, $\eta^2_p = 0.6$], with a steeper slope for the post-inflection segment [$M = 0.12$ (0.12)] compared to the pre-inflection line [$M = -0.008$ (0.08)]. The effect of condition and the two-way interaction was not significant.

The analyses on reaction times showed a similar pattern of results (see **Figure 2B**). The preliminary ANOVA with condition and sample numerosity as within-subject effects revealed a faster response in the dots-to-digit compared to the dots-to-dots condition [$F(1, 27) = 114.8$, $p < 0.001$, $\eta^2_p = 0.81$]. The main effect of numerosity [$F(2.72, 73.46) = 102.44$, $p < 0.001$, $\eta^2_p = 0.79$] and the interaction between numerosity and condition [$F(1.74, 46.88) = 15.34$, $p < 0.001$, $\eta^2_p = 0.36$] were also significant. The ANOVA comparing the two ranges similarly showed a significant effect of condition [$F(1, 27) = 80.22$, $p < 0.001$, $\eta^2_p = 0.75$] and numerical range [$F(1, 27) = 174.61$, $p < 0.001$, $\eta^2_p = 0.87$] and a significant interaction between the two factors [$F(1, 27) = 10.26$, $p < 0.01$, $\eta^2_p = 0.27$]. The *post-hoc* tests revealed faster performance in the small range than in the large numerical range in both the dots-to-dots [$M_{\text{small}} = 629.17$ (130.3), $M_{\text{large}} = 872.03$ (160.1), $p < 0.001$, $d = -2.09$] and dots-to-digit [$M_{\text{small}} = 493.10$ (78.03), $M_{\text{large}} = 629.63$ (191.54), $p < 0.001$, $d = -1.17$] conditions. Moreover, the participants were overall faster in the dots-to-digit compared to the dots-to-dots condition, both in the small ($p < 0.001$, $d = -0.96$) and large ($p < 0.001$, $d = -1.7$) ranges.

We then fitted individually a segmented model on mean reaction times as a function of numerosity with a fixed breakpoint based on individual subitizing thresholds estimated from error rates. A comparison of the pre- and post- inflection slopes showed a significant effect of range [$F(1, 26) = 25.91$, $p < 0.001$, $\eta^2_p = 0.5$], with a steeper positive slope in the post-inflection range [$M = 96.82$ (101.31)] compared to the pre-inflection one [$M = 24.42$ (73.16)]. The effect of the condition and the two-way interaction was not significant.

### Correlations With LSC-SUA

We investigated the relationship between performance in the delayed match-to-sample task and arithmetic competence (see **Table 1**). A significant negative correlation was found between BIS in both the dots-to-dots and dots-to-digit conditions and MC timing. Notably, BIS in the dots-to-dots condition also showed a positive relationship with MC and AC scores.

## Electrophysiological Results

**Figure 3** represents the ERP waveforms evoked by sample numerosity onset (**Figure 3A**) and the averaged sub-epochs after sample stimulus offset (**Figure 3B**) on the left and right

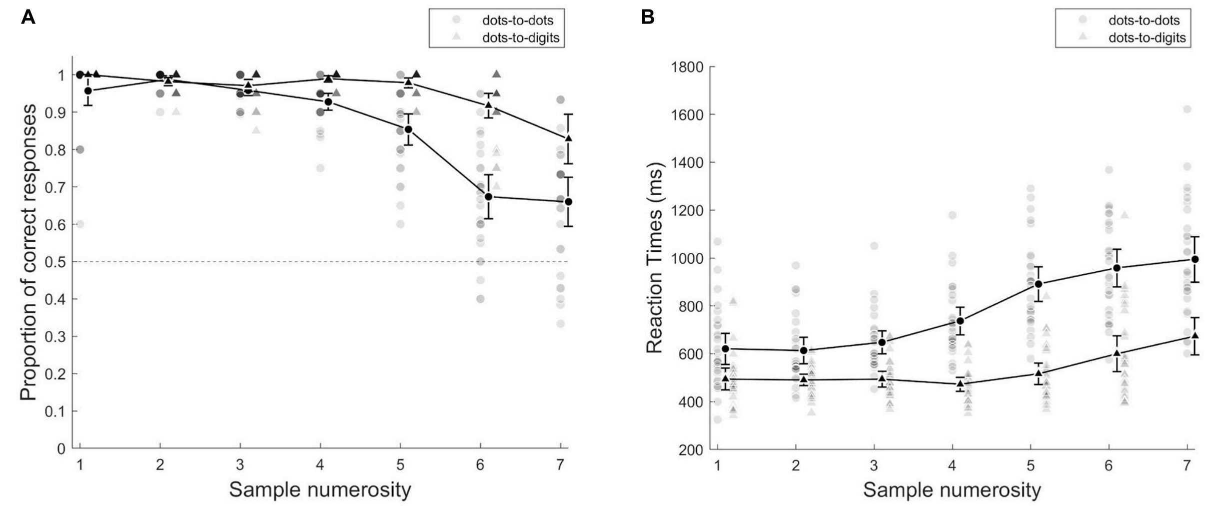


**FIGURE 2 | (A)** Proportion of correct responses and **(B)** mean reaction times as a function of sample numerosity. Individual points represent single participants, and group averages are presented with black lines, separately for each condition. Error bars indicate 95% CI.

ROIs. Analyses of the components evoked by sample onset (i.e., P1 and N2) are reported in **Supplementary Figure 1**. For the analysis of the delay period, mean amplitude and peak latency for the D-N250 and D-P300 components and mean amplitude for the later 320–550-ms time window was inserted in separate three-way repeated-measure ANOVAs, with sample numerosity (7 levels: 2–7), condition (2 levels: dots-to-dots and dots-to-digit), and hemisphere (2 levels: left and right ROIs) as within-subject variables.

As reported in **Table 2**, we found a significant effect of sample numerosity, condition, and hemisphere on mean amplitude for each of the time windows. The interaction between numerosity and condition was significant in the D-N250 and D-P300 components. Notably, the interaction between numerosity and hemisphere was significant in all the time windows; we, therefore, conducted separate repeated measures ANOVA for left and right ROIs. Since numerosity also interacted with the condition in all but one of the time windows, we retained the condition factor in these follow-up ANOVAs. We finally compared the slopes of mean amplitude as a function of numerosity in different ranges and conditions. The overall results of the latter analyses are shown in **Table 3** and **Figures 4**, **5** and are presented below separately by component.

The three-way repeated measures ANOVA on peak latency for the D-N250 and D-P300 components did not yield significant main effects or interactions. Therefore, the peak latency data were not further analyzed.

## D-N250 Component

For the D-N250 component, in the left ROIs, the ANOVA on mean amplitude revealed a significant interaction between condition and numerosity, and significant main effects (see **Table 3**). To better understand the effect of sample numerosity on the two conditions, a segmented model was then fitted on mean amplitude as a function of sample numerosity separately for the two conditions. No difference emerged between the inflection points in the dots-to-dots [M(SD) = 4.55 (1.06)] and dots-to-digit [M(SD) = 4.34 (1.07)] condition, and neither showed a significant correlation with behavioral breakpoints in the corresponding condition. **Figure 5** represents the slopes of pre- and post-inflection segments within each condition. A two-way repeated measure ANOVA revealed a significant difference between pre- and post- inflection segments [$F(1, 27) = 33.1$, $p < 0.001$, $\eta^2_p = 0.55$], with a steeper negative slope for the pre-inflection line [M(SD) = −1.46 (1.54)] compared to the post-inflection one [M(SD) = 0.36 (1.44)]. A main effect of condition [$F(1, 27) = 7.43$, $p = 0.01$, $\eta^2_p = 0.22$] was also found because of an overall more negative slope for dots-to-digit [$M = -0.87$ (1.68)] compared to dots-to-dots [$M = -0.23$ (1.30)] condition. The interaction between condition and range was, however, not significant.

In the right ROIs, the ANOVA on mean amplitude revealed a significant main effect of condition and numerosity, whereas the interaction between the two was not significant. A higher mean amplitude was found in the dots-to-digit [M(SD) = 0.11 (2.36)] compared to dots-to-dots condition [M(SD) = −0.83

**TABLE 1 |** Descriptive statistics of scores in the two calculation subtests from LSC-SUA and correlation with balanced integration scores (BIS) from match-to-sample task.

| | | | | Pearson r | |
|---|---|---|---|---|---|
| | | **M** | **SD** | **BIS dots-to-dots** | **BIS dots-to-digits** |
| LSC-SUA | MC scores (z) | 0.10 | 0.94 | 0.40* | 0.38 |
| | MC total time (z) | −0.42 | 0.91 | −0.51* | −0.46* |
| | AC scores (z) | 0.37 | 0.91 | 0.43* | 0.29 |

**p < 0.05 (FDR corrected with alpha = 0.05).*

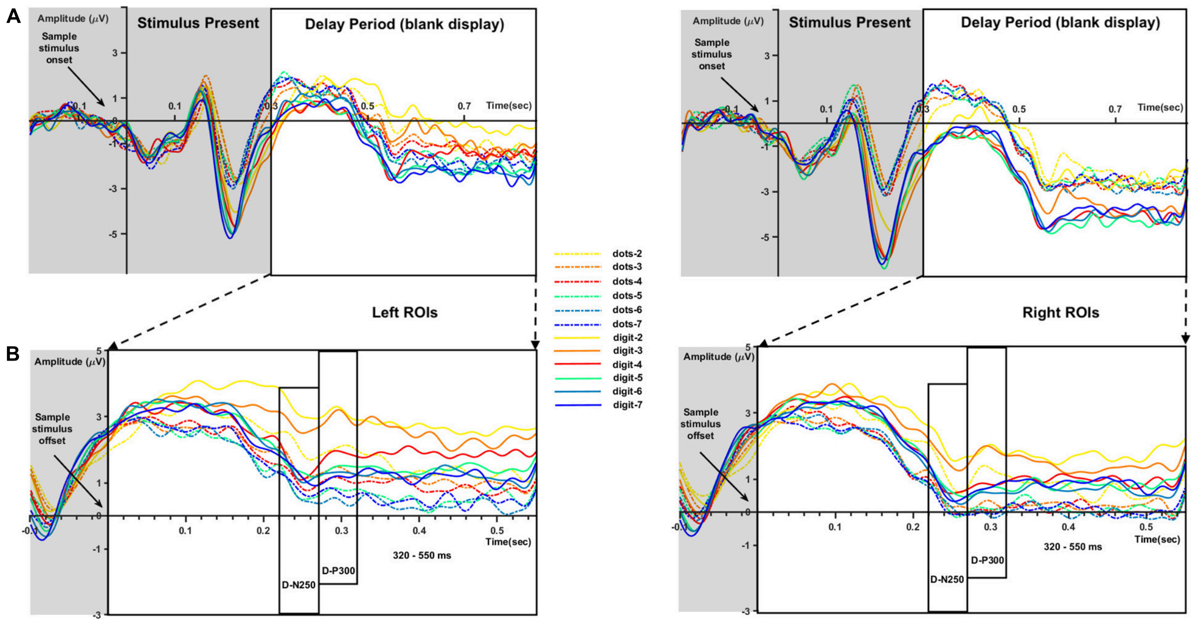


**FIGURE 3 |** ERP waveforms for each sample numerosity (different colors) and condition (dots-to-dots and dots-to-digit with full and dashed lines, respectively) in left and right regions of interest (ROIs). **(A)** Waveforms evoked by sample numerosity onset. **(B)** Extracted time windows corresponding to the delay period (blank display). We considered for analysis the negative deflections around 250 ms after sample numerosity offset (D-N250), the positive deflections around 300 ms (D-P300), and the final slow waveforms between 320 and 550 ms.

**TABLE 2 |** Mean amplitude three-way repeated measures ANOVA results for numerosity, condition, and hemisphere effects on each time window.

| | | D-N250 | | D-P300 | | 320–550 ms | |
|---|---|---|---|---|---|---|---|
| | *df* | *F* | $\eta_p^2$ | *F* | $\eta_p^2$ | *F* | $\eta_p^2$ |
| Numerosity | 5, 135 | 20.85*** | 0.44 | 19.04*** | 0.41 | 6.70*** | 0.20 |
| Condition | 1, 27 | 35.18*** | 0.57 | 48.44*** | 0.64 | 63.19*** | 0.70 |
| Hemisphere | 1, 27 | 13.87*** | 0.34 | 16.10*** | 0.37 | 7.19** | 0.21 |
| Numerosity × Condition | 5, 135 | 2.65* | 0.09 | 4.18** | 0.13 | 1.35 | 0.05 |
| Numerosity × Hemisphere | 5, 135 | 3.81** | 0.12 | 3.31** | 0.11 | 3.58** | 0.12 |
| Condition × Hemisphere | 1, 27 | 0.44 | 0.02 | 0.13 | 0.01 | 0.78 | 0.03 |
| Numerosity × Condition × Hemisphere | 5, 135 | 0.73 | 0.03 | 1.98 | 0.07 | 0.51 | 0.02 |

**p < 0.05, **p < 0.01, ***p < 0.001.*

(2.44)]. However, in order to investigate the relationship between neural and behavioral inflection points, in this case, bilinear models were also estimated at the individual level separately for each condition. Breakpoints from individual segmented models were not significantly different in the two conditions [M(SD): dots-to-dots = 4.58 (1.37), dots-to-digits = 4.11 (1.21)], and no significant correlation was found between neural and behavioral breakpoints. Slope analysis revealed a significant effect of range [$F(1, 27) = 7.35$, $p = 0.01$, $\eta^2_p = 0.21$], with a steeper amplitude slope in pre-inflection [$M = -1.08$ (2.35)] compared to post-inflection [$M = 0.06$ (1.52)] segments. The effect of the condition and the interaction were not significant.

## D-P300 Component

Considering the next positive peak component, the ANOVA on mean amplitude revealed significant two-way interactions and main effects of numerosity and condition on both left and right ROIs.

Breakpoints from the individual segmented model did not differ between the two conditions, both in the left [M(SD): dots-to-dots = 4.2 (1.16), dots-to-digit = 4.53 (1.16)], and in the right [M(SD): dots-to-dots = 4.27 (1.21), dots-to-digit = 4.39 (1.19)] ROIs, and no significant correlation was found between inflection points estimated from neural activity and behavioral performance.

Slope analysis for the left ROIs showed only a main effect of numerosity range [$F(1, 27) = 16.41$, $p < 0.001$, $\eta^2_p = 0.38$], with a steeper negative slope in the pre-inflection segment [$M = -1.4$ (2.19)] than in the post-inflection range [$M = 0.15$ (1.42)]. The effect of condition and the two-way interaction were not significant. Similarly, for the right ROIs, a main effect of range was found, with a steeper decreasing trend for pre-inflection

**TABLE 3 |** Mean amplitude two-way repeated measures ANOVA results for numerosity and condition effects on left and right regions of interest (ROIs) separately for each time window.

| | | **Left ROIs** | | **Right ROIs** | |
|---|---|---|---|---|---|
| | ***df*** | ***F*** | $\eta_p^2$ | ***F*** | $\eta_p^2$ |
| **D-N250** | | | | | |
| Numerosity | 5, 135 | 23.44*** | 0.28 | 10.60*** | 0.15 |
| Condition | 1, 27 | 34.50*** | 0.09 | 21.57*** | 0.09 |
| Numerosity × Condition | 5, 135 | 2.69* | 0.02 | 1.55 | 0.01 |
| **D-P300** | | | | | |
| Numerosity | 5, 135 | 18.46*** | 0.25 | 11.84*** | 0.16 |
| Condition | 1, 27 | 36.67*** | 0.12 | 36.69*** | 0.14 |
| Numerosity × Condition | 5, 135 | 3.79** | 0.02 | 3.38** | 0.03 |
| **320–550 ms** | | | | | |
| Numerosity | 5, 135 | 8.55*** | 0.14 | 3.53** | 0.06 |
| Condition | 1, 27 | 59.47*** | 0.17 | 41.85*** | 0.16 |
| Numerosity × Condition | 5, 135 | 1.67 | 0.01 | 0.69 | 0.01 |

*$^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.*

[$M = -1.28$ (2.64)] than for post-inflection [$M = 0.21$ (1.25)] slopes [$F(1, 27) = 18.07$, $p < 0.001$, $\eta^2_p = 0.4$]. Moreover, in this hemisphere, we found a significant effect of condition, with a steeper overall decreasing trend for dots-to-dots [$M = -0.95$ (2.18)] compared to dots-to-digit [$M = -0.13$ (1.71)] condition [$F(1, 27) = 5.5$, $p = 0.03$, $\eta^2_p = 0.17$]. No significant interaction between condition and range was found.

### The 320–550-ms Time Window

In this time window, the ANOVA on mean amplitude revealed in both hemispheres only significant main effects of numerosity and condition, with no two-way interactions. The dots-to-digit condition showed a greater mean amplitude than the dots-to-dots condition, both on the left [$M(SD)$: $M_{digit} = 1.15$ (2.5), $M_{dots} = -0.35$ (2.51)] and right [$M_{digit} = 0.17$ (2.69), $M_{dots} = -1.17$ (2.9)] ROIs.

Similar with the previous two components, results from the individual segmented models did not reveal a significant difference between the breaking point of the two conditions, both in the left [M(SD): dots-to-dots = 4.47 (0.8), dots-to-digit = 4.14 (1.26)] and in the right [M(SD): dots-to-dots = 4.16 (0.96), dots-to-digit = 4.29 (1.28)] ROIs. Moreover, no correlation emerged between the inflection points estimated from mean amplitude and behavioral breakpoints in the corresponding condition.

For the slope comparison, no significant effect was found in the left hemisphere. However, in the right hemisphere, a significant range effect was found [$F(1, 27) = 8.6$, $p = 0.007$, $\eta^2_p = 0.24$] with a steeper decreasing trend for pre-inflection [$M = -0.83$ (2.09)] compared to post-inflection [$M = 0.2$ (1)] segments. The effect of condition and the interaction were not significant.

## DISCUSSION

In this study, we investigated the encoding and maintenance of quantity information related to small and large numerosities in a delayed numerosity match-to-sample task. To this end, we analyzed electrophysiological activity between the presentation of a sample array of dots and a test numerical stimulus in non-symbolic or symbolic format. Importantly, the participants had to assess the number of objects in the sample stimulus because of the explicit comparison (same/different) requested after the test presentation (Sella et al., 2013). Through this manipulation, we aimed at individuating differences in the neural encoding of small and large numerosities, and to explore potential variations due to the use of cross-format matching. The numerical task allowed us to relate neural encoding and behavioral performance, as well as determine the functional relevance of our investigation, drawing a connection with real-world arithmetic abilities.

## Subitizing and Estimation

Task performance was modulated by numerical range. More specifically, both accuracy and reaction time results indicated better performance in the small range than in larger numerosities. In addition, the response pattern was different in the two ranges, as shown by the slope analysis. In the estimation range, accuracy showed a steady decrease, and reaction time showed a specular increase with numerosity. Conversely, compared to the estimation range, accuracy and reaction times remained relatively stable in the subitizing range. These results are in line with previous findings showing a differential pattern of response connected with small and large numerosities, compatible with the OTS vs. ANS distinction (Piazza, 2010; Choo and Franconeri, 2014). The estimated subitizing limit (between three and five items) fits well with existing literature on the capacity of the OTS in adults (Revkin et al., 2008; Burr et al., 2010), although some authors reported higher thresholds (Anobile et al., 2019). Moreover, the analysis on individual thresholds revealed a difference in the subitizing limit between the fully non-symbolic and mixed format conditions, in line with previous evidence showing that task settings can influence the subitizing span (e.g., Maldonado Moscoso et al., 2020). Despite our interpretation of the current findings as supporting the idea of a distinct mechanism underlying subitizing, we have previously mentioned that other accounts for this behavioral pattern have been proposed, postulating a single underlying mechanism (Sengupta et al., 2017; Cheyette and Piantadosi, 2020). However, our behavioral findings are complemented by electrophysiological results that clearly show modulation of parietal-occipital neural activity by numerical range.

The analyses of electrophysiological data focused on the time window between sample stimulus offset and test stimulus onset to investigate the representation of numerosity during the memory period before the comparison and response selection process. Note that the same logic has been widely used in neurophysiological studies investigating the coding of numerosity by single neurons during the delay period in a match-to-sample task (Nieder, 2005). In line with our initial hypotheses, in the memory retention interval after sample stimulus offset, we individuated two short time windows and one continuous slow wave sensitive to the number of dots in the array. Significant differences in amplitude between numerosities were found in a negative peak around 250 ms (D-N250) after

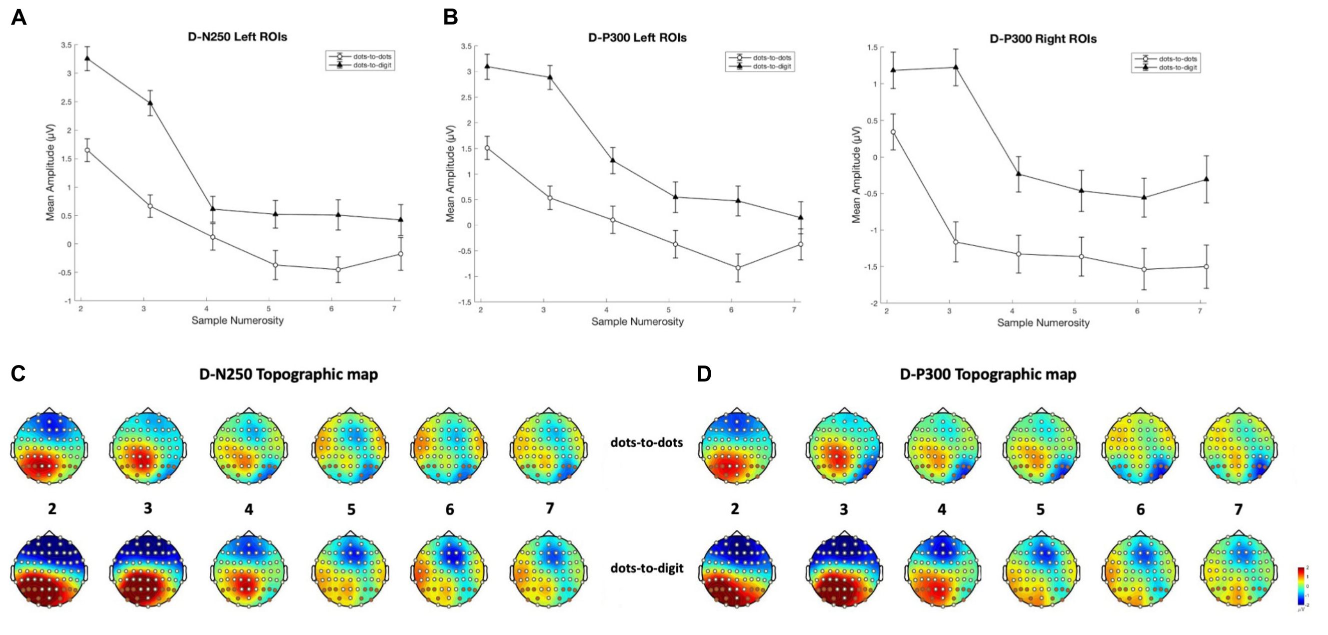


**FIGURE 4 |** Mean amplitude results and topographical maps. Top row: mean amplitude for each sample numerosity in the two separate conditions for components **(A)** D-N250 on left ROIs and **(B)** D-P300 on both left and right ROIs. Error bars represent the standard error of the mean. Bottom row: topography for each sample numerosity in the two conditions for components **(C)** D-N250 and **(D)** D-P300. The maps were obtained from 220 to 270 ms for D-N250 and 270 to 320 ms for D-P300.

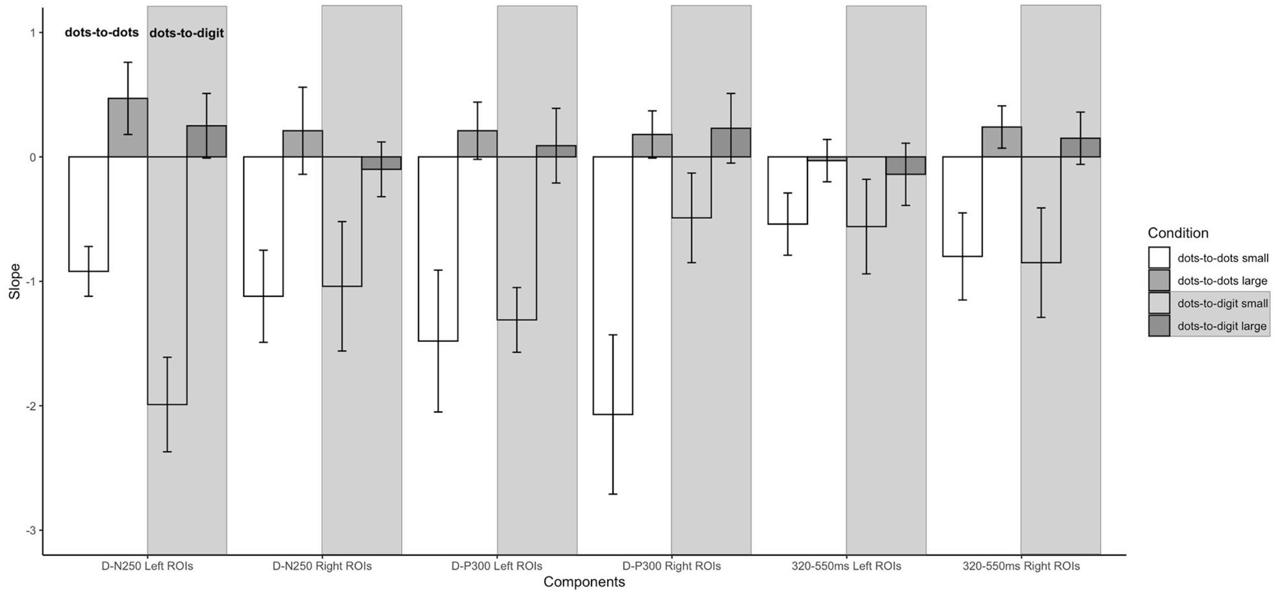


**FIGURE 5 |** Slopes of pre- and post-inflection segments within each condition for each ERP component. Error bars represent the standard error of the mean.

stimulus offset, a subsequent positive peak (D-P300), and a later slow wave between 320 and 500 ms. More specifically, similar to previous results (Hyde and Spelke, 2012), we found a modulation of activity by numerical magnitude for small sets of items. Indeed, our analyses revealed a trend of increasing negativity for progressively larger sets of items (see **Figure 4**) up to numerosity 4, visible in all the time windows except for the 320–550-ms time window. Instead, we failed in finding a clear modulation by numerical magnitude in the estimation range, where neural activity presented consistently shallower slopes in comparison to the subitizing range. Although the investigated timing deviates from previous studies in number research, this result is generally in line with evidence differentiating the neural response to small and large numerosities (Hyde and Spelke, 2009).

These results are also similar to previous working memory studies showing that the magnitude of negative slow waves in the retention period increases, as a function of the number of elements to hold in memory, up to a capacity limit (McCollough et al., 2007). A possible interpretation would then relate our result as reflecting working memory processes involved in maintaining the representation of individual items in the sample, during the delay interval. Different from working memory research, however, the current task did not require tracking location, color, or other combinations of perceptual characteristics of the individual items, especially in the mixed-format condition where dot arrays had to be compared with Arabic numerals. Set size modulation on neural signatures of working memory in the context of an enumeration task has been previously reported by Pagano et al. (2014). More specifically, the authors presented an increase in CDA amplitude with numerosity changes in the subitizing range. The set-size modulation of neural activity in our numerical match-to-sample task could then suggest a similar implication of working memory processing even in the absence of an explicit need to track the information of single items. More specifically, in our results, the differential pattern shown by activity in relation to small and large numerosities supports the idea that individuation and working memory capacity would be involved in small exact numerical processing, in line with the distinction of the subitizing phenomenon from the intrinsic nature of the numerical magnitude representation of the ANS (Feigenson et al., 2004). Indeed, strikingly similar results emerge from working memory tasks involving the encoding of location or color (Marois and Todd, 2004) and enumeration tasks of small sets of objects (Cutini et al., 2014): in both cases, an increase of neural activity in the posterior parietal cortex with set size, leveling at approximately four elements, has been reported. However, even if behavioral studies have consistently related working memory capacity with subitizing limits (Piazza et al., 2011), we did not find a correspondence between behavioral subitizing capacity and neural activity, which also showed consistent threshold estimates across format conditions. This result is in line with previous research that failed in finding an association between set size CDA modulation and behavioral subitizing span, whereas a correlation has been found with an earlier component (N2pc) linked to spatial attention and object individuation, suggesting that other domain-general mechanisms may also play an important role in small numerosity processing (Pagano et al., 2014).

Alternatively, since the delayed match-to-sample task did not specifically target components associated with previous working memory studies, we must also consider the possibility that the present modulation of neural activity could be more directly related to number encoding. Comparative studies using single-cell recording showed that during a numerosity match-to-sample task, some neurons in the prefrontal and posterior parietal cortices of monkeys activate maximally in the delay period, following the sensory presentation of a specific number of items, both in small and large numerical ranges (Nieder and Merten, 2007). Other studies have revealed single neurons in the lateral intraparietal sulcus activating with a monotonic modulation as a function of numerosity (Roitman et al., 2007). These results leave room for speculation that the current modulation could be associated with the encoding of numerical information, especially through summation coding, rather than domain-general mechanisms. However, such an interpretation is difficult to reconcile with the differential modulation in the small and large ranges, especially since it has been demonstrated that monotonic modulation of neural activity in response to large numerosity is detectable from early processing stages (Park et al., 2016). Moreover, even though a monotonic change in neurophysiological activity has been reported for both small and large numerosities, even in early processing stages, differences between the two ranges in polarity, latency, and anatomical location of response suggest functional dissociation (Fornaciai and Park, 2017). In particular, while in the large range numerical modulation has been associated with approximate numerical coding, in the small range, it is thought to reflect the amount of attention allocated to individual objects (Hyde and Spelke, 2009). Still, our failure to find a clear modulation by magnitude in the estimation range could be attributed to differences in paradigm and numerical range, since the previous studies used passive viewing and larger sets of items compared to the range used in this study. Further research is then needed to better differentiate between domain-general and domain-specific effects. We suggest that the match-to-sample task could be an optimal ground for a similar investigation, offering insights into the retention of numerical information in relation to the first quantity, as well as into the process of numerical discrimination between the sample and test numerosities. Indeed, a previous investigation using a match-to-sample paradigm has found a modulation of neural activity in response to larger numerosities, but as a variation of amplitude in response to the second test stimulus, depending on the numerical distance from the sample set size (Paulsen and Neville, 2008; Paulsen et al., 2010).

## Behavioral and Neural Response to the Format Change

Analysis of the behavioral results in the two different conditions revealed higher performance in the cross-format presentation of sample and test stimuli, compared to the fully non-symbolic condition. After a non-symbolic sample, in the estimation range, the participants were significantly faster and more accurate when the test stimulus was a digit rather than a dot array, and a similar difference in speed also emerged in the subitizing range. These results are in direct contrast with evidence reporting a cognitive cost for the integration of symbolic and non-symbolic numerical information. Lyons et al. (2012) found that responses in a comparison task were slower and less accurate when adult participants had to compare an array of dots with a digit than when they had to judge two non-symbolic stimuli or two numerals, interpreting such effect as evidence of a dissociation between representations of symbolic and non-symbolic numerosities. Instead, the current results point toward facilitation in mixed-format compared to fully non-symbolic condition, coherent with the idea of a shared representation in which numerals differ from non-symbolic magnitudes in terms of sharpness of tuning. In this view, the finer tuning elicited from numbers would decrease uncertainty in the comparison, resulting in better and faster performance. Similar results were also reported by Marinova et al. (2020), who found that a cross-format presentation in a comparison task leads to an intermediate performance level between fully non-symbolic and fully symbolic conditions, respectively, associated with lower and higher accuracies. However, since they failed to find a ratio effect elicited by numerals, the authors concluded that different cognitive systems were implicated in the processing of symbolic and non-symbolic stimuli. Unfortunately, the lack of a fully symbolic condition in this study did not allow for the testing of a similar effect. Further investigations could better explore performance differences across formats, using a larger numerosity range or several ratios between a sample and test stimulus, to avoid the risk of ceiling effects. In addition, we must notice that the present facilitation in the mixed format could have been partially influenced by a fixed order of the dots-to-dots and dots-to-digit conditions, presented always in the first and second blocks, respectively.

At the neural level, the two conditions exhibited an overall similar pattern of result: in both the dots-to-dots and dots-to-digit conditions, ERP amplitudes were modulated by numerosity in all the examined time windows. In addition, even when a significant interaction between numerosity and condition emerged, such as in left D-N250 and bilateral D-P300, the two conditions revealed similar modulation between subitizing and estimation ranges. Since in both conditions the sensory stimulation before the delay period consisted of an array of dots, the similarity might appear to be a trivial result. On the other hand, the significant modulation of numerosity in the mixed-format condition, which required the comparison with a symbolic digit, seems more likely to reflect the encoding of the number of objects compared to the previous interpretations relating the amplitude change to the memory rehearsal of individual items. However, a possible reconciliation between the two hypotheses is offered by Pagano et al. (2014), who proposed that the involvement of working memory in numerical tasks could be connected to the necessity to maintain an active representation of individual items during mapping with a specific numerical value.

We hypothesized that cross-format presentation could induce a more precise encoding of sample numerosity. In line with this hypothesis, a difference in amplitude between the two conditions emerged during the entire delay period, with responses to dots-to-digit trials eliciting an overall more positive activity compared to the dots-to-dots condition. Importantly, given the non-symbolic nature of the sample stimulus in both conditions, a difference in amplitude under numerical modulation between the two blocks could be attributed to the task-relevant information held in memory rather than potential discrepancies in sensory stimulation. This phenomenon is acknowledged in working memory studies, where differences in neurophysiological activity have been reported in response to identical stimuli, depending on features that participants were required to focus on (Woodman and Vogel, 2008). Moreover, we found a significant difference in slope present in the left hemisphere on D-N250, where the amplitude in the dots-to-digit condition showed a steeper decrease compared to fully non-symbolic trials. Instead, contrary to our hypothesis, we found the opposite effect in the right ROIs on the later D-P300 component, where the amplitude showed a stronger decrease in the dots-to-dots condition compared to the mixed format block. This result is consistent with a top-down modulation on the encoding of numerical magnitude due to specific task settings. A larger spacing of neural activity between adjacent numerosities could be the expression of an enhancement of the functional coding of the numerical information, emerging at different time points in left parietal sites in preparation of a comparison with an exact numeral and in the right hemisphere before non-symbolic targets. Converging evidence suggests that while bilateral regions near IPS have been associated with non-symbolic numerical processing, left-lateralized frontoparietal circuits could be related to the processing of exact numerical information, possibly related to the involvement of linguistic mechanisms in symbolic numerical processing (Ansari, 2007; Piazza et al., 2007). However, these results must be interpreted with caution, as it must be noted that evidence in favor of hemispheric specialization emerged only in analyses on individual slopes.

Despite the pre-delay baseline correction, we cannot exclude that the effects on the delay period might be influenced by the initial processing of the visual stimuli. In particular, the difference in amplitude between the two conditions could be ascribed to the effect of attentional processes involved in the initial processing phase and persisted in the memory period. This is consistent with a specular modulation of N2 (see **Figure 2** and **Supplementary Figure 1**). On one hand, the increased attentional engagement in the dots-to-digit condition might be related to task difficulty (Dong et al., 2015). However, while format condition has been shown to impact behavioral performance mainly in the estimation range, at the neural level, this effect was consistent for small and large numerosities. Moreover, behavioral results indicate better performance in the dots-to-digit compared to the dots-to-dots condition. On the other hand, attentional differences could also be more specifically attributed to the involvement of different individuation mechanisms in the two blocks, such as groupitizing strategies (Starkey and McCandliss, 2014), which have been shown to rely on attentional resources (Maldonado Moscoso et al., 2020). Indeed, the dots-to-digit condition, requiring higher enumeration precision, could encourage participants to cluster the elements of the arrays to facilitate enumeration. This interpretation is also consistent with our behavioral results, showing better performance and a higher subitizing threshold in the dots-to-digit condition (also see Maldonado Moscoso et al., 2020; Anobile et al., 2021). However, we must note that the sample stimulus was present on screen for only 300 ms, discouraging exact enumeration in both conditions and that the location of the dots in the array was randomly selected to avoid consistent grouping. We believe that future investigations specifically addressing this hypothesis could benefit from information regarding gaze movement during the initial processing phase (Schindler et al., 2020).

Finally, we must also acknowledge the possibility of confounding effects deriving from a blocked procedure. In particular, we cannot exclude that the overall difference in amplitude between the two conditions, both in the early time window and, as a sustained effect, the delay period, could be partially associated with reduced neural activity in the dots-to-dots condition connected with repetition suppression effects due to the uniform presentation of dot arrays (Summerfield and de Lange, 2014).

## Basic Numerical Abilities and Mathematical Skills

The correlational analysis revealed a significant link between performance in the match-to-sample task and math abilities. More specifically, performance in the dots-to-digit trials was related to the speed in the mental calculation test, and performance in both dots-to-dots was also related to the mental calculation scores. In addition, performance in the dots-to-dots condition correlated with approximate computation abilities. The significant relationship between speed of arithmetic computation and overall performance in the mixed-format condition is in line with previous evidence reporting a reliable link between math scores and symbolic comparison (Schneider et al., 2017). A similar connection with performance in fully non-symbolic presentation is particularly relevant for the ongoing debate on the connection between magnitude processing and broader mathematical skills. Recently, it has been proposed that non-symbolic numerical representation and symbolic math abilities would present a stronger link during the first steps of arithmetic knowledge acquisition, progressively differentiating because of increasing experience with formal mathematics (Ansari, 2008). This interpretation is supported by contradicting results on the relationship between mathematical skills and numerical acuity in non-symbolic comparison or estimation tasks (Castronovo and Göbel, 2012; Price et al., 2012; Szkudlarek et al., 2021). In contrast, the current result highlighted a relationship between magnitude processing and arithmetic computation in educated adults, suggesting a more continuous link during the life span. One possible explanation of the current result lies in the nature of the measures used. Previous studies have, revealed that non-symbolic processing could be differentially related to different aspects of math abilities (Lourenco et al., 2012). In particular, arithmetic computations that cannot be solved *via* rote memory, such as in the LSC tests administered in this investigation, are thought to rely more on magnitude processing and show a stronger relationship with precision in non-symbolic tasks (Piazza et al., 2010). Similar reasoning applies to the a match-to-sample paradigm as a measure of basic numerosity processing, which could partially diverge from acuity measures more commonly derived from comparison and estimation tasks. Notably, previous associations between performance in a similar computerized task and several tests of numerical competence have been reported in typically developing children and children with Down syndrome (Sella et al., 2013, 2021).

The nature of the relationship between numerosity processing and math skills is still a matter of debate since evidence of bidirectional influence between the precision of magnitude representation and math knowledge has been found (Elliott et al., 2019). One view proposes that magnitude processing might scaffold the acquisition of symbolic numerals, with an impact on broader symbolic math abilities (Libertus et al., 2013). In this view, even after formal learning, approximate number representation could provide support to basic operations, helping in the intuitive detection of errors (Feigenson et al., 2013). This interpretation is also supported by the current results, in particular by the significant correlation emerging between performance in non-symbolic trials and scores in the approximate computation test, where the participants had to quickly individuate a correct response by means of comparing the order of magnitudes of the multiple choices provided. However, even if our results are in support of a link between non-symbolic and symbolic representations and arithmetic abilities in adults, precise conclusions on the direction of this relationship cannot be drawn from the current correlational analysis.

# CONCLUSION

To conclude, our results provide new evidence for a functional dissociation between subitizing and estimation mechanisms. During an explicit numerical match-to-sample task, in addition to the behavioral signatures of OTS and ANS, we found that the two ranges were characterized by differential modulation of parietal-occipital neural activity by numerical information. Using ERPs, we demonstrated clear similarities with working memory neural signatures during the retention or encoding period of small numerical quantities, suggesting an implication of domain-general mechanism in small number processing. We also report evidence of top-down modulation of neural processing based on the demands of the task. Differences in ERP amplitude and numerical modulation suggest a qualitative variation in the encoding of numerical information in response to identical stimuli as a function of single vs. mixed-format comparison. The relevance of the current results is further emphasized by the significant relationship between performance in the delayed numerosity match-to-sample task and real-life mathematical skills, thereby supporting its validity for assessing basic number processing skills.

# DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

# ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Psychological Science Ethics Committee of the University of Padova. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

WF, GD, CS, and MZ contributed to the conception and design of the study. WF and GD performed the data collection. WF, SD, CS, and MZ contributed to the analysis and interpretation of data. WF and SD wrote the first draft of the manuscript. All the authors contributed to manuscript revision and approved the submitted version of the manuscript.

## FUNDING

This study was supported by a Cariparo Foundation Excellence grant (NUMSENSE) to MZ, and by a grant from MIUR (Dipartimenti di Eccellenza DM 11/05/2017 n.262) to the Department of General Psychology, University of Padova. WF was supported by a China Scholarship Council grant (CSC, 201908330302).

## ACKNOWLEDGMENTS

We thank Lorenzo Esposito for his assistance in data collection.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2021.750582/full#supplementary-material

## REFERENCES

Andres, M., Seron, X., and Olivier, E. (2005). Hemispheric lateralization of number comparison. *Cogn. Brain Res.* 25, 283–290. doi: 10.1016/j.cogbrainres.2005.06.002

Anobile, G., Arrighi, R., and Burr, D. C. (2019). Simultaneous and sequential subitizing are separate systems, and neither predicts math abilities. *J. Exp. Child. Psychol.* 178, 86–103. doi: 10.1016/j.jecp.2018.09.017

Anobile, G., Castaldi, E., Maldonado Moscoso, P. A., Arrighi, R., and Burr, D. (2021). Groupitizing improves estimation of numerosity of auditory sequences. *Front. Hum. Neurosci.* 15:687321. doi: 10.3389/fnhum.2021.687321

Anobile, G., Cicchini, G. M., and Burr, D. C. (2014). Separate mechanisms for perception of numerosity and density. *Psychol. Sci.* 25, 265–270. doi: 10.1177/0956797613501520

Anobile, G., Turi, M., Cicchini, G. M., and Burr, D. C. (2015). Mechanisms for perception of numerosity or texture-density are governed by crowding-like effects. *J. Vis.* 15, 1–12. doi: 10.1167/15.5.4

Ansari, D. (2007). Does the parietal cortex distinguish between "10," "ten," and ten dots? *Neuron* 53, 165–167. doi: 10.1016/j.neuron.2007.01.001

Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nat. Rev. Neurosci.* 9, 278–291. doi: 10.1038/nrn2334

Ansari, D., Lyons, I. M., Van Eimeren, L., and Xu, F. (2007). Linking visual attention and number processing in the brain: the role of the temporo-parietal junction in small and large symbolic and nonsymbolic number comparison. *J. Cogn. Neurosci.* 19, 1845–1853. doi: 10.1162/jocn.2007.19.11.1845

Barriga-Paulino, C. I., Rodríguez-Martínez, E. I., Rojas-Benjumea, M. Á, and Gómez, C. M. (2014). Slow wave maturation on a visual working memory task. *Brain Cogn.* 88, 43–54. doi: 10.1016/j.bandc.2014.04.003

Barriga-Paulino, C. I., Rojas Benjumea, M. Á, Rodríguez-Martínez, E. I., and Gómez González, C. M. (2015). Fronto-temporo-occipital activity changes with age during a visual working memory developmental study in children, adolescents and adults. *Neurosci. Lett.* 599, 26–31. doi: 10.1016/j.neulet.2015.05.017

Bugden, S., Woldorff, M. G., and Brannon, E. M. (2019). Shared and distinct neural circuitry for nonsymbolic and symbolic double-digit addition. *Hum. Brain Mapp.* 40, 1328–1343. doi: 10.1002/hbm.24452

Bulthé, J., De Smedt, B., and 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. doi: 10.1016/j.neuroimage.2013.10.049

Burr, D. C., Turi, M., and Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *J. Vis.* 10, 1–10. doi: 10.1167/10.6.20

Cappelletti, M., Barth, H., Fregni, F., Spelke, E. S., and Pascual-Leone, A. (2007). rTMS over the intraparietal sulcus disrupts numerosity processing. *Exp. Brain Res.* 179, 631–642. doi: 10.1007/s00221-006-0820-0

Castaldi, E., Vignaud, A., and Eger, E. (2020). NeuroImage mapping subcomponents of numerical cognition in relation to functional and anatomical landmarks of human parietal cortex. *Neuroimage* 221:117210. doi: 10.1016/j.neuroimage.2020.117210

Castronovo, J., and Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS One* 7:e33832. doi: 10.1371/journal.pone.0033832

Chen, Q., and Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: a meta-analysis. *Acta Psychol. (Amst)* 148, 163–172. doi: 10.1016/j.actpsy.2014.01.016

Cheyette, S. J., and Piantadosi, S. T. (2020). A unified account of numerosity perception. *Nat. Hum. Behav.* 4, 1265–1272. doi: 10.1038/s41562-020-00946-0

Choo, H., and Franconeri, S. L. (2014). Enumeration of small collections violates Weber's law. *Psychon. Bull. Rev.* 21, 93–99. doi: 10.3758/s13423-013-0474-4

Coubart, A., Izard, V., Spelke, E. S., Marie, J., and Streri, A. (2014). Dissociation between small and large numerosities in newborn infants. *Dev. Sci.* 17, 11–22. doi: 10.1111/desc.12108

Cutini, S., Scatturin, P., Basso Moro, S., and Zorzi, M. (2014). Are the neural correlates of subitizing and estimation dissociable? An fNIRS investigation. *Neuroimage* 85, 391–399. doi: 10.1016/j.neuroimage.2013.08.027

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

Decarli, G., Paris, E., Tencati, C., Nardelli, C., Vescovi, M., Surian, L., et al. (2020). Impaired large numerosity estimation and intact subitizing in developmental dyscalculia. *PLoS One* 15:e0244578. doi: 10.1371/journal.pone.0244578

Dehaene, S. (1996). The organization of brain activations in number comparison: event-related potentials and the additive-factors method. *J. Cogn. Neurosci.* 8, 47–68. doi: 10.1162/jocn.1996.8.1.47

Dehaene, S. (2003). The neural basis of the weber-fechner law: a logarithmic mental number line. *Trends Cogn. Sci.* 7, 145–147. doi: 10.1016/S1364-6613(03)00055-X

Dong, S., Reder, L. M., Yao, Y., Liu, Y., and Chen, F. (2015). Individual differences in working memory capacity are reflected in different ERP and EEG patterns to task difficulty. *Brain Res.* 1616, 146–156. doi: 10.1016/j.brainres.2015.05.003

Drew, T., Horowitz, T. S., Wolfe, J. M., and Vogel, E. K. (2012). Neural measures of dynamic changes in attentive tracking load. *J. Cogn. Neurosci.* 24, 440–450. doi: 10.1162/jocn_a_00107

Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., and Kleinschmidt, A. (2009). Deciphering cortical number coding from human brain activity patterns. *Curr. Biol.* 19, 1608–1615. doi: 10.1016/j.cub.2009.08.047

Eger, E., Pinel, P., Dehaene, S., and Kleinschmidt, A. (2015). Spatially invariant coding of numerical information in functionally defined subregions of human parietal cortex. *Cereb. Cortex* 25, 1319–1329. doi: 10.1093/cercor/bht323

Eger, E., Sterzer, P., Russ, M. O., Giraud, A. L., and Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron* 37, 719–726. doi: 10.1016/S0896-6273(03)00036-9

Elliott, L., Feigenson, L., Halberda, J., and Libertus, M. E. (2019). Bidirectional, longitudinal associations between math ability and approximate number system precision in childhood. *J. Cogn. Dev.* 20, 56–74. doi: 10.1080/15248372.2018.1551218

Emerson, R. W., and Cantlon, J. F. (2015). Continuity and change in children's longitudinal neural responses to numbers. *Dev. Sci.* 18, 314–326. doi: 10.1111/desc.12215

Ester, E. F., Drew, T., Klee, D., Vogel, E. K., and Awh, E. (2012). Neural measures reveal a fixed item limit in subitizing. *J. Neurosci.* 32, 7169–7177. doi: 10.1523/JNEUROSCI.1218-12.2012

Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002

Feigenson, L., Libertus, M. E., and Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Dev. Perspect.* 7, 74–79. doi: 10.1111/cdep.12019

Feldmann-Wüstefeld, T. (2021). Neural measures of working memory in a bilateral change detection task. *Psychophysiology* 58, 1–22. doi: 10.1111/psyp.13683

Fornaciai, M., and Park, J. (2017). Distinct neural signatures for very small and very large numerosities. *Front. Hum. Neurosci.* 11:21. doi: 10.3389/fnhum.2017.00021

Gallistel, C. R., and Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition* 44, 43–74. doi: 10.1016/0010-0277(92)90050-R

Green, C. S., and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature* 423, 534–537. doi: 10.1038/nature01647

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., and Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11116–11120. doi: 10.1073/pnas.1200196109

Halberda, J., Mazzocco, M. M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–668. doi: 10.1038/nature07246

Heine, A., Wißmann, J., Tamm, S., De Smedt, B., Schneider, M., Stern, E., et al. (2013). An electrophysiological investigation of non-symbolic magnitude processing: numerical distance effects in children with and without mathematical learning disabilities. *Cortex* 49, 2162–2177. doi: 10.1016/j.cortex.2012.11.009

Hyde, D. C., and Spelke, E. S. (2009). All numbers are not equal: an electrophysiological investigation of small and large number representations. *J. Cogn. Neurosci.* 21, 1039–1053. doi: 10.1162/jocn.2009.21090

Hyde, D. C., and Spelke, E. S. (2012). Spatiotemporal dynamics of processing nonsymbolic number: an event-related potential source localization study. *Hum. Brain Mapp.* 33, 2189–2203. doi: 10.1002/hbm.21352

Ikkai, A., McCollough, A. W., and Vogel, E. K. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *J. Neurophysiol.* 103, 1963–1968. doi: 10.1152/jn.00978.2009

Jevons, W. S. (1871). The power of numerical discrimination. *Nature* 3, 281–282. doi: 10.1038/003281a0

Kaufman, E. L., and Lord, M. W. (1949). The discrimination of visual number. *Am. J. Psychol.* 62, 498–525.

Leibovich, T., and Ansari, D. (2016). The symbol-grounding problem in numerical cognition: a review of theory, evidence, and outstanding questions. *Can. J. Exp. Psychol.* 70, 12–23. doi: 10.1037/cep0000070

Leibovich-Raveh, T., Lewis, D. J., Kadhim, S. A. R., and Ansari, D. (2018). A new method for calculating individual subitizing ranges. *JNC* 4, 429–447. doi: 10.5964/jnc.v4i2.74

Libertus, M. E., Feigenson, L., and Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learn. Individ. Differ.* 25, 126–133. doi: 10.1016/j.lindif.2013.02.001

Libertus, M. E., Woldorff, M. G., and Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behav. Brain Funct.* 3, 1–15. doi: 10.1186/1744-9081-3-1

Liesefeld, H. R., and Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behav. Res. Methods* 51, 40–60. doi: 10.3758/s13428-018-1076-x

Liu, R., Schunn, C. D., Fiez, J. A., and Libertus, M. E. (2018). The integration between nonsymbolic and symbolic numbers: evidence from an EEG study. *Brain Behav.* 8, 1–12. doi: 10.1002/brb3.938

Lourenco, S. F., Bonny, J. W., Fernandez, E. P., and Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18737–18742. doi: 10.1073/pnas.1207212109

Luck, S. J. (2014). *An Introduction To The Event-Related Potential Technique.* Cambridge, MA: MIT press.

Lyons, I. M., Ansari, D., and Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *J. Exp. Psychol. Gen.* 141, 635–641. doi: 10.1037/a0027248

Maldonado Moscoso, P. A., Castaldi, E., Burr, D. C., Arrighi, R., and Anobile, G. (2020). Grouping strategies in number estimation extend the subitizing range. *Sci. Rep.* 10, 1–10. doi: 10.1038/s41598-020-71871-5

Marinova, M., Sasanguie, D., and Reynvoet, B. (2020). Numerals do not need numerosities: robust evidence for distinct numerical representations for symbolic and non-symbolic numbers. *Psychol. Res.* 85, 764–776. doi: 10.1007/s00426-019-01286-z

Marois, R., and Todd, J. J. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428, 751–754.

Mazza, V., Pagano, S., and Caramazza, A. (2013). Multiple object individuation and exact enumeration. *J. Cogn. Neurosci.* 25, 697–705.

Mazzocco, M. M. M., Feigenson, L., and Halberda, J. (2011b). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS One* 6:e23749. doi: 10.1371/journal.pone.0023749

Mazzocco, M. M. M., Feigenson, L., and Halberda, J. (2011a). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Dev.* 82, 1224–1237. doi: 10.1111/j.1467-8624.2011.01608.x

McCollough, A. W., Machizawa, M. G., and Vogel, E. K. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex* 43, 77–94.

Merkley, R., Shimi, A., and Scerif, G. (2016). Electrophysiological markers of newly acquired symbolic numerical representations: the role of magnitude and ordinal information. *ZDM Math. Educ.* 48, 279–289. doi: 10.1007/s11858-015-0751-y

Montesano, L., Valenti, A., and Cornoldi, C. (2020). *LSC-SUA Prove Di Lettura, Comprensione Del Testo, Scrittura E Calcolo: Batteria Per La Valutazione Dei DSA E Altri Disturbi In Studenti Universitari E Adulti*, Vol. 1. Trento: Erickson, 1–181.

Mussolin, C., Mejias, S., and Noël, M. P. (2010). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition* 115, 10–25. doi: 10.1016/j.cognition.2009.10.006

Nasr, K., Viswanathan, P., and Nieder, A. (2019). Number detectors spontaneously emerge in a deep neural network designed for visual object recognition. *Sci. Adv.* 5, 1–11. doi: 10.1126/sciadv.aav7903

Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nat. Rev. Neurosci.* 6, 177–190. doi: 10.1038/nrn1626

Nieder, A., and Merten, K. (2007). A labeled-line code for small and large numerosities in the monkey prefrontal cortex. *J. Neurosci.* 27, 5986–5993. doi: 10.1523/JNEUROSCI.1056-07.2007

Nieder, A., and Miller, E. K. (2004). A parieto-frontal network for visual numerical information in the monkey. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7457–7462. doi: 10.1073/pnas.0402239101

Nieder, A., Diester, I., and Tudusciuc, O. (2006). Temporal and spatial enumeration processes in the primate parietal cortex. *Science* 313, 1431–1435. doi: 10.1126/science.1130308

Notebaert, K., Nelis, S., and Reynvoet, B. (2011). The magnitude representation of small and large symbolic numbers in the left and right hemisphere: an event-related fMRI study. *J. Cogn. Neurosci.* 23, 622–630. doi: 10.1162/jocn.2010.21445

Oostenveld, R., and Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clin. Neurophysiol.* 112, 713–719. doi: 10.1016/S1388-2457(00)00527-7

Pagano, S., Lombardi, L., and Mazza, V. (2014). Brain dynamics of attention and working memory engagement in subitizing. *Brain Res.* 1543, 244–252. doi: 10.1016/j.brainres.2013.11.025

Park, J., Dewind, N. K., Woldorff, M. G., and Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cereb. Cortex* 26, 748–763. doi: 10.1093/cercor/bhv017

Paulsen, D. J., and Neville, H. J. (2008). The processing of non-symbolic numerical magnitudes as indexed by ERPs. *Neuropsychologia* 46, 2532–2544. doi: 10.1016/j.neuropsychologia.2008.04.003

Paulsen, D. J., Woldorff, M. G., and Brannon, E. M. (2010). Individual differences in nonverbal number discrimination correlate with event-related potentials and measures of probabilistic reasoning. *Neuropsychologia* 48, 3687–3695. doi: 10.1016/j.neuropsychologia.2010.08.014

Pelegrina, S., Molina, R., Rodríguez-Martínez, E. I., Linares, R., and Gómez, C. M. (2020). Age-related changes in selection, recognition, updating and maintenance information in WM. An ERP study in children and adolescents. *Biol. Psychol.* 157:107977. doi: 10.1016/j.biopsycho.2020.107977

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends Cogn. Sci.* 14, 542–551. doi: 10.1016/j.tics.2010.09.008

Piazza, M., and Eger, E. (2016). Neural foundations and functional specificity of number representations. *Neuropsychologia* 83, 257–273. doi: 10.1016/j.neuropsychologia.2015.09.025

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition* 116, 33–41. doi: 10.1016/j.cognition.2010.03.012

Piazza, M., Fumarola, A., Chinello, A., and Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition* 121, 147–153. doi: 10.1016/j.cognition.2011.05.007

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., and Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron* 44, 547–555. doi: 10.1016/j.neuron.2004.10.014

Piazza, M., Mechelli, A., Price, C. J., and Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: an fMRI study. *Brain Res.* 1106, 177–188. doi: 10.1016/j.brainres.2006.05.104

Piazza, M., Pinel, P., Le Bihan, D., and Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron* 53, 293–305. doi: 10.1016/j.neuron.2006.11.022

Pinal, D., Zurrón, M., and Díaz, F. (2014). Effects of load and maintenance duration on the time course of information encoding and retrieval in working memory: from perceptual analysis to post-categorization processes. *Front. Hum. Neurosci.* 8:165. doi: 10.3389/fnhum.2014.00165

Price, G. R., Palmer, D., Battista, C., and Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychol. (Amst)* 140, 50–57. doi: 10.1016/j.actpsy.2012.02.008

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., and Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychol. Sci.* 19, 607–614. doi: 10.1111/j.1467-9280.2008.02130.x

Roitman, J. D., Brannon, E. M., and Platt, M. L. (2007). Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS Biol.* 5:e208. doi: 10.1371/journal.pbio.0050208

Rubinsten, O., Dana, S., Lavro, D., and Berger, A. (2013). Processing ordinality and quantity: ERP evidence of separate mechanisms. *Brain Cogn.* 82, 201–212. doi: 10.1016/j.bandc.2013.04.008

Sasanguie, D., De Smedt, B., and Reynvoet, B. (2017). Evidence for distinct magnitude systems for symbolic and non-symbolic number. *Psychol. Res.* 81, 231–242. doi: 10.1007/s00426-015-0734-1

Schindler, M., Schovenberg, V., and Schabmann, A. (2020). Enumeration processes of children with mathematical difficulties: an explorative eye-tracking study on subitizing, groupitizing, counting, and pattern recognition. *Learn. Disabil. A Contemp. J.* 18, 193–211.

Schleifer, P., and Landerl, K. (2011). Subitizing and counting in typical and atypical development. *Dev. Sci.* 14, 280–291. doi: 10.1111/j.1467-7687.2010.00976.x

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., et al. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev. Sci.* 20:e12372. doi: 10.1111/desc.12372

Sella, F., Lanfranchi, S., and Zorzi, M. (2013). Enumeration skills in down syndrome. *Res. Dev. Disabil.* 34, 3798–3806. doi: 10.1016/j.ridd.2013.07.038

Sella, F., Onnivello, S., Lunardon, M., Lanfranchi, S., and Zorzi, M. (2021). Training basic numerical skills in children with Down syndrome using the computerized game "The Number Race.". *Sci. Rep.* 11, 1–14. doi: 10.1038/s41598-020-78801-5

Sengupta, R., Bapiraju, S., and Melcher, D. (2017). Big and small numbers: empirical support for a single, flexible mechanism for numerosity perception. *Attent. Percept. Psychophys.* 79, 253–266. doi: 10.3758/s13414-016-1221-5

Simon, T. J., Peterson, S., Patel, G., and Sathian, K. (1998). Do the magnocellular and parvocellular visual pathways contribute differentially to subitizing and counting? *Percept. Psychophys.* 60, 451–464.

Sokolowski, H. M., Fias, W., Bosah Ononye, C., and Ansari, D. (2017). Are numbers grounded in a general magnitude processing system? A functional neuroimaging meta-analysis. *Neuropsychologia* 105, 50–69. doi: 10.1016/j.neuropsychologia.2017.01.019

Soltész, F., and Szucs, D. (2014). Neural adaptation to non-symbolic number and visual shape: an electrophysiological study. *Biol. Psychol.* 103, 203–211. doi: 10.1016/j.biopsycho.2014.09.006

Starkey, G. S., and McCandliss, B. D. (2014). The emergence of "groupitizing" in children's numerical cognition. *J. Exp. Child Psychol.* 126, 120–137. doi: 10.1016/j.jecp.2014.03.006

Starr, A., Libertus, M. E., and Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18116–18120. doi: 10.1073/pnas.1302751110

Stoianov, I., and Zorzi, M. (2012). Emergence of a "visual number sense" in hierarchical generative models. *Nat. Neurosci.* 15, 194–196. doi: 10.1038/nn.2996

Summerfield, C., and de Lange, F. P. (2014). Expectation in perceptual decision making: neural and computational mechanisms. *Nat. Rev. Neurosci.* 15, 745–756. doi: 10.1038/nrn3838

Szkudlarek, E., Park, J., and Brannon, E. M. (2021). Failure to replicate the benefit of approximate arithmetic training for symbolic arithmetic fluency in adults. *Cognition* 207:104521. doi: 10.1016/j.cognition.2020.104521

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011:879716. doi: 10.1155/2011/879716

Temple, E., and Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7836–7841. doi: 10.1073/pnas.95.13.7836

Testolin, A., Dolfi, S., Rochus, M., and Zorzi, M. (2020). Visual sense of number vs. sense of magnitude in humans and machines. *Sci. Rep.* 10, 1–13. doi: 10.1038/s41598-020-66838-5

Trick, L. M., and Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychol. Rev.* 101:80.

Tudusciuc, O., and Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14513–14518. doi: 10.1073/pnas.0705495104

Turconi, E., Jemel, B., Rossion, B., and Seron, X. (2004). Electrophysiological evidence for differential processing of numerical quantity and order in humans. *Cogn. Brain Res.* 21, 22–38. doi: 10.1016/j.cogbrainres.2004.05.003

van Hoogmoed, A. H., and Kroesbergen, E. H. (2018). On the difference between numerosity processing and number processing. *Front. Psychol.* 9:1650. doi: 10.3389/fpsyg.2018.01650

van Marle, K., Chu, F. W., Li, Y., and Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Dev. Sci.* 17, 492–505. doi: 10.1111/desc.12143

Vandierendonck, A. (2021). On the utility of integrated speed-accuracy measures when speed-accuracy trade-off is present. *J. Cogn.* 4:22. doi: 10.5334/joc.154

Viswanathan, P., and Nieder, A. (2013). Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11187–11192. doi: 10.1073/pnas.1308141110

Woodman, G. F., and Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychon. Bull. Rev.* 15, 223–229. doi: 10.3758/PBR.15.1.223

Zorzi, M., and Testolin, A. (2018). An emergentist perspective on the origin of number sense. *Philos. Trans. R. Soc. B Biol. Sci.* 373:20170043. doi: 10.1098/rstb.2017.0043

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers
in Human Neuroscience

ORIGINAL RESEARCH
published: 25 February 2022
doi: 10.3389/fnhum.2022.752190

# Young Children Intuitively Divide Before They Recognize the Division Symbol

**Emily Szkudlarek*, Haobai Zhang, Nicholas K. DeWind and Elizabeth M. Brannon**

*Department of Psychology, University of Pennsylvania, Philadelphia, PA, United States*

Children bring intuitive arithmetic knowledge to the classroom before formal instruction in mathematics begins. For example, children can use their number sense to add, subtract, compare ratios, and even perform scaling operations that increase or decrease a set of dots by a factor of 2 or 4. However, it is currently unknown whether children can engage in a true division operation before formal mathematical instruction. Here we examined the ability of 6- to 9-year-old children and college students to perform symbolic and non-symbolic approximate division. Subjects were presented with non-symbolic (dot array) or symbolic (Arabic numeral) dividends ranging from 32 to 185, and non-symbolic divisors ranging from 2 to 8. Subjects compared their imagined quotient to a visible target quantity. Both children (Experiment 1 $N = 89$, Experiment 2 $N = 42$) and adults (Experiment 3 $N = 87$) were successful at the approximate division tasks in both dots and numeral formats. This was true even among the subset of children that could not recognize the division symbol or solve simple division equations, suggesting intuitive division ability precedes formal division instruction. For both children and adults, the ability to divide non-symbolically mediated the relation between Approximate Number System (ANS) acuity and symbolic math performance, suggesting that the ability to calculate non-symbolically may be a mechanism of the relation between ANS acuity and symbolic math. Our findings highlight the intuitive arithmetic abilities children possess before formal math instruction.

**Keywords: mathematical ability, number sense, division, arithmetic, approximate number system, approximate arithmetic**

OPEN ACCESS

**Edited by:**
*Roberto Arrighi,*
*University of Florence, Italy*

**Reviewed by:**
*Fabrizio Doricchi,*
*Sapienza University of Rome, Italy*
*Maria Dolores de Hevia,*
*Centre National de la Recherche Scientifique (CNRS), France*
*Attila Krajcsi,*
*Eötvös Loránd University, Hungary*

***Correspondence:**
*Emily Szkudlarek*
*emilyszkudlarek@gmail.com*

**Specialty section:**
*This article was submitted to Cognitive Neuroscience, a section of the journal Frontiers in Human Neuroscience*

**Received:** *02 August 2021*
**Accepted:** *24 January 2022*
**Published:** *25 February 2022*

**Citation:**
*Szkudlarek E, Zhang H, DeWind NK and Brannon EM (2022) Young Children Intuitively Divide Before They Recognize the Division Symbol. Front. Hum. Neurosci. 16:752190. doi: 10.3389/fnhum.2022.752190*

## INTRODUCTION

Arithmetic skills underlie the entire elementary school math curriculum (National Governors Association Center for Best Practices, and Council of Chief State School Officers, 2019). Mastery of early arithmetic begins a cascade that unlocks the opportunity to study more advanced branches of mathematics such as algebra, geometry and calculus. According to the US Common Core Standards children learn arithmetic operations in a sequence starting with addition and subtraction, then multiplication, and finally division beginning in grade 3. Division is commonly introduced as the inverse of multiplication, and children's early understanding of division is mediated via multiplication. Only later in more advanced math education do these representations diverge (Campbell, 1997; Mauro et al., 2003). There is neural and behavioral evidence that division remains more effortful than the other basic arithmetic operations even into adulthood (Ischebeck et al., 2009; Rosenberg-Lee et al., 2011). These findings suggest that division is the most difficult of the four basic arithmetic operations. However, this greater difficultly may be a function of how formal division is taught and not a fundamental aspect of the operation.

Children have some basic intuitions about division before they formally learn how to divide. These basic intuitions may derive from insights into practical mathematics in the world around them, called intuitive action schemas (Riley, 1984; Jitendra and Hoff, 1996; Correa et al., 1998). One hypothesized action schema that supports division is children's knowledge of how to fairly distribute items amongst people (Blake and McAuliffe, 2011; Shaw and Olson, 2012; Sheskin et al., 2016; Hamamouche et al., 2020).

Another way in which children may begin to form a concept of division is through many-to-one counting to solve multiplication and division word problems. For example, when kindergarteners were presented with the problem "Tad has 15 guppies. He put 3 guppies in each jar. How many jars did Tad put guppies in?" children demonstrated many-to-one counting by counting out 15 guppies into groups of 3, and then counting the number of groups (Carpenter et al., 1993). However, these strategies usually require external support and small set sizes.

A third way children could develop an intuitive sense of division is through experience with their non-symbolic sense of number. The Approximate Number System (ANS) allows children to approximately represent, compare, estimate, and calculate with large sets of objects (Feigenson et al., 2004). A substantial body of work demonstrates that adults, children, infants, and non-human primates can use ANS representations to add and subtract arrays of objects (McCrink and Wynn, 2004; Pica et al., 2004; Barth et al., 2005, 2006; Knops et al., 2009; McNeil et al., 2011; Gunderson et al., 2012; Pinheiro-Chagas et al., 2014; Xenidou-Dervou et al., 2014; Cantlon et al., 2015). Young children can also perform scaling operations on large arrays of objects and multi-step operations (Barth et al., 2009; McCrink and Spelke, 2010, 2016; McCrink et al., 2013, 2016). By using their ANS, children can even solve addend unknown algebra problems (Kibbe and Feigenson, 2015, 2017), and compare ratios of discrete sets of items (Falk et al., 2012). This work indicates that ANS representations can be used in a variety of non-symbolic and approximate mathematical contexts.

It is still an open question whether children can use their ANS to compute a true non-symbolic, approximate division operation. In the approximate scaling task used by McCrink and colleagues a child saw a large set of items which were then hidden behind a white box. A 'dividing wand' appeared on top of the box and the child was told "Look! They're getting divided". During training, the child watched the dividing wand halve (or quarter in another experiment) a set of objects. During testing, the child compared their imagined quotient to a target set of objects and picked the larger set. Children's accuracy varied as a function of the ratio between the halved or quartered array and the visible target array, a hallmark of ANS representations. There are two ways in which this scaling task differs from a non-symbolic, approximate division task. First, scaling operations are a specific case of a division operation where the divisor is held constant. A true division operation requires both the dividend and divisor to hold multiple values. Second, this task is not entirely non-symbolic because there is a specific one-to-one correspondence between the 'dividing wand' symbol and a given divisor. Thus, it is unknown whether children can use their non-symbolic, approximate sense of number to perform non-symbolic division.

The first goal of the current experiment was to determine whether young children can intuitively divide large quantities with their ANS. To answer this question, we developed a novel non-symbolic division paradigm where both the dividend and the divisor are non-symbolic quantities that vary from trial to trial. Using multiple divisors within one subject allowed us to ask whether children truly have an intuitive sense of division, or whether children are limited to the halving or quartering operations demonstrated previously (McCrink and Spelke, 2010, 2016). To determine whether intuitive division operates over ANS representations, we tested whether accuracy on our non-symbolic division task was dependent on the ratio between the quotient and a target comparison value. Ratio dependent accuracy is a hallmark of the ANS (Feigenson et al., 2004). As a stronger test of our hypothesis, we also independently measured each child's ANS acuity using a dot comparison task and examined the correlation between intuitive division accuracy and ANS acuity. If children indeed use their ANS to perform approximate division, children with better ANS acuity should perform more accurately on our approximate division task.

As a further test of children's intuitive division competence, participants also completed a *symbolic*, approximate division task. This task was animated in the same way as the non-symbolic, approximate division task; however, the dot-array dividends and targets were replaced with numerals. Previous research demonstrates that children can perform symbolic, approximate addition and subtraction (Gilmore et al., 2007), mixed symbolic to non-symbolic ratio comparisons (Kalra et al., 2020) and fully symbolic, approximate ratio comparisons (Szkudlarek and Brannon, 2021) before formal instruction. Successful performance on our symbolic, approximate division task would indicate that intuitive division performance is not specific to the numerical magnitude representation afforded by dot arrays, but rather to numerical magnitude representation. Thus, in the current experiment we test whether children's intuitive division abilities can extend to symbolic division.

Our second goal was to explore how approximate division skill relates to formal teaching about the division operation. If non-symbolic and symbolic approximate division tasks have any use in pedagogical context, they may be most helpful before formal division teaching begins. Our sample included children aged 6 to 9, which spans the age range before and during the beginning of formal division instruction. To ensure that children's intuitive large number division skill was not dependent on prior instruction about the division operation, we quantified children's level of symbolic division knowledge with a test of their symbolic, exact division skill. Then, we tested whether children can successfully approximately divide before they have formal knowledge of division as a math operation.

The third and final goal of the current experiment was to examine whether intuitive division skill provides a link between ANS acuity and formal mathematics. Prior work has demonstrated that ANS acuity and symbolic math performance are correlated in children and adults (Chen and Li, 2014; Fazio et al., 2014; Schneider et al., 2016). However, recent findings suggest that performing a mathematical operation non-symbolically and approximately may be a better predictor of symbolic math ability than ANS acuity in both children and adults (Pinheiro-Chagas et al., 2014; Matthews et al., 2016; Starr et al., 2016; Szkudlarek and Brannon, 2021). In the context of the current experiment, sharper ANS acuity may allow for better non-symbolic division calculation. In turn, better non-symbolic division ability may provide students stronger conceptual models of division operations. This stronger concept of division may lead to a sturdier ability to tackle symbolic, exact division calculation in the classroom. Accordingly, we predict a significant correlation between non-symbolic division accuracy and formal mathematical skill, as measured with the Key-Math-3 Numeration test (Connolly, 2007). Furthermore, we predict that intuitive division accuracy will mediate the correlation between ANS acuity and performance on the Key-Math-3 Numeration test. We tested this mediation hypothesis in both children and university undergraduates. Our mediation hypothesis is particularly interesting for adults because there is currently a lack of theorized mechanisms for why the relation between ANS acuity and symbolic mathematics persists into adulthood. For example, the theory that sharper ANS acuity promotes the initial learning of number words does not explain why ANS acuity would still be linked to math skills in adulthood (Odic et al., 2015). If the ability to model arithmetic operations using the ANS is a mechanism of the link to symbolic math, sharper ANS acuity could indirectly impact complex math abilities later in development, opening up further paths of inquiry to explore this relation.

We explored intuitive division ability across three Experiments. In Experiment 1, 6-9 year old children completed non-symbolic and symbolic division tasks and measures of ANS acuity and formal math to examine whether children could successfully perform intuitive division before formal knowledge of the division operation, and whether this ability could serve as a pathway between ANS acuity and symbolic math ability. In Experiment 2, we tested a new cohort of 6-9 year old children on the division tasks using different numerical values to rule out alternative calculation strategies and to replicate the finding that children can perform intuitive division before formal division knowledge. In Experiment 3, we examined university undergraduate's ability to perform intuitive division, and whether this ability can continue to provide a bridge between ANS acuity and formal math ability into adulthood.

# MATERIALS AND METHODS

## Child Experiments

### Subjects

Eighty-nine 6-9 year-old children participated in Experiment 1 (mean age = 7.9 years-old, standard deviation = 1.1 years; 50 female, 39 male). Written parental consent was collected in accordance with a protocol accepted by the University of [blinded] Institutional Review Board. Thirty-two additional children were consented but did not complete both the non-symbolic and symbolic division tasks due to absence and were excluded. The parents of 88 children in the sample completed a detailed demographics questionnaire. 86% identified as Black or African American, 10% as White, 2% as Asian, and 2% as more than one race. A large proportion of children came from families with household incomes of $50,000 or less (7% $150,000+, 6% $150,000 - $100,000, 4% $75,000 - $50,000, 45% $50,000 - $25,000, 30% $25,0000 - $0, and 8% chose not to report). All subjects were recruited from six after school programs in the Philadelphia, PA area. A subset of the children who completed both the non-symbolic and symbolic division tasks completed additional assessments (Dot comparison, $n$ = 84; Key-Math Numeration subtest, $n$ = 89; Division knowledge assessment, $n$ = 82; the Woodcock-Johnson Basic Reading Skills cluster, $n$ = 77; and a measure of numeral identification, $n$ = 80). All participants received a small toy as a thank you gift.

To replicate our results and ensure that children were indeed performing approximate division to solve our division tasks, forty-two children (mean age = 7.9, standard deviation = 1.2 years, 16 female, 12 male, 14 did not report) were tested in Experiment 2 on the same division tasks with different numerical values from Experiment 1. None of the children participated in Experiment 1. The parents of twenty-five participants completed our demographic form. 56% identified as Black or African American, 12% identified as White, 4% as more than one race, and 28% chose not to report. The majority of children came from households with incomes of $50,000 or less (8% $100,000+, 12% $75,000 - $50,000, 32% $50,000 - $25,000, 8% $25,0000 - $0, and 40% chose not to report). Children in Experiment 2 were tested on the non-symbolic and symbolic division tasks, a measure of numeral identification and the formal division test.

### Procedure

Children in both experiments completed all tasks individually with an experimenter in a quiet room at their after-school program. Children completed the non-symbolic and symbolic division tasks first and the order of the two tasks were counterbalanced across children. The order in which all other tasks were administered was random across participants and was dependent on the duration of the task and the child's availability. The approximate division tasks and the dot comparison task were run in MATLAB and programmed using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). The approximate division tasks, video of the tasks, and data are available online at OSF (https://osf.io/g5y27/?view_only=0a2ab8862d9d4e95b2426cb11657e78d).

The non-symbolic and symbolic division tasks and dot comparison task were run on a 15-inch touch screen laptop computer. Children in Experiment 1 completed the division tasks, dot comparison task, Key-Math-3 Numeration subtest, Woodcock-Johnson Reading Cluster, formal division test and numeral identification task for a total of 45-60 min across 2-3 days. Children also completed a math anxiety questionnaire, but these results are not reported in the current manuscript. Children in Experiment 2 completed both division tasks, numeral identification test, and formal division task in one session. All children received stickers to maintain motivation.

## Experimental Tasks

### Introduction to the Non-Symbolic and Symbolic Division Tasks

Children were introduced to a bee named "Buzz" on the computer screen. The experimenter read the following story: "Buzz flies to flowers to find food to bring back to his hive. Buzz lands on the flower to get the food, and some of the food sticks to him. When Buzz flies away from the flower, some of the food falls down onto the flower." A picture on the computer screen showed Buzz carrying four dots above a flower with two petals. Four dots fell toward the flower with two dots landing on each petal. The children were told, "The same amount of food falls on each petal of the flower." The children were then shown Buzz above a flower with four petals and eight dots falling toward the flower. The children were told, "We can see the food falling down toward the flower. See how even if the flower looks different, the same amount of food falls onto each petal." Children were then shown Buzz above a flower with eight petals and eight dots falling toward the flower. The instructions were repeated one more time. The experimenter never mentioned the number of dots or petals.

### Non-Symbolic Division

Please see **supplementary videos 1**-**4** for a video recording of the division tasks. After the initial instructions, the demonstration phase of the game began (**Figure 1A**). Children watched an animated set of dots fall onto the petals of a flower. On demonstration trials, children could see how many dots fell onto each flower petal (i.e., they saw the result of the division operation). The initial number of dots is the dividend, the number of petals is the divisor, and the number of dots that fall onto one petal is the quotient. After the dots fell onto the flower petals the flower disappeared and one of the petals from the flower moved to the middle of the left side of the screen. A new flower petal with dots already inside appeared on the right side of the screen. The experimenter asked, "Which petal has more food?" The child was told they should touch the petal to indicate their answer. The trial did not progress until the child made their response, but the experimenter encouraged the child to make their choice quickly. Once the child touched a petal, a happy bee with the words "Great job!" appeared for the correct response or a sad bee with the words "try again!" appeared for an incorrect response. Then, a screen appeared with Buzz in the center. The child was told to touch Buzz to continue playing the game. Touching Buzz started the next trial. Children completed eight demonstration trials. The purpose of these trials was to ensure children understood that the same number of dots fall into each petal of the flower, and that their job was to pick the petal that had more food. These trials were not used in any analyses because the result of the division operation was visible to the child on these trials. During the demonstration phase flowers had 2, 5, and 8 petals.

After completing the demonstration phase children were told, "Now it is a foggy day in the garden, but Buzz still needs your help. You won't have enough time to count all the food Buzz is carrying, and because of the weather, you won't be able to see the amount of food that falls onto each petal. Instead, you'll need to *imagine* how many pieces of food are on each petal. Remember, the same amount of food falls onto each petal." After these instructions, the experimenter started the first experimental trial (**Figure 1C**). An array of dots and an empty flower appeared on the left side of the screen like before, but this time when the dots began to fall toward the flower a fog appeared over the flower. Children could still see the outline of the flower petals and the dots falling, but the dots were obscured before they were distributed onto the petals. Children could thus no longer see how many dots landed on each flower petal but had to mentally divide the dots by the number of petals to infer the quotient. Like in the demonstration, the flower disappeared and one of the petals moved up to the middle of the left side of the screen. However, this time the inside of the petal was foggy (gray) so that the child could not see the number of dots inside the petal. Another flower petal appeared on the right side of the screen with dots visible inside. Then the experimenter said, "Ok, which petal has more food? Try and imagine how many pieces of food are on this [gesture to left] petal even though you can't see them!" Children then responded by touching the petal they thought had the greater quantity of dots and received feedback. Children completed 32[1] trials with automated feedback. During these 32 trials children saw flowers with 2, 5, or 8 petals in random order. To test whether children could generalize to new divisors, after the completion of the first 32 trials children completed 24[2] more trials with 3 or 6 petals without automated feedback. Throughout all trials the experimenter never mentioned any number words. Note that the animation always occurred on the left side of the screen so that children could anticipate where to attend, however, the correct choice (i.e., larger quantity) was counterbalanced across all trials. Accuracy and reaction time were recorded for each trial.

### Symbolic Division

The symbolic division task was identical to the non-symbolic version, except that all dividends and comparison targets were displayed with Arabic numerals instead of dots (**Figure 1D**). The instructions and the numerical values remained identical between both versions of the division tasks. Children received eight demonstration trials at the beginning of the task (**Figure 1B**). During these trials, the numeral at the top of the screen (dividend) fell toward the flower and split apart into the numeral representing quotients (e.g., the numeral "32" split into four numerals "8" and each numeral "8" fell onto one of the four petals in the above example). Children then completed 32 trials with

---

[1] 11 subjects ran a longer version of the division tasks (53 trials with feedback and 40 trials without feedback) but the number of trials for the remaining subjects was reduced due to time constraints at the after-school programs where testing occurred.

[2] One subject ran 9 extra trials without feedback due to a computer malfunction. Two subjects ran a majority (19 or 20 out of 24) but not all 24 trials due to an early pick up time. These subjects are included in the analysis.

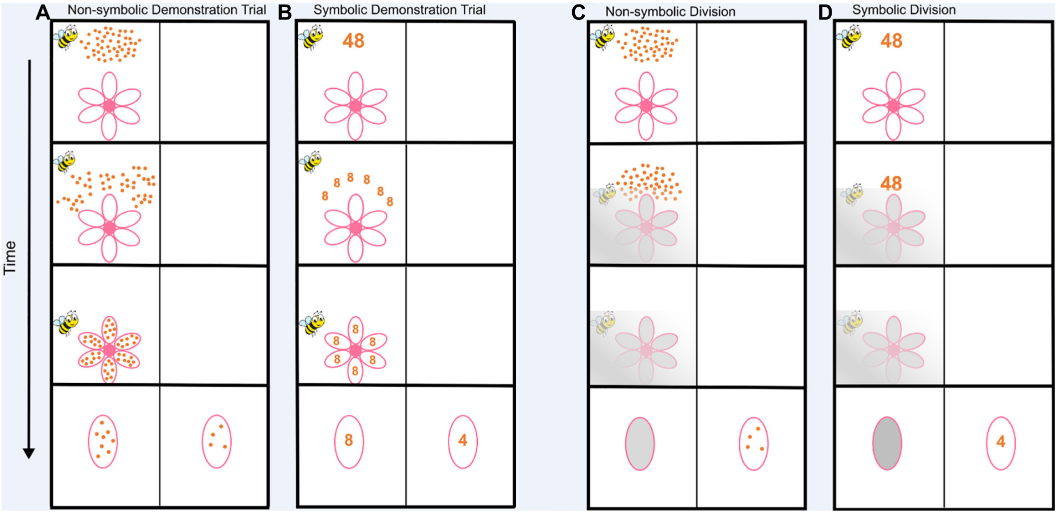


**FIGURE 1 |** Schematic of the non-symbolic and symbolic division tasks. **(A)** Demonstration trials for the non-symbolic task. Children and adults watched as the dots on the top of the screen fell onto the petals below. Then, one of the flower petals moved toward the center of the left side of the screen and the target petal appeared with a new quantity of dots to compare on the right side of the screen. Participants responded by touching the petal with the greater quantity of dots. Participants completed 8 demonstration trials. The demonstration trials are not included in any analysis. **(B)** The demonstration trials for the symbolic version of the task. Participants watched the numeral at the top of the screen split apart and change into the numerals that landed on the flower at the bottom of the screen. **(C)** Experimental trials for the non-symbolic division task. The task was identical to the demo trials, except that as the dots fell to the bottom of the screen a cloud appeared that obscured how many dots fell onto each petal. Then the obscured petal moved to the middle of the left side of the screen and subjects had to imagine how many dots were on this petal and compare it to the visible target quantity. **(D)** The experimental trials for the symbolic version of the task. Participants watched the dividend numeral move and disappear into the fog without viewing the quotient. Then participants compared their imagined quotient to the new target number on the right side of the screen. For a video of the division tasks please see https://osf.io/g5y27/?view_only=b57c188ca72f4b48a0447fdff1470dc9.

feedback and 24 trials without feedback. For example, on a given trial a child might see the numeral "32" and the numeral would then float down behind the fog onto a flower with four petals. The child would have to imagine "8" on each petal. If the child were presented with a target comparison petal with the numeral "4" the correct answer would be the foggy petal (8 > 4).

## Numerical Values for the Division Tasks

We chose numerical values for the approximate division tasks to ensure that children had to pay attention to all three numbers (dividend, divisor, target) to solve the task successfully. All target comparison numbers were drawn from the same values as the possible quotients. The quotients used (8, 10, 13, 17, 22, 29, 37, 48) were chosen to be approximately evenly spaced on a log scale (**Supplementary Figure 1**). This allowed the ratio between the quotient and the target to be independent of the magnitude of the divisor, dividend and quotient. This is important because the difficulty between comparing any two numerosities is dependent on the ratio between them (Feigenson et al., 2004). By including an equal number of trials at each ratio for each divisor we could compare accuracy at each divisor, and test whether participants could generalize to novel divisors after learning the non-symbolic division task.

Experiment 1 included 32 trials with feedback testing divisors of 2, 5, and 8 and 24 trials without feedback testing divisors of 3 and 6 (**Supplementary Table 1**). In Experiment 2 children completed 32 trials with feedback testing divisors of 2, 5, and 8 and 24 trials without feedback testing divisors of 3, 4, and 6 (**Supplementary Table 2**). The numerical values chosen in Experiment 2 ensured that participants would have chance level performance if they chose their answer based on the median target value displayed.

## Dot Comparison Task

Two dot arrays appeared on a black screen for 750 ms. The arrays were then occluded, and the task was to touch the numerically larger array. Children completed 200 trials with feedback on every trial. The number of dots ranged from 8 to 32. The stimuli were created to evenly sample a stimulus space that varied by the ratio between the number, size, and the spacing of the dots. To encourage greater reliability of the measurement, trial level difficulty was titrated (Lindskog et al., 2013). The titration procedure calculated the percentage correct over the last five trials. The ratio between the two dot arrays moved to an easier ratio if accuracy was 3 out of 5 or less, stayed the same if accuracy was 4 out of 5, and moved to a more difficult ratio if the accuracy was 5 out of 5. A quantitative index of each child's ANS acuity was calculated as a Weber fraction ($w$) as specified in DeWind et al. (2015). This model accounts for the effects of non-numerical features of dot arrays on numerical discrimination and has been shown to provide more reliable estimates of $w$ (DeWind and Brannon, 2016).

## Numeral Identification Task

The numerals 1-30 were printed and displayed individually on index cards. The numerals were displayed in random order, and the child was asked "What number is this?" Accuracy was recorded. The majority of children in our sample successfully recognized all numerals 1-30 (69 out of 79 participants tested) and so this task was not included in subsequent analyses.

### Key Math-3 Diagnostic Assessment

The Numeration section of the Key Math-3 Diagnostic Assessment Form B (Connolly, 2007) is a test of general basic math skills like place value, counting, the relative magnitude of numbers. It also tests understanding of fractions, decimals, and percentages. We used the age standardized scale score.

### Woodcock-Johnson IV Test of Cognitive Abilities

Participants' reading abilities were assessed using the "Basic Reading Skills" cluster of the Woodcock-Johnson. This cluster is comprised of the "Letter-Word Identification" and "Word Attack" subtests. In the "Letter-Word Identification" subtest, participants named letters and read words aloud. In "Word Attack," participants read nonsense words and identified letter sounds. We used the age standardized Basic Reading Skills score.

### Formal Division Test

We created a test of 15 questions that examined children's addition and division knowledge. Six items were word problems, eight items were symbolic arithmetic problems, and one item required the experimenter to show the child a picture of the division symbol (÷) and ask, "Do you know what this symbol is?" For each arithmetic problem, the child was shown a flashcard with the arithmetic equation as the experimenter read the problem aloud. The test questions are reproduced in **Supplementary Table 3**. A division knowledge score was calculated based on a child's accuracy on the four symbolic division problems (range 0-4) and whether or not they could identify the division symbol.

## Experiment 3

### Subjects

Participants were eighty-seven undergraduates (mean age 20.7 years old, 51 female). Written and informed consent was collected in accordance with a protocol accepted by the University of [blinded] Institutional Review Board. Seven participants did not return to complete the second session and were thus excluded from the mediation analysis. The data from two dot comparison scores and two fraction magnitude comparison scores were lost due to computer error.

### Procedure

Adults completed all tasks in two sessions that took place on separate days no more than 3 days apart and received course credit as compensation. Testing occurred in a quiet room on a touch screen desktop computer. During the first session adults completed the non-symbolic and symbolic division tasks in counterbalanced order, the vocabulary test, and the division strategy questionnaire. During the second session subjects completed an addition verification task, a dot comparison task, and a fraction magnitude comparison task in counterbalanced order. Subjects completed a math anxiety questionnaire, but this data is not included in the current report.

### Non-Symbolic and Symbolic Division Tasks

The tasks and instructions for adults were identical to those described for children in Experiment 1. The participants were told that this task was created for use with children to explain the presence of the cartoon bee and storyline.

### Dot Comparison Task

The task was the same as that described for the children.

### Vocabulary Test

Subjects answered 42 multiple choice vocabulary questions in 5 min. The questions were taken from the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al., 1976). Performance was calculated as the number of problems answered correctly minus 1/4 of the number incorrect to discourage guessing.

### Addition Verification Test

One and two digit addition and subtraction problems were displayed horizontally with a proposed answer (e.g., $27 + 52 = 79$). Subjects pressed the F or J key (counterbalanced) if the statement was correct and the F or J key if it was incorrect. On incorrect trials (50% of all trials) the sum displayed was ± 10 or ± 2 from the correct sum as modeled after (Klein et al., 2010). Participants had 10 seconds to make a response. Subjects completed two blocks of 96 trials each. Performance was quantified as the median reaction time on correctly rejected incorrect equations.

### Fraction Magnitude Comparison Task

Subjects viewed two fractions displayed in the middle of the screen in white on a black background. The goal of the task was to pick the fraction greater in magnitude by pressing the F key for the left fraction or the J key for the right fraction. The stimuli were the same as used in Fazio et al. (2015). Accuracy and reaction time were recorded.

### Division Strategy Questionnaire

The goal of this questionnaire was to examine the strategies adults used to solve the non-symbolic and symbolic division tasks.

The full questionnaire and results are reported in the **Supplementary Table 4**.

## Analysis Plan

### Alternative Heuristic Analysis

We conducted a series of analyses to test the possibility that participants were using an alternative heuristic instead of approximately dividing. If participants attempted to compare the divisor (number of petals) to the target comparison number when making their response performance would not exceed chance expectations since the target was greater than the divisor on all trials. Alternatively, participants could attempt to compare only the dividend to the target comparison number. The target was greater than the dividend on only three trials (3/56), and so we confirmed that children and adults performed at above chance levels when excluding those three trials.

We next examined whether participants used a heuristic where they constructed a mental model of the median of the target value across all trials and evaluated whether the target on a given trial was more or less than the median target value. Using this heuristic, subjects would pick the target value if it were greater than the median target value and they would pick the imagined quotient if the target were less than the median target value. The stimulus set used in Experiment 1 was not designed to rule out this alternative strategy, however, the stimulus set constructed for Experiment 2 ensured that participants could not score above chance if they relied on this strategy. Thus, above chance performance in Experiment 2 rules out the possibility that children rely on the median target strategy to solve the approximate division tasks. We tested for use of the median target heuristic in adults by calculating accuracy on the few trials where the median target strategy was ineffective.

### Mediation Analysis

We ran mediation models to test the hypothesis that non-symbolic division mediates the relationship between ANS acuity and formal mathematics ability. We removed any outlier scores greater or less than three times the interquartile range for children and adults. This process removed four ANS acuity scores from the child dataset, and 3 symbolic division scores from the adult dataset. We used the natural log transformation on both child and adult ANS acuity scores (Child ANS acuity Shapiro-Wilk W = 0.96; adult ANS acuity W = 0.91, W value close to 1 represents a normal distribution). Bivariate correlations and descriptive statistics are reported for children in Experiment 1 in **Supplementary Table 5** and adults in **Supplementary Table 6**. To ensure that correlations between measures were not simply due to age in the children, we partialed out age from our measures of ANS acuity and symbolic and non-symbolic division.

To measure formal math ability we used the Key-Math-3 Numeration subtest in children, and accuracy on the fraction magnitude comparison test in adults. We did not run a mediation model using the Addition Verification measure in adults because this measure was not significantly correlated with ANS acuity (**Supplementary Table 6**; $r = -0.14$, $p = 0.21$). Mediation analyses test for a significant indirect effect that accounts for some portion of the original direct effect. The goal of this analysis was to examine whether non-symbolic division skill accounts for significant variance in the relation between ANS acuity and symbolic math ability in both children and adults. A significant mediation would be consistent with our hypothesis that non-symbolic division calculation is a mechanism of the relation between ANS acuity and symbolic math, though we cannot test for causality in our cross-sectional design. A significant mediation result in our adult participants would support the idea that non-symbolic calculation skill underlies the small but significant relation between ANS acuity and symbolic math in adults expert in symbolic number (Schneider et al., 2016).

## RESULTS

### Non-Symbolic Division Performance

Children and adults performed well above chance expectations on both the feedback (children 77%, $t_{88} = 27.4$, $p < 0.001$, $d = 2.9$; adults 89%, $t_{86} = 60.1$, $p < 0.001$, $d = 6.4$) and no feedback (children 73%, $t_{88} = 19.8$, $p < 0.001$, $d = 2.1$; adults 88%, $t_{86} = 51.2$, $p < 0.001$, $d = 5.5$) phases of the non-symbolic division task (**Figure 2**). These data indicate successful generalization of the division operation to novel divisors and demonstrate that both children and adults engaged in approximate division.

In Experiment 2, with stimuli designed to rule out a median target alternative strategy, we replicated the finding that children can perform non-symbolic intuitive division and generalize to novel divisors with a different set of numerical values chosen from the same stimulus space (**Figure 2**; feedback 69% accuracy $t_{41} = 12.8$, $p < 0.001$, $d = 2.0$; no feedback 74% accuracy $t_{41} = 15.9$, $p < 0.001$, $d = 2.5$).

### Symbolic Division Performance

Children and adults performed well above chance on both the feedback (children 72%, $t_{88} = 17.4$, $p < 0.001$, $d = 1.8$; adults 96%, $t_{86} = 72.8$, $p < 0.001$, $d = 7.8$) and no feedback (children 67%, $t_{88} = 11.5$, $p < 0.001$, $d = 1.2$; adults 95%, $t_{86} = 63.6$, $p < 0.001$, $d = 6.8$) phases of the symbolic division task (**Figure 3**). We replicated this above chance performance with children in Experiment 2 (feedback 62% accuracy $t_{41} = 5.99$, $p < 0.001$, $d = 0.92$; no feedback 60% accuracy $t_{41} = 6.01$, $p < 0.001$, $d = 0.93$).

### Adult and Child Division Format Effect

We compared the relative performance of adults and children in Experiment 1 because these experiments were run using the same numerical values. We ran a mixed effects ANOVA predicting overall performance on the division tasks with a main effect of task format (symbolic or non-symbolic) and age group (adult or child), an interaction between format and age, and a random effect of subject. There was a main effect of age group on division performance, ($F_{1,174} = 332.8$, $p < 0.001$) and a significant age by task format interaction ($F_{1,174} = 106.1$, $p < 0.001$). The main effect indicated that adults performed with higher accuracy overall ($t_{174} = 18.2$, $p < 0.001$, $d = 2.8$). Follow up tests on the interaction indicated that adults performed with higher accuracy on the symbolic as compared to the non-symbolic version of the division task (paired *t*-test: $t_{86} = -10.5$, $p < 0.001$, $d = 1.3$), whereas children showed the opposite effect. Children performed significantly better on the non-symbolic version of the task (paired t-test: $t_{88} = 5.52$, $p < 0.001$, $d = 0.54$). This format effect held even among children who could recognize all the numerals 1-30, indicating that greater accuracy on the non-symbolic task was not due to a lack of basic numeral knowledge ($t_{78} = 5.05$, $p < 0.001$, $d = 0.54$).

We ran the same analysis on median reaction time on correct trials. There was a main effect of age ($F_{1,174} = 46.6$, $p < 0.001$) driven by the fact that adults were faster than children (1.11 seconds vs. 1.84 seconds $t_{174} = 11.0$, $p < 0.001$). There was also a significant main effect of task format ($F_{1,174} = 23.6$, $p < 0.001$) driven by the fact that both adults and children were faster to perform approximate division on the non-symbolic than the symbolic task (paired *t*-test $t_{175} = 4.87$, $p < 0.001$; median RT adult non-symbolic = 1.01 s, median RT adult symbolic = 1.10 s, median RT non-symbolic child = 1.62 s, median RT symbolic

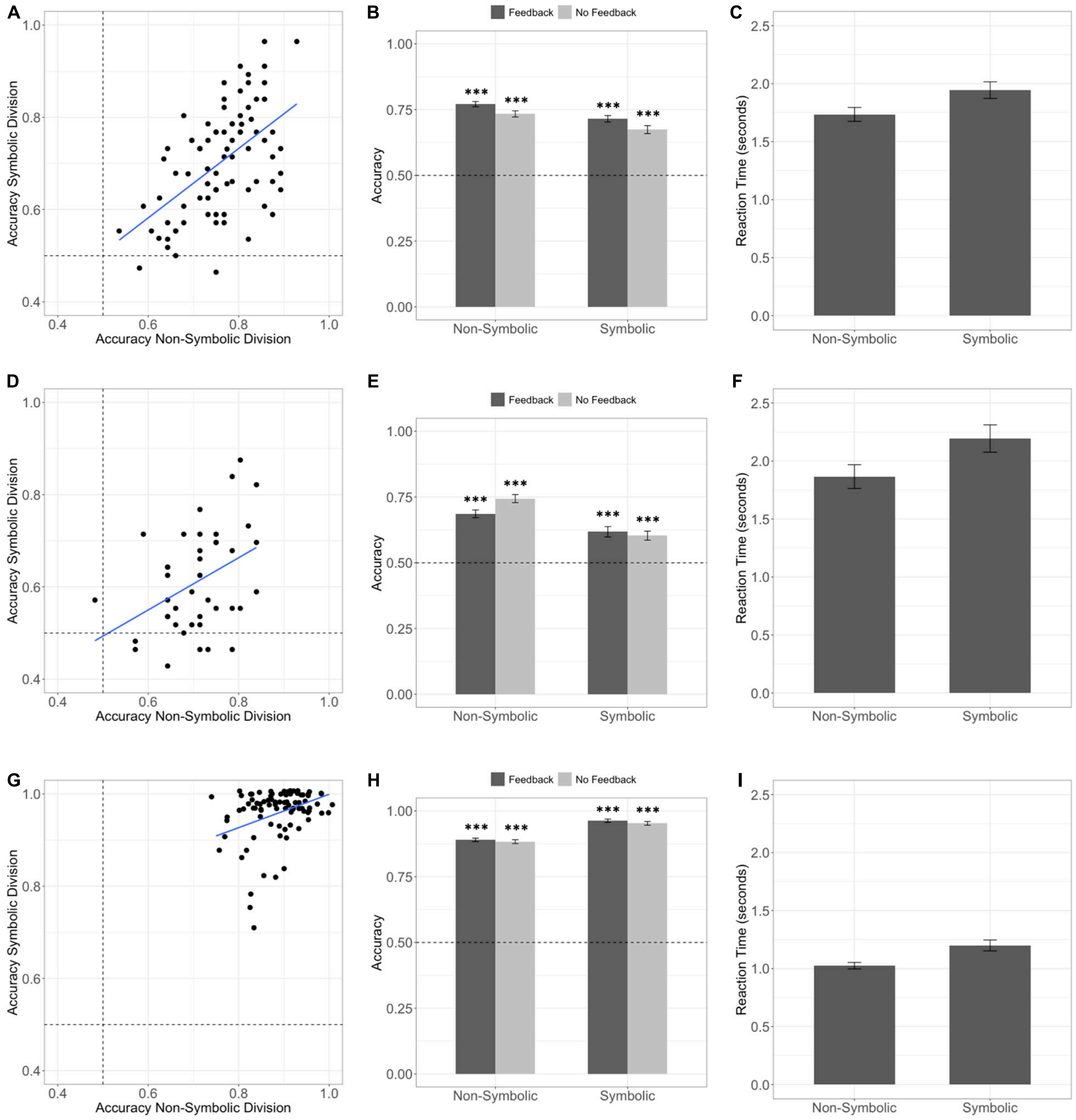


**FIGURE 2 |** Children and adults can successfully perform approximate division over non-symbolic and symbolic operands. The dotted line depicts chance performance. Error bars depict the standard error of the mean. ***$p < 0.001$ **(A)** Scatter plot depicting approximate non-symbolic and symbolic division overall accuracy in Experiment 1. **(B)** Children in Experiment 1 performed with above chance accuracy on both symbolic and non-symbolic division tasks on trials with feedback. Children also performed significantly above chance on both tasks during the no-feedback phase of the task where participants needed to divide with novel divisors. **(C)** Group level means of median reaction time on correct trials for the non-symbolic and symbolic tasks in Experiment 1. **(D)** Scatter plot depicting approximate non-symbolic and symbolic division overall accuracy in Experiment 2. **(E)** Children in Experiment 2 performed with above chance accuracy on both symbolic and non-symbolic division tasks on trials with feedback. Children also performed significantly above chance on both tasks during the no-feedback phase of the task where participants needed to divide with novel divisors. **(F)** Group level means of median reaction time on correct trials for the non-symbolic and symbolic tasks in Experiment 2. **(G)** Scatter plot depicting approximate non-symbolic and symbolic division overall accuracy in Experiment 3 with adult subjects. Points are jittered horizontally to reveal overlapping data points. **(H)** Adults in Experiment 3 performed with above chance accuracy on both symbolic and non-symbolic division tasks on trials with feedback and trials without feedback with novel divisors. **(I)** Group level means of median reaction time on correct trials for the non-symbolic and symbolic tasks in Experiment 3.

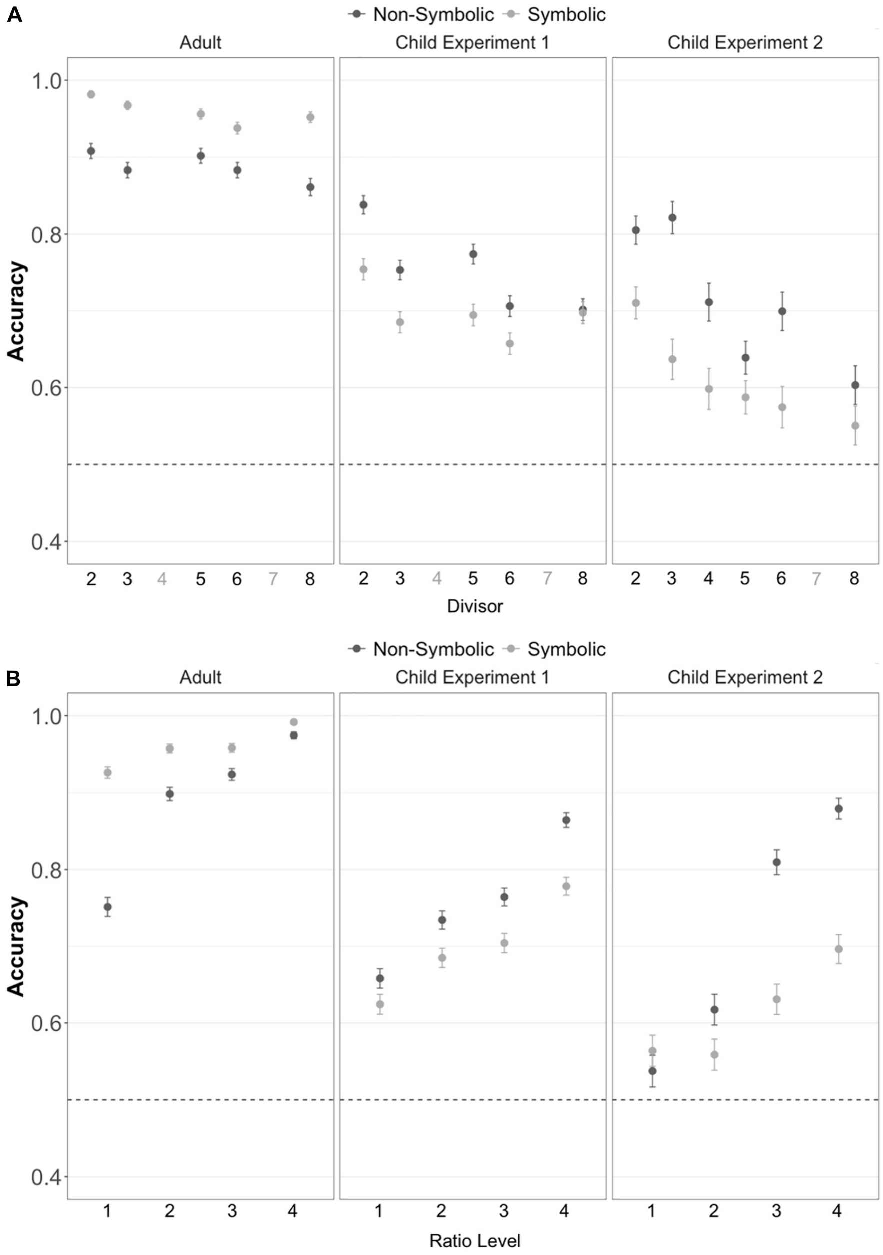


**FIGURE 3 |** Children and adults can successfully perform approximate division across **(A)** varied divisors and **(B)** ratios between the quotient and comparison quantity. Error bars represent the standard error of the mean. Divisors depicted in gray on the x-axis were not used in the experiment. Ratio level 1 ≈0.8, Ratio level 2 ≈0.6, Ratio level 3 ≈0.45, Ratio level 4 ≈0.35. The dotted line represents chance performance.

child = 1.81 s). There was no significant format by age interaction ($F_{1,174} = 0.194$, $p = 0.66$).

## Effect of Divisor on Division Accuracy

In Experiment 1, children performed significantly above chance (50%) on the non-symbolic division task independently for all divisors tested (divisor 2 $t_{88} = 25.8$, $p < 0.001$, $d = 2.7$; divisor 3 $t_{88} = 19.3$, $p < 0.001$, $d = 2.0$; divisor 5 $t_{88} = 22.9$, $p < 0.001$, $d = 2.4$; divisor 6 $t_{88} = 14.9$, $p < 0.001$, $d = 1.6$; divisor 8 $t_{88} = 13.5$, $p < 0.001$, $d = 1.4$). In Experiment 2, children also performed significantly above chance (50%) on the non-symbolic division task independently for all divisors tested (divisor 2 $t_{41} = 13.0$, $p < 0.001$, $d = 2.0$; divisor 3 $t_{41} = 11.9$, $p < 0.001$, $d = 1.8$; divisor 4 $t_{41} = 9.29$, $p < 0.001$, $d = 1.4$; divisor 5 $t_{41} = 7.32$, $p < 0.001$, $d = 1.3$; divisor 6 $t_{41} = 10.2$, $p < 0.001$, $d = 1.6$; divisor 8 $t_{41} = 4.81$, $p < 0.001$, $d = 0.74$). The same pattern of results was found in Experiment 3 with adult subjects (divisor 2 $t_{86} = 41.5$, $p < 0.001$, $d = 4.4$; divisor 3 $t_{86} = 38.8$, $p < 0.001$, $d = 4.2$; divisor 5 $t_{86} = 41.6$, $p < 0.001$, $d = 4.5$; divisor 6 $t_{86} = 37.8$, $p < 0.001$, $d = 4.0$; divisor 8 $t_{86} = 29.1$, $p < 0.001$, $d = 3.1$).

On the symbolic division tasks, children and adults also performed with above chance accuracy on all divisors independently, with the exception of the divisor 8 in Experiment 2 where participants performed marginally above chance (Experiment 1 divisor 2 $t_{88} = 14.1$, $p < 0.001$, $d = 1.5$; divisor 3 $t_{88} = 11.2$, $p < 0.001$, $d = 1.2$; divisor 5 $t_{88} = 13.5$, $p < 0.001$, $d = 1.4$; divisor 6 $t_{88} = 8.79$, $p < 0.001$, $d = 0.93$; divisor 8 $t_{88} = 11.6$, $p < 0.001$, $d = 1.2$; Experiment 2 divisor 2 $t_{41} = 7.87$, $p < 0.001$, $d = 1.2$; divisor 3 $t_{41} = 5.16$, $p < 0.001$, $d = 0.80$; divisor 4 $t_{41} = 3.71$, $p < 0.001$, $d = 0.57$; divisor 5 $t_{41} = 3.07$, $p = 0.004$, $d = 0.47$; divisor 6 $t_{41} = 2.44$, $p = 0.02$, $d = 0.38$; divisor 8 $t_{41} = 1.85$, $p = 0.07$, $d = 0.29$; Experiment 3 divisor 2 $t_{86} = 90.8$, $p < 0.001$, $d = 9.7$; divisor 3 $t_{86} = 74.2$, $p < 0.001$, $d = 8.0$; divisor 5 $t_{86} = 54.4$, $p < 0.001$, $d = 5.8$; divisor 6 $t_{86} = 46.8$, $p < 0.001$, $d = 5.0$; divisor 8 $t_{86} = 50.1$, $p < 0.001$, $d = 5.4$). These results indicate that above chance accuracy on the division tasks was not dependent on any single divisor. Thus participants were able to divide across multiple divisors flexibly.

### Effect of the Ratio Between the Target and Quotient on Division Accuracy

To test whether accuracy was dependent on the ratio between the quotient and the comparison target value, we fit a generalized linear mixed effects model (GLMM) following a binomial error distribution predicting whether each trial was correct with the ratio between the quotient and target as a fixed effect and a random effect of subject. Six models were fit, one for each of the three experiments and two task formats. For the non-symbolic division task this model indicated significant main effects of ratio for both adults and children in both Experiments 1 and 2 (**Figure 4** adult $\beta = 0.18$, $z = 3.28$, $p = 0.001$; child experiment 1 $\beta = 0.62$, $z = 13.8$, $p < 0.001$; child experiment 2 $\beta = 0.34$, $z = 5.95$, $p < 0.001$). For the symbolic division task this model also indicated significant main effects of ratio for both adults and children in both experiments (**Figure 4**; adult $\beta = 0.33$, $z = 3.37$, $p < 0.001$; child experiment 1 $\beta = 0.55$, $z = 13.5$, $p < 0.001$; child experiment 2 $\beta = 0.41$, $z = 7.67$, $p < 0.001$).

### Effect of Formal Division Knowledge on Approximate Division

There were 40 children who could not identify the division symbol in Experiment 1. Children who could not identify the division symbol successfully completed both the non-symbolic and symbolic division tasks with above chance accuracy (**Figure 5**; non-symbolic 74% $t_{39} = 16.5$, $p < 0.001$, $d = 2.6$; symbolic 67% $t_{39} = 11.2$, $p < 0.001$, $d = 1.8$). In Experiment 2, we replicated the finding that children who could not identify the division symbol nevertheless performed at above chance level on our approximate division tasks (non-symbolic 68% $t_{21} = 11.6$, $p < 0.001$, $d = 2.5$; symbolic 56% $t_{21} = 3.31$ $p = 0.003$, $d = 0.71$).

There were 51 children who could not solve any of the four simple symbolic division problems on our formal division test in Experiment 1 (for example, $6 \div 3 = ?$). These simple division problems were read aloud so that children who could not recognize the ÷ symbol could rely on the spoken word 'divided' to solve these problems. Children who could not solve symbolic division problems were nevertheless significantly above chance on both approximate division tasks (**Figure 5**; non-symbolic 74% $t_{50} = 19.0$, $p < 0.001$, $d = 2.7$; symbolic 67% $t_{50} = 11.7$ $p < 0.001$, $d = 1.6$). In Experiment 2 we replicated these results (non-symbolic 68% $t_{26} = 13.1$, $p < 0.001$, $d = 2.5$; symbolic 57% $t_{26} = 4.06$ $p < 0.001$, $d = 0.78$). Together, these data indicate that formal knowledge of division is not necessary to solve the approximate division tasks, in either symbolic or non-symbolic format.

For a comparison of approximate division accuracy between children who did and did not demonstrate formal division knowledge please see the **Supplementary Material**.

### Alternative Heuristic Analysis

Adults and children in Experiments 1 and 2 performed with above chance accuracy on both the non-symbolic and symbolic approximate division tasks when excluding the trials where subjects could compare the dividend and the target to get the correct answer, indicating participants did not rely on this heuristic to complete the division tasks (adults non-symbolic $t_{86} = 62.4$, $p < 0.001$; adults symbolic $t_{86} = 71.1$, $p < 0.001$; children experiment 1 non-symbolic 75% $t_{88} = 27.3$, $p < 0.001$; children experiment 1 symbolic 69% $t_{88} = 15.1$, $p < 0.001$; children experiment 2 non-symbolic $t_{41} = 16.2$, $p < 0.001$; children experiment 2 symbolic 60% $t_{41} = 6.32$, $p < 0.001$).

We next examined whether performance was consistent with participants creating a mental model of the median target value to solve the division task. The stimulus set in Experiment 2 with children was designed to rule out this heuristic, and as reported above, children performed with above chance accuracy on both the symbolic and non-symbolic division tasks in Experiment 2 (**Figures 2**, **3**). This indicates children can solve our approximate division task without use of a median target heuristic. The stimulus set used by adult subjects was not designed to rule out this heuristic, however, the accuracy rate was above chance on the subset of trials that could not be solved using the median target heuristic (non-symbolic 64%, $t_{86} = 6.64$, $p < 0.001$; symbolic 91%, $t_{86} = 28.4$, $p < 0.001$).

### Mediation Analysis

For children, ANS acuity was a significant predictor of a participant's score on the Key-Math-3 Numeration subtest (standardized $\beta = -0.34$, $p = 0.002$) and of accuracy on the non-symbolic division task (standardized $\beta = -0.29$, $p = 0.009$). ANS acuity continued to be a significant predictor of the score on the Numeration subtest after controlling for the mediator, non-symbolic division accuracy, however the strength of this relation was lessened (ANS acuity standardized $\beta = -0.27$, $p = 0.02$; non-symbolic division accuracy standardized $\beta = 0.23$, $p = 0.04$). We tested the significance of this reduction using a non-parametric bootstrap estimation with 5,000 simulations of the indirect effect using the "mediation" package in R (Tingley et al., 2014; indirect effect = −0.07, 95% CI = [−0.17 −0.01], $p = 0.03$). The direct effect was also significant, indicating a mediation (direct effect = −0.27, 95% CI = [−0.48.05], $p = 0.02$). The proportion mediated was 0.20 ($p = 0.03$, 95% CI = [0.01.69]). Thus, sharper ANS acuity was associated

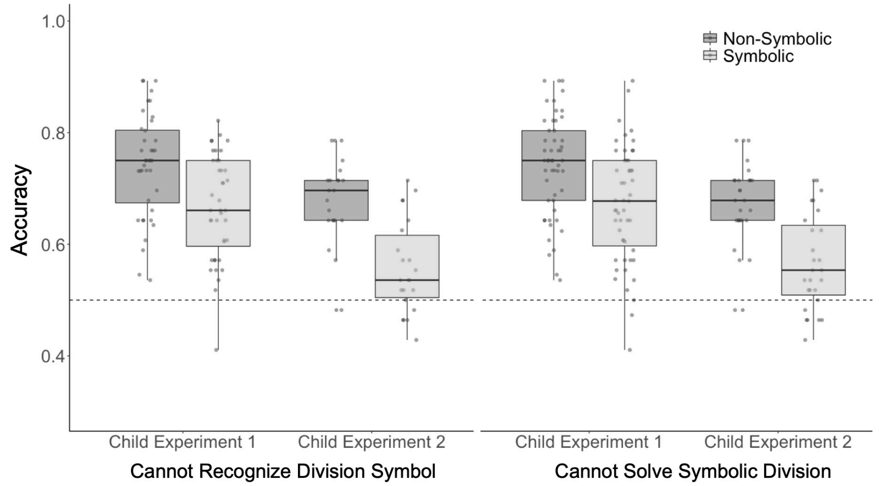


**FIGURE 4 |** Children who cannot recognize the division symbol (÷) or cannot solve symbolic division are able to perform non-symbolic and symbolic approximate division. Children's performance on the non-symbolic and symbolic division tasks broken down by their performance on the formal division test. The feedback and no feedback trials are combined for each task. Boxes indicate median and first and third quartiles. The whiskers indicate 1.5 multiplied by the interquartile range from the first and third quartiles. The dotted line represents chance performance. Each point reflects one participant's accuracy.

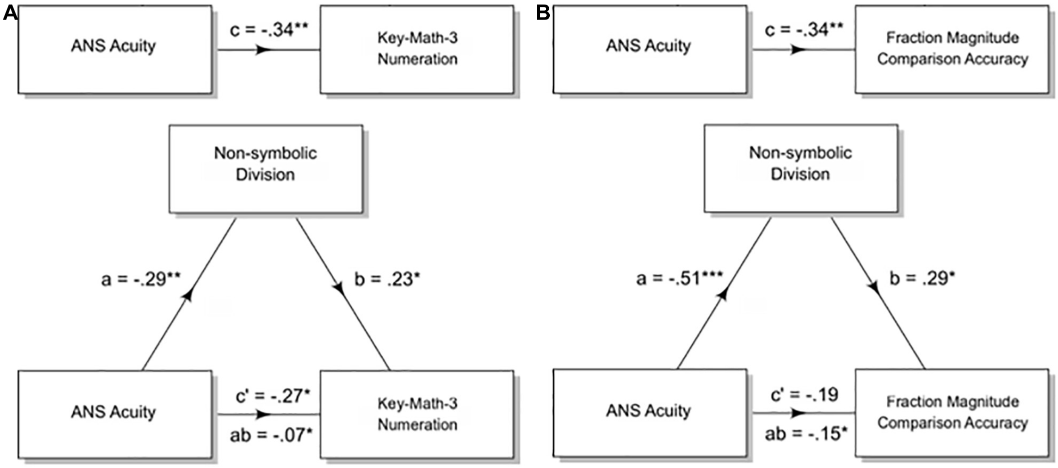


**FIGURE 5 |** Non-symbolic division mediates the effect of ANS acuity on symbolic math abilities in adults and children. Mediation analyses test for a significant indirect effect (the product of the standardized coefficients a and b) that accounts for some portion of the original direct effect (c). The remaining direct effect is represented as c'. The models in this figure test whether non-symbolic division performance mediates the relation between ANS acuity and a measure of formal math skills in children (Key-Math-3 Numeration) and adults (Fraction Magnitude Comparison). **(A)** Non-symbolic division accuracy partially mediates the relation between ANS acuity and a child's score on the Key-Math-3 Numeration section. Both the indirect (ab) and the direct path c' are significant. **(B)** Non-symbolic division accuracy fully mediates the relation between ANS acuity and accuracy on the fraction magnitude comparison test. The remaining direct effect (c') is no longer significant, while the indirect effect (ab) is significant as tested with a bootstrap estimate approach. * < 0.05, ** < 0.01, *** < 0.001.

with 0.07 standard deviations higher Key-Math-3 Numeration score as mediated through non-symbolic division accuracy (**Figure 5A**). This finding is in line with our hypothesis, however, when we partialed out the relation between the Woodcock-Johnson Reading Cluster and the Numeration subtest, ANS acuity was no longer significantly correlated with scores on the Numeration subtest (ANS acuity standardized $\beta = -0.18$, $p = 0.13$). This indicates that the relation between ANS acuity and the Numeration subtest is not specific to math skills, but rather to general academic performance.

For adults, ANS acuity was a significant predictor of accuracy on the fraction magnitude comparison test (standardized $\beta = -0.34$, $p = 0.003$) and of accuracy on the non-symbolic division task (standardized $\beta = -0.51$, $p < 0.001$). ANS acuity was no longer a significant predictor of accuracy on the fraction magnitude test after controlling for the mediator, non-symbolic division accuracy (ANS acuity standardized $\beta = -0.19$, $p = 0.12$; non-symbolic division accuracy standardized $\beta = 0.29$, $p = 0.02$). Non-symbolic division accuracy mediated the relation between ANS acuity and accuracy on the fraction magnitude comparison test. The indirect effect was significant when tested with a bootstrap estimation approach with 5,000 simulations (indirect effect = −0.15, 95% CI = [−0.28, −0.03], $p$ = 0.01). The direct effect was not significant, indicating a mediation (direct effect = −0.19, 95% CI = [-0.4 6, 0.07], $p$ = 0.16). The proportion mediated was 0.43 (95% CI = [0.08, 1.5]), $p$ = 0.02. Thus, sharper ANS acuity was associated with 0.15 standard deviations higher fraction magnitude comparison accuracy as mediated through non-symbolic division accuracy (**Figure 5B**). The indirect effect remained significant when controlling for the relation between a participants' vocabulary score and fraction magnitude comparison accuracy (indirect effect = −0.15, 95% CI = [−0.29, −0.03], $p$ = 0.009, direct effect = −0.19, 95% CI = [−0.46, 0.08], $p$ = 0.17, proportion mediated = 0.44, CI = [08, 1.6], $p$ = 0.02).

## DISCUSSION

The current experiments are the first to demonstrate that elementary school children and adults can approximately divide over both non-symbolic arrays and numerals. Our task required a true non-symbolic, approximate division computation that integrates the relations between a dividend, divisor, and quotient. Successful completion of the two division tasks was not dependent on formal knowledge of division. Children who could not recognize the division symbol nor solve simple division problems were nevertheless successful at performing non-symbolic division, and more surprisingly, they were also able to complete the division task when the dividend and target comparison number were represented symbolically with Arabic numerals. These findings highlight the depth of intuitive math knowledge that children possess before formal education.

We found that task format differentially impacted children and adult's division accuracy. Whereas adults were significantly better at the symbolic compared to the non-symbolic division task, children were significantly better at the non-symbolic task. The timing of when symbols facilitate more accurate arithmetic calculations may mark an important conceptual milestone in mathematical development. One possibility is that making a switch to more accurate computation within the symbolic number system earlier in development is a better scaffold for increasingly complex computation. Alternatively, continuing to root a mathematical operation in its underlying concrete representation may be a better foundation for understanding complex math concepts. Future research can test whether the timing of this transition is longitudinally predictive of later math achievement, how other characteristics of the learner impact a child's non-symbolic and symbolic arithmetic accuracy over time, and whether instructional practices can impact the timing of this transition.

One exciting implication of these findings is the possibility to introduce the division concept early in math education via large number approximate calculation using both non-symbolic quantities and numerals. Future research can explore whether an explicit linking between non-symbolic division and division using numerals can aid formal division understanding. The theoretical framework of concreteness fading may be a particularly useful method for implementing such an intervention (Fyfe et al., 2015; Fyfe and Nathan, 2018). A progression from practice with approximate non-symbolic division, to approximate symbolic division to exact symbolic division may be a way to link children's intuitions about division to formal division knowledge. Another theoretical framework that has shown promise in linking intuitive math knowledge to symbolic math learning is implicit analogical transfer (Sidney and Thompson, 2019). Under this framework, 'warming up' activities are used to promote successful analogical transfer between current and future knowledge. In the context of the current findings, intuitive division tasks could be used to activate children's intuitive understanding of a division topic before a lesson in formal division. Under both frameworks, grounding abstract arithmetic concepts in children's intuitive understanding of arithmetic may boost children's conceptual understanding of arithmetic operations and their confidence in their own skill to perform such calculations. Incorporating numerical symbols within an intuitive division context may function as a pedagogical bridge for developing a deeper and more robust division concept in children, which ultimately could promote stronger symbolic, exact division calculation skill.

The current experiments also examined whether approximate division could be a mechanism of the known relation between ANS acuity and symbolic mathematics (Schneider et al., 2016). Two pieces of evidence strongly suggest that the ability to non-symbolically and approximately divide is grounded in the Approximate Number System. First, accuracy for all subjects was modulated by the ratio between the target and quotient in both non-symbolic and symbolic format, indicating use of an approximate strategy when making their choice. Ratio dependent discrimination of quantity is a hallmark of the ANS (Feigenson et al., 2004). Second, accuracy on both division tasks was significantly correlated with participant's ANS acuity as independently measured with a dot comparison task. The division operation joins a growing number of mathematical operations that can be represented using the ANS before formal math education including addition, subtraction, scaling, ratio comparison, and solving for X (Barth et al., 2005; McCrink and Wynn, 2007; Kibbe and Feigenson, 2015; McCrink et al., 2016).

With this evidence that approximate division is rooted in the ANS, we then tested the second aspect of our hypothesis – that approximate division ability is correlated with symbolic math skill. In line with our hypothesis, non-symbolic division mediated the relation between ANS acuity and symbolic math in both children and adults. Sharper ANS acuity may facilitate greater accuracy in a student's conceptual model of a division operation, and this conceptual model may function as a scaffold for formal symbolic computation. Thus, the mechanism for the established link between ANS acuity and symbolic math ability may be rooted in the computational abilities allowed by the ANS, and not in the acuity of the ANS *per se*. Having a strong mental model of what it means to divide (or engage in other operations such as subtraction or multiplication) may in turn create a strong foundation for the learning of abstract mathematical concepts. The significant mediation effect in adults suggests that adults continue to use approximate mental models to calculate, even once they have knowledge of exact calculation techniques.

Unexpectedly, in children, while non-symbolic division ability was a mediator of the relation between ANS acuity and Key-Math-3 performance, this mediation effect was no longer significant when controlling for the correlation between children's scores on the Woodcock-Johnson Reading Cluster and their performance on the Key-Math-3 Numeration test. When controlling for reading ability, ANS acuity was no longer correlated with scores on the Key-Math-3. It is possible this result is due to shared correlations between an unknown additional skill, such as inhibition or executive function (Fuhs and McNeil, 2013), and our measures of math, reading, and ANS acuity. However, we do not interpret this finding as evidence that ANS acuity is meaningfully related to reading skill, but rather as evidence of the strong correlation between math and reading skills is typical in children of this age that is attributed to extrinsic academic factors (Wang et al., 2015; Cantin et al., 2016). In the current study, we found a correlation of r = 0.62 between Key-Math-3 and Woodcock Johnson Reading Cluster scores after controlling for age of the participants. This strong correlation between math and reading scores left little variance to partition in the mediation model. In adults, partialing out vocabulary skill from fraction magnitude accuracy ($r = 0.02$) did not impact our mediation effect. Thus, this unexpected finding is most likely due to measurement rather than theoretical importance of reading skill in the relation between the ANS and symbolic math. Indeed, we do not find this pattern of results in the adult experiment.

In conclusion, our study highlights that children have strong intuitive math abilities before they begin formal math education. We found that children are remarkably good at dividing large numbers regardless of whether they were presented non-symbolically or symbolically, and this ability is not limited to simply halving or quartering. Children's extraordinary success at approximate division with large quantities suggests the possibility that introducing non-symbolic arithmetic calculation early in math education may be beneficial for formal arithmetic learning. We hope that our findings inspire future studies to test the efficacy of math instruction that emphasizes grounding highly abstract mathematical concepts in intuitive math abilities.

## DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://osf.io/g5y27/?view_only=0a2ab8862d9d4e95b2426cb11657e78d.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by University of Pennsylvania Human Subjects Electronic Research Application (HS-ERA). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

ES, HZ, ND, and EB designed the study. ES and HZ wrote the code to run the experimental tasks and collected the data. ES analyzed the data and wrote the first draft of the manuscript. All authors edited the final manuscript.

## FUNDING

The research reported here was supported by a National Science Foundation Division of Research on Learning Award #1660973 to EB, and a Ruth L. Kirschstein Predoctoral Individual National Research Service Award F31 HD095579-01 to ES.

## ACKNOWLEDGMENTS

We thank Luis Rosario, Maria Arevalo-Gonzalez, Wendy Arce, Andreas Nolan, Justin Roberson, Nuwar Ahmed, and Bonnie Zuckerman for help with data collection and management, and Stephanie Bugden for helpful discussions about the manuscript. We also thank the parents and after school programs that participated in this research.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnhum.2022.752190/full#supplementary-material

## REFERENCES

Barth, H., Baron, A., Spelke, E., and Carey, S. (2009). Children's multiplicative transformations of discrete and continuous quantities. *J. Exp. Child Psychol.* 103, 441–454. doi: 10.1016/j.jecp.2009.01.014

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., and Spelke, E. (2006). Non-symbolic 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., and Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proc. Nat. Acad. Sci. USA* 102, 14116–14121. doi: 10.1073/pnas.0505512102

Blake, P. R., and McAuliffe, K. (2011). "I had so much it didn't seem fair": Eight-year-olds reject two forms of inequity. *Cognition* 120, 215–224. doi: 10.1016/j.cognition.2011.04.006

Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436.

Campbell, J. I. (1997). On the relation between skilled performance of simple division and multiplication. *J. Exp. Psychol. Learn. Memory Cogn.* 23:1140. doi: 10.1037//0278-7393.23.5.1140

Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., and Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *J. Exp. Child Psychol.* 146, 66–78. doi: 10.1016/j.jecp.2016.01.014

Cantlon, J. F., Merritt, D. J., and Brannon, E. M. (2015). Monkeys display classic signatures of human symbolic arithmetic. *Animal Cogn.* 2015:9425. doi: 10.1007/s10071-015-0942-5

Carpenter, T. P., Ansell, E., Franke, M. L., Fennema, E., and Weisbeck, L. (1993). Models of problem solving: a study of kindergarten children's problem-solving processes. *J. Res. Math. Edu.* 24:428. doi: 10.2307/749152

Chen, Q., and Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychol.* 148, 163–172. doi: 10.1016/j.actpsy.2014.01.016

Connolly, A. J. (2007). *Keymath-3 Diagnostic Assessment: Manual Forms A and B*. Minneapolis, MN: SAGE Publications.

Correa, J., Bryant, P., and Nunes, T. (1998). Young children's understanding of division: the relationship between division terms in a noncomputational Task. *J. Edu. Psychol.* 90, 321–329. doi: 10.1037/0022-0663.90.2.321

DeWind, N. K., Adams, G. K., Platt, M. L., and Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition* 142, 247–265. doi: 10.1016/j.cognition.2015.05.016

DeWind, N. K., and Brannon, E. M. (2016). Significant inter-test reliability across approximate number system assessments. *Front. Psychol.* 7:310. doi: 10.3389/fpsyg.2016.00310

Ekstrom, R. B., French, J. W., Harman, H. H., and Dermen, D. (1976). *Manual for Kit of Factor-Referenced Cognitive Tests*. Princeton NJ: Education Testing Service.

Falk, R., Yudilevich-Assouline, P., and Elstein, A. (2012). Children's concept of probability as inferred from their binary choices—Revisited. *Edu. Stud. Math.* 81, 207–233. doi: 10.1007/s10649-012-9402-1

Fazio, L. K., Bailey, D. H., Thompson, C. A., and Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *J. Exp. Child Psychol.* 123, 53–72. doi: 10.1016/j.jecp.2014.01.013

Fazio, L. K., DeWolf, M., and Siegler, R. S. (2015). Strategy use and strategy choice in fraction magnitude comparison. *J. Exp. Psychol. Learn. Mem. Cogn.* 2015:153. doi: 10.1037/xlm0000153

Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002

Fuhs, M. W., and McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. *Dev. Sci.* 16, 136–148. doi: 10.1111/desc.12013

Fyfe, E. R., McNeil, N. M., and Borjas, S. (2015). Benefits of "concreteness fading" for children's mathematics understanding. *Learn. Instruct.* 35, 104–120. doi: 10.1016/j.learninstruc.2014.10.004

Fyfe, E. R., and Nathan, M. J. (2018). Making "concreteness fading" more concrete as a theory of instruction for promoting transfer. *Edu. Rev.* 2018, 1–20. doi: 10.1080/00131911.2018.1424116

Gilmore, C. K., McCarthy, S. E., and Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature* 447, 589–591. doi: 10.1038/nature05850

Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Dev. Psychol.* 48, 1229–1241. doi: 10.1037/a0027433

Hamamouche, K., Chernyak, N., and Cordes, S. (2020). Sharing scenarios facilitate division performance in preschoolers. *Cogn. Dev.* 56:100954. doi: 10.1016/j.cogdev.2020.100954

Ischebeck, A., Zamarian, L., Schocke, M., and Delazer, M. (2009). Flexible transfer of knowledge in mental arithmetic—An fMRI study. *NeuroImage* 44, 1103–1112. doi: 10.1016/j.neuroimage.2008.10.025

Jitendra, A. K., and Hoff, K. (1996). The effects of schema-based instruction on the mathematical word-problem-solving performance of students with learning disabilities. *J. Learn. Disabil.* 29, 422–431. doi: 10.1177/002221949602900410

Kalra, P. B., Binzak, J. V., Matthews, P. G., and Hubbard, E. M. (2020). Symbolic fractions elicit an analog magnitude representation in school-age children. *J. Exp. Child Psychol.* 195:104844. doi: 10.1016/j.jecp.2020.104844

Kibbe, M. M., and Feigenson, L. (2015). Young children 'solve for *x* 'using the Approximate Number System. *Dev. Sci.* 18, 38–49. doi: 10.1111/desc.12177

Kibbe, M. M., and Feigenson, L. (2017). A dissociation between small and large numbers in young children's ability to "solve for *x*" in non-symbolic math problems. *Cognition* 160, 82–90. doi: 10.1016/j.cognition.2016.12.006

Klein, E., Moeller, K., Dressel, K., Domahs, F., Wood, G., Willmes, K., et al. (2010). To carry or not to carry — Is this the question? Disentangling the carry effect in multi-digit addition. *Acta Psychol.* 135, 67–76. doi: 10.1016/j.actpsy.2010.06.002

Kleiner, M., Brainard, D., and Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception* 36:14.

Knops, A., Viarouge, A., and Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. *Atten. Percep. Psychophys.* 71, 803–821. doi: 10.3758/APP.71.4.803

Lindskog, M., Winman, A., Juslin, P., and Poom, L. (2013). Measuring acuity of the approximate number system reliably and validly: The evaluation of an adaptive test procedure. *Front. Psychol.* 4:510. doi: 10.3389/fpsyg.2013.00510

Matthews, P. G., Lewis, M. R., and Hubbard, E. M. (2016). Individual differences in nonsymbolic ratio processing predict symbolic math performance. *Psychol. Sci.* 27, 191–202. doi: 10.1177/0956797615617799

Mauro, D. G., LeFevre, J.-A., and Morris, J. (2003). Effects of problem format on division and multiplication performance: division facts are mediated via multiplication-based representations. *J. Exp. Psychol. Learn. Memory Cogn.* 29, 163–170. doi: 10.1037/0278-7393.29.2.163

McCrink, K., Shafto, P., and Barth, H. (2016). The relationship between non-symbolic multiplication and division in childhood. *Quart. J. Exp. Psychol.* 2016, 1–17. doi: 10.1080/17470218.2016.1151060

McCrink, K., and Spelke, E. S. (2010). Core multiplication in childhood. *Cognition* 116, 204–216. doi: 10.1016/j.cognition.2010.05.003

McCrink, K., and Spelke, E. S. (2016). Non-symbolic division in childhood. *J. Exp. Child Psychol.* 142, 66–82. doi: 10.1016/j.jecp.2015.09.015

McCrink, K., Spelke, E. S., Dehaene, S., and Pica, P. (2013). Non-symbolic halving in an Amazonian indigene group. *Dev. Sci.* 16, 451–462. doi: 10.1111/desc.12037

McCrink, K., and Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychol. Sci.* 15, 776–781. doi: 10.1111/j.0956-7976.2004.00755.x

McCrink, K., and Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychol. Sci.* 18, 740–745. doi: 10.1111/j.1467-9280.2007.01969.x

McNeil, N. M., Fuhs, M. W., Keultjes, M. C., and Gibson, M. H. (2011). Influences of problem format and SES on preschoolers' understanding of approximate addition. *Cogn. Dev.* 26, 57–71. doi: 10.1016/j.cogdev.2010.08.010

National Governors Association Center for Best Practices, and Council of Chief State School Officers (2019). *Common Core State Standards for Mathematics*. Washington, DC: National Governors Association Center for Best Practices & Council of Chief State School Officers.

Odic, D., Le Corre, M., and Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition* 138, 102–121. doi: 10.1016/j.cognition.2015.01.008

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442.

Pica, P., Lemer, C., Izard, V., and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503. doi: 10.1126/science.1102085

Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., et al. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One* 9:e111155. doi: 10.1371/journal.pone.0111155

Riley, M. (1984). *Schema Knowledge Structures for Representing and Understanding Arithmetic Story Problems*. Washington, DC: IES.

Rosenberg-Lee, M., Chang, T. T., Young, C. B., Wu, S., and Menon, V. (2011). Functional dissociations between four basic arithmetic operations in the human posterior parietal cortex: A cytoarchitectonic mapping study. *Neuropsychologia* 49, 2592–2608. doi: 10.1016/j.neuropsychologia.2011.04.035

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., et al. (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Dev. Sci.* 20:e12372. doi: 10.1111/desc.12372

Shaw, A., and Olson, K. R. (2012). Children discard a resource to avoid inequity. *J. Exp. Psychol. General* 141, 382–395. doi: 10.1037/a0025907

Sheskin, M., Nadal, A., Croom, A., Mayer, T., Nissel, J., and Bloom, P. (2016). Some equalities are more equal than others: quality equality emerges later than numerical equality. *Child Dev.* 87, 1520–1528. doi: 10.1111/cdev.12544

Sidney, P. G., and Thompson, C. A. (2019). Implicit analogies in learning: supporting transfer by warming up. *Curr. Direct. Psychol. Sci.* 28, 619–625. doi: 10.1177/0963721419870801

Starr, A., Roberts, R. C., and Brannon, E. M. (2016). Two potential mechanisms underlying the link between approximate number representations and symbolic math in preschool children. *Proc. Cogn. Sci. Soc.* 2015:30408.

Szkudlarek, E., and Brannon, E. M. (2021). First and second graders successfully reason about ratios with both dot arrays and arabic numerals. *Child Dev.* 2021:13470. doi: 10.1111/cdev.13470

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). mediation: R package for causal mediation analysis. *J. Stat. Softw.* 59, 1–38.

Wang, Z., Soden, B., Deater-Deckard, K., Lukowski, S. L., Schenker, V. J., Willcutt, E. G., et al. (2015). Development in reading and math in children from different SES backgrounds: The moderating role of child temperament. *Dev. Sci.* 2015:12380. doi: 10.1111/desc.12380

Xenidou-Dervou, I., van Lieshout, E. C. D. M., and van der Schoot, M. (2014). Working memory in nonsymbolic approximate arithmetic processing: a dual-task study with preschoolers. *Cogn. Sci.* 38, 101–127. doi: 10.1111/cogs.12053

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

# Advantages of publishing in Frontiers

**OPEN ACCESS**

Articles are free to read
for greatest visibility
and readership

**FAST PUBLICATION**

Around 90 days
from submission
to decision

**HIGH QUALITY PEER-REVIEW**

Rigorous, collaborative,
and constructive
peer-review

**TRANSPARENT PEER-REVIEW**

Editors and reviewers
acknowledged by name
on published articles

**Frontiers**
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

**Visit us:** www.frontiersin.org
**Contact us:** frontiersin.org/about/contact

**REPRODUCIBILITY OF RESEARCH**

Support open data
and methods to enhance
research reproducibility

**DIGITAL PUBLISHING**

Articles designed
for optimal readership
across devices

**FOLLOW US**

@frontiersin

**IMPACT METRICS**

Advanced article metrics
track visibility across
digital media

**EXTENSIVE PROMOTION**

Marketing
and promotion
of impactful research

**LOOP RESEARCH NETWORK**

Our network
increases your
article's readership