DOCTORAL THESIS

Innate Predispositions in Numerosity Cognition: A Cross-Cultural, Developmental, and Comparative Perspective

Ph.D. candidate: Elena Eccher

Supervisor: Prof. Giorgio Vallortigara
Co-supervisor: Prof. Manuela Piazza

Doctoral Programme in *Cognitive and Brain Sciences*

CIMeC – Center for Mind/Brain Science

University of Trento

XXXVI Cycle
2020 – 2024

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the Doctorate in Cognitive and Brain Sciences, University of Trento.

I hereby declare that the contents and organization of this dissertation constitute original work and do not compromise in any way the rights of third parties, including those relating to the security of personal data.

“It is through science that we prove,
but through intuition that we discover”

(Henri Poincare)

Abstract

Numerical cognition and its underlying mechanisms have been extensively investigated because of their crucial role in the survival and reproduction of biological organisms. Evidence shows that animals possess a "*number sense*" to estimate quantities without counting, which assists behaviors such as foraging and social interaction. Thus, contrary to traditional beliefs that link numerical concepts only to language and culture, a growing body of research demonstrates that non-symbolic numerical cognition is shared across species, ages, and cultures. This number sense relies on the Approximate Number System (ANS), which is a fundamental cognitive mechanism that enables the estimation and comparison of quantities without the need for language or formal numerical symbols, such as digits or counting. The ANS supports the ability to make approximate judgments about the number of objects or events in a set and to perform basic arithmetic operations, such as addition and subtraction, although with inherent imprecision. Research has shown that many species, from primates to invertebrates, possess these abilities. For example, primates can perform numerical tasks and display ratio-dependent accuracy similar to humans, while domestic chicks show proto-arithmetic abilities comparable to those of human infants'. These findings suggest that the ANS might represent an evolutionarily ancient capability for processing numerical information, supporting the ability to navigate numerical tasks in natural environments. However, ongoing debates continue about the balance between innate predispositions and cultural influences on the ANS. In this thesis, we aimed to investigate key features of the ANS across different populations and species to deepen our understanding of the possible innate components of this system.

In the first study, we investigated the Spatial-Numerical Association (SNA) phenomenon. Since its discovery, SNA has been largely attributed to cultural factors, such as reading and writing habits, while more recent research indicates that SNA is also manifested in individuals lacking cultural exposure, including non-human animals and pre-verbal children, suggesting that SNA may be influenced by both evolutionary and cultural factors. Thus we explored whether different SNA behaviors could emerge as a function of culture, age, and stimuli presentation. Results from Himba adults, Italian preschoolers, and Italian adults revealed different SNA patterns in explicit and implicit tasks. Specifically, non-symbolic numerosities elicited a left-to-right SNA across all the populations in an implicit task, suggesting that SNAs are not solely the product of cultural influence. Additionally, we further investigated how symbolic and non-symbolic numerosities elicit asymmetrical SNA effects in literate adults. Our findings suggest that symbolic and non-symbolic numerosities might trigger distinct hemispheric activations thus affecting the directionality of SNAs.

In the second study, we aimed to expand our understanding of numerical cognition in non-human animals (i.e., domestic chicks, *Gallus gallus*) by extending research from the visual domain to the auditory domain. Our findings demonstrated that the predispositions observed with visual stimuli are also present with acoustic stimuli, highlighting the consistency of the ANS across different sensory modalities. The role of potential confounding variables and the need for further research are discussed.

In the third study, we examined another key aspect of the ANS, namely cross-modal numerical transfer. For both humans and non-human animals, the ability to assess and compare numerical quantities across different sensory modalities has been reported. Moreover, it has been demonstrated that multisensory information can improve numerical accuracy in children. Therefore, cross-modal numerical ability may offer evolutionary advantages, providing additional support for the idea that the ANS is shared and maintained across different species. We investigated this ability specifically in domestic chicks by presenting them with auditory and visual stimuli, both in spontaneous choice and imprinting paradigms. While our preliminary results may be not yet conclusive, they suggest the potential for cross-modal numerical abilities in chicks. Limitations of the current studies and proposals for further research are discussed.

In summary, this thesis advances our understanding of numerical cognition by exploring SNA and numerosity perception across populations not subjected to cultural inputs. The findings point to a complex interaction between innate mechanisms and cultural influences, with implications for both human and

Acknowledgment

I would like to extend my deepest gratitude to my supervisor, Professor Giorgio Vallortigara, for his invaluable support and guidance throughout my PhD years. Your dedication and commitment to research have been truly inspiring. From you, I have learned that nothing is too small, too simple, or too trivial to be worth studying. Thank you for challenging me and giving me the opportunity to fall in love with this work.

I also wish to deeply thank my co-supervisor, Professor Manuela Piazza. I am profoundly grateful for the opportunity to work under your supervision. Your insightful feedback, patience, and expertise have been instrumental in shaping this research. Thank you for providing me not only academic mentorship but also personal encouragement that greatly contributed to my growth as a researcher.

Much gratitude also goes to my oversight committee members, Professor Manuela Piazza and Professor Eugenio Parise, for offering crucial feedback and constructive recommendations on my research.

A special mention to Dr. Orsola Rosa Salva and Professor Cinzia Chiandetti for generously sharing their invaluable expertise on chick studies. My research would not have been possible without your support and knowledge. A similar thank goes to Dr. Marco Buiatti, who generously shared his valuable time and expertise with me for parallel projects throughout my PhD.

Thanks must also go to the amazing collaborators Dr. Serge Caparos, Dr. Mathilde Josserand, and Dr. Esther Boissin for their invaluable contributions to organizing and conducting the field mission to test the Himba people. I would also like to extend my gratitude to the Himba community, the children and their families, and the local (Nigerian and Italian) institutions and collaborators who played an essential role in the success of this project.

In addition, I would like to thank all the students who assisted in collecting the data for this thesis. Working with you as your supervisor has enriched me not only on a scientific level but also on a deeply human one.

Finally, I would like to thank each and every one of my lab mates, who have enriched me through every conversation, formal or informal. You have been an integral and necessary part of my journey. Thank you.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1. The Approximate Number System and its basic properties

Numbers are a crucial aspect of everyday life, not only for human adults frequently performing explicit arithmetic tasks (e.g. when doing groceries, counting totals and changes) but also for non-human animals. For instance, “How many predators?” or “How much food?” are number-based pieces of information that significantly influence animal behavior. In the past, it has long been thought that the development of numerical concepts was linked to the use of language, with the idea that mathematical ability originated as an "abstraction" from human language (Bloom, 1994; Chomsky, 2006). However, a growing body of literature has demonstrated that non-symbolic numerical cognition is not a human prerogative (see Butterworth et al., 2018; Giurfa, 2019; Nieder, 2019; Vallortigara, 2015, 2017, 2018). The idea that also non-human animals need to comprehend and process numerical information is grounded in its highly ecologically adaptive value (Haun et al., 2010; Nieder, 2020). On daily basis animals need to make choices for their survival in the wild, from hunting/foraging strategies (Hanus & Call, 2007; Panteleeva et al., 2013; Watts & Mitani, 2002) to dealing with social contexts and mating activities (Benson-Amram et al., 2018; Grinnell et al., 1995; Heinsohn, 1997; Kitchen, 2004; Lyon, 2003; McComb et al., 1994; Nieder, 2019; Wilson et al., 2001). These decisions are not made on precise counting processes but are rather based on an approximate estimation of quantities. This ability to access numerical information about a set of objects has been commonly referred to as “*Number Sense*”, and it refers to skills such as recognizing quantities without counting, understanding numerical magnitude, and performing simple arithmetic operations (Dehaene, 2011). Thus, when talking about *numbers sense* and numerosity (as in opposition to “numbers”), cognitive scientists normally want to stress the non-symbolic nature of the information taken into account. It has been hypothesized that this basic cognitive mechanism for the representation of numerosities is one of the “Core Knowledge system” that is available since birth, routed in our evolutionary history rather than resulting from learning processes (Spelke, 2000). Indeed, evidence of numerical competencies is present not only in human societies lacking formal education (O’Shaughnessy et al., 2023; Pica et al., 2004) but also at pre-scholar stages of development (Di

Giorgio et al., 2019; Izard et al., 2009; Lipton & Spelke, 2004). In the last decades, it has been proposed that numerosity perception mostly relies on two different core systems: the Object Tracking System (OTS) and the Approximate Number System (ANS). The first accounts for tracking small sets of objects (up to 3-4 elements), while the other yields a noisy representation of (both small and large) numerosities. Both systems seem to depend on an evolutionary heritage, rather than emerge through individual learning or cultural transmission, as they appear both in humans and non-human animals (Feigenson et al., 2004). This is why, in this framework, studies on human adults of different cultures, pre-schoolers, and non-human animals can easily allow parallelisms between their findings. For the aim of this elaborate, I will focus only on the Approximate Number System.

The ANS has been described as the system that underlies the capacity to estimate numerosity (Butterworth et al., 2018; Dehaene, 2011; Gallistel and Gelman, 2000, 1992; Vallortigara, 2017). It makes it possible to discriminate between sets of objects with different numerosity and to manipulate numerosities to the point of performing arithmetical operations (Brannon, 2002; McCrink & Wynn, 2004, 2009; Wynn, 1992). This system does not appear to have an upper boundary in numerosity perception (Cordes et al., 2001), but rather to be dependent on the numerical ratio of the comparisons. Indeed, as a consequence of its inexactness the ANS cannot distinguish between two quantities that are too close, and several studies have demonstrated that both accuracy and reaction times in numerosity discrimination follow Weber's law, both in human (Sasanguie et al., 2017) and non-human animals (Ditz & Nieder, 2016; Skorupski et al., 2018). The Weber law states that, given an initial stimulus (a set of multiple items), the minimum increment in numerosity necessary to perceive a change in the stimulus is a constant proportion of its numerosity. In other words, the ability to detect a difference in numerosity between two stimuli depends on their ratio, not their absolute difference (see Gallistel & Gelman, 1992). As a consequence of this ratio-dependent limit, the ANS has two main signature effects: the *distance effect* and the *size effect.* The distance effect refers to the difficulty in distinguishing between quantities that are close together compared to those that are farther apart (e.g. distinguishing a set of 5 from a set of 6 dots is harder than distinguishing a set of 5 from a set of 11), while the size effect makes it harder, given the same numerical distance, to differentiate between sets with a large number of items compared to sets with fewer items (e.g. distinguish a set of 8 from a set of 12 dots is easier than distinguish a set of 30 from one of 34 dots, Nieder, 2019). The ability to discriminate two quantities depending on their ratio does not attain full acuity until quite late in development (Halberda & Feigenson, 2008). As with other cognitive functions, the ANS develops with age. At birth, humans can discriminate a set of stimuli of different numerosities only starting from a 1:3 ratio (Izard et al., 2009), while at 6 months a 1:2 ratio is sufficient but infants fail if the arrays are presented with a 2:3 ratio (Xu & Spelke, 2000). At 10 months old instead, babies succeed even at the 2:3 ratio (Xu & Arriaga, 2007). The trajectory of the ANS keeps increasing in acuity during development (Halberda & Feigenson, 2008) and at adultage, on average, it is possible to discriminate ratios differing by about 7:8 (Piazza, 2011; Pica et al., 2004; Figure 1.1).

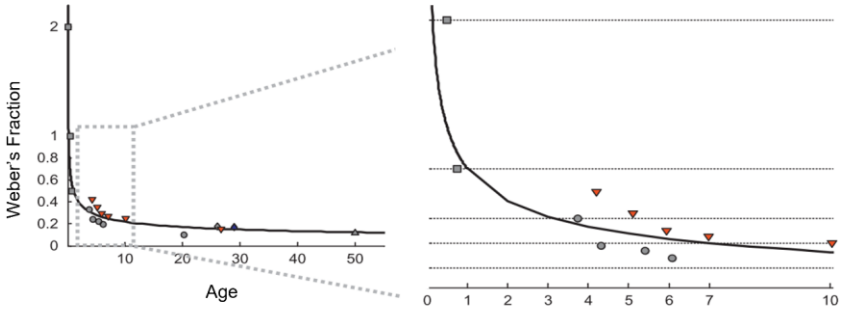


Figure 1.1: Development of ASN system with age
Acuity of the ANS develops with age as demonstrated by the power law fit (black curve) across several studies at different ages (see Piazza., 2011 for details). *Modified from Piazza et al., 2011*

While in the traditional view, the numerical distance and size effects are often attributed to a single underlying ratio effect, as both symbolic and non-symbolic numbers engage the ANS, some recent findings challenge this view as they found dissociability between the two effects (Hohol et al., 2020; Krajcsi et al., 2016, 2022). For example, it has been demonstrated that in a numerosity comparison task, there was a strong correlation between distance and size effects only for non-symbolic numerosities (i.e. dots) but not for symbolic numerosities (i.e. Indo-Arabic numbers). The authors suggested that symbolic number comparison may not rely on the ANS, but rather on an alternative representation, such as the discrete semantic system (DSS), where distance and size effects arise from distinct sources: distance effect from semantic unit differences and size effect from symbol frequency (Krajcsi et al., 2018). The role of stimuli modality presentation, symbolic vs non-symbolic, seems therefore to be relevant in determining even the most "traditional" effect in numerosity comparison tasks.

The differential processing of symbolic and non-symbolic numerosities emerges also in studies that investigate the correlation between numerosity acuity and mathematical competencies. On one side it has been suggested that the development of ANS acuity might reflect not only the maturational process of the neural system per se but also the role of cultural factors that normally correlate with age, such as experience and formal mathematical education. It has been demonstrated that after training with a computer game designed to enhance number sense, particularly with games aiming to train numerical comparison skills, 7-to-9-year-old children with math learning disabilities significantly improved their performances in various core number sense tasks, such as number comparison itself but also counting and arithmetical operations (Wilson et al., 2006). In another study, it has been demonstrated that at 3 and 4 years the performance in numerosity comparison was linked to proficiency with symbolic numbers, even after controlling for short-term memory, IQ, and age (Mussolin et al., 2012, Figure 1.2). Piazza and colleagues (2013) demonstrated that also in an indigenous group, the Mundurucú (who possess a limited numerical vocabulary and have varying degrees of access to formal mathematical education) the level of education significantly correlated with proficiency in a numerosity discrimination task.

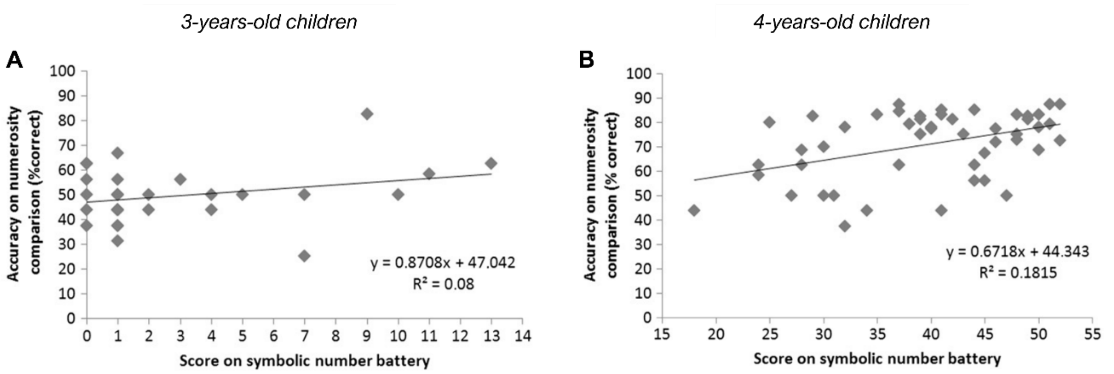


Figure 1.2: Numerosity comparison and symbolic number proficiency correlate at young age
Panel A) Numerosity comparison accuracy (y-axis) correlates with symbolic number proficiency (x-axis) in 3-year-olds
Panel B) Numerosity comparison accuracy (y-axis) correlates with symbolic number proficiency (x-axis) in 4-year-olds
Modified from Mussolin et., 2012

On the other side, research by Schneider and colleagues (Schneider et al., 2017) found through a meta-analysis that magnitude processing is consistently associated with mathematical competence across the lifespan and different tasks, measures, and mathematical domains. However, this association is generally stronger for symbolic than for non-symbolic numerical magnitude processing. This analysis therefore underscores the importance of examining symbolic and non- symbolic numerosities separately to gain a more comprehensive understanding of numerosity processing and its related effects (as will be discussed also later in the thesis).

The possible correlation between more daily life activities with numbers and ANS acuity is consistent with data collected in an Indigenous population. In a study investigating numerical abilities in Munduruku, a population with reduced lexical terms for representing numerosities, Pica and colleagues (2004) found that a Munduruku group of subjects was indeed able to discriminate non-symbolic numerosities similarly to a French-speakers control group. However, Munduruku adults' Weber fraction was estimated to be 0.17, whereas French adults' was estimated to be 0.12, meaning that the French had a better acuity for number discrimination. The changes in acuity in ANS, especially if related to the acquisition of symbolic competence of numbers, is an important aspect to take into account when discussing if and how symbolic versus non-symbolic stimuli presentation can influence the performance at numerosity comparison tasks. This will be later considered in the discussion paragraph of Study 1. Now we will discuss some basic features of ANS, especially those concerning non-human animals and preschool children, as an introduction to the experimental chapters.

1.1.1. ANS across species

As previously mentioned, the ability to process numerical quantities, or numerosities, has been extensively studied across a broad range of non-human animals. Research has shown that many species, including birds, mammals, reptiles, amphibians, and even some invertebrates, possess the cognitive ability to estimate, compare, and sometimes manipulate numerical quantities (see for example Lorenzi et al., 2021 for a review). Starting from primates (see Beran et al., 2015 for an extensive review), it has been demonstrated that they can perform tasks involving numerical judgments and basic arithmetic (Anderson et al., 2007; Beran et al., 2008; Brannon & Terrace, 1998; Thomas & Chase, 1980). For example, Cantlon and Brannon in 2006 demonstrated that two female rhesus macaques were not only able to learn to order numerosities from 1 to 9 in ascending order but also to subsequently generalize this ordering ability to novel numerosities (Cantlon & Brannon, 2006). It has also been demonstrated, in several different studies, that primates show a ratio-dependent acuity similar to humans. For example, orangutans (*Pongo pygmaeus*) discriminate between two quantities by the function of the ratio between them (e.g., 2 vs. 3 with a ratio of 0.67 was easier than 5 vs. 6 with a ratio of 0.83; Call, 2000). Comparable spontaneous food quantity discrimination was observed in several other non-human primates (Hanus & Call, 2007; Shumaker et al., 2001). Numerical abilities have been documented in other mammals, including rats (Davis & Albert, 1986), dogs (Ward & Smuts, 2007), cats (Pisa & Agrillo, 2009), lions (McComb et al., 1994), elephants (Perdue et al., 2012), and several other species (see Lorenzi et al., 2021 for a review).

As we will see more in detail in the next chapters, domestic chicks (*Gallus gallus domesticus*) have been broadly investigated for their numerical abilities. For example, it has been demonstrated that 3-day-old chicks preferentially approach the numerosity they have been reared with instead of a novel one (Rugani, Regolin, et al., 2010). Also, basic arithmetic abilities appear to be present at birth in chicks. In a study conducted by Rugani and colleagues (2009), chicks were reared with five identical objects. In the subsequent test phase after the initial disappearance of two sets of similar objects (4 objects behind one screen and 1 behind the other), some objects were visibly transferred, one by one, between the screens before the chick was released into the arena. The chicks naturally chose the screen hiding the larger number of elements, regardless of the directional cues from the initial and final movements, suggesting that the animals performed basic arithmetic, such as adding or subtracting objects, during their first experiences.

Even more interesting is the evidence of numerical cognition in more phylogenetically distant taxa. Using a delayed match-to-sample task, abilities similar to those found in human and non-human primates have been found in crows. Ditz and Neider investigated the numerical discrimination abilities of two carrion crows (*Corvus corone*) across a wide range of numerosities (i.e. from 1 to 30) and demonstrated that these birds relied on an analog number system for discrimination, displaying the characteristic Weber-fraction pattern (Ditz & Nieder, 2016). The ability to distinguish between larger and smaller quantities is a fundamental skill observed even in invertebrates (see Bortot et al., 2021 for a review). For instance, quantity discrimination has been studied in mealworm beetles (*Tenebrio molitor*, Carazo et al., 2009) and it has been demonstrated that both honeybees and ants use proto-counting strategies to orient themselves (Chittka & Geiger, 1995; Dacke & Srinivasan, 2008; Wittlinger et al., 2006). Another example of how much numerical cognition is present and varied in nature, even among animals with simpler brains, is seen in honeybees (*Apis mellifera*). Research has shown that honeybees can employ different numerical strategies to solve foraging tasks. In one study, Howard and colleagues demonstrated that bees could learn a "less-than" or "greater-than" rule through appetitive-aversive differential conditioning (Howard et al., 2018). Another study by Bortot and colleagues found that honeybees, after training, spontaneously used an absolute-value strategy rather than a relative one to solve a similar task (Bortot et al., 2019). The authors of the latter study suggested that the different results between the two studies might be due to differences in the experimental paradigms. As will be discussed later, paradigm designs are crucial in studies of numerical cognition.

The observed qualitative and quantitative similarities in human and non-human behaviors in numerical-based tasks support the idea of ANS as a shared, evolutionarily ancient mechanism for numerical representation. Moreover, non-human animals have also demonstrated more complex forms of numerical cognition, including the ability to map numbers onto space and to perceive numerical correspondences across different sensory modalities, which will be discussed in the following paragraphs.

1.1.2. Numerical cross-modal transfer

Cross-modal numerical comparison or transfer refers to the ability of both human and non-human animals to assess and compare numerical quantities presented through different sensory modalities, such as visual and auditory formats. Barth and colleagues (2003) demonstrated that when adults engage in numerical discrimination tasks across different sensory modalities, they do so with no additional costs compared to performing the tasks within the same sensory modality. In a following study, they demonstrated that even preschoolers can both compare and add large numerosities showing the same level of accuracy within a single modality (i.e. arrays of dots) and across two modalities and formats (i.e. dot arrays and tone sequences; Barth et al., 2005). At 6-months infants can predict the number of appearing objects based on the number of anticipatory sounds not only when tested with small numerosities (Kobayashi et al., 2005), but also when tested with larger numerosities. Feigenson tested 6-month-olds with multimodal presented numerosities from 4 to 8 and found that infants showed an acuity ratio following Weber’s law in the same way as in experiments that have been done with unimodal testing (Feigenson, 2011). Similarly, at birth human newborns consistently fixated longer on displays that matched the number of auditory sequences they were familiarized with, but only when the numerical ratio was 1:3 (i.e. 4 vs. 12, 6 vs. 18) and not for the smaller 1:2 ratio (i.e. 4 vs. 8) (Izard et al., 2009). Demonstrating that cross-modal comparison is achieved also for large numerosities is a crucial aspect of proving *abstract* representation of the number sense. As discussed by Izard and colleagues in their study on newborns, testing larger numerosities avoids the possibility that participants are memorizing a multimodal representation of unitary objects that can be individually tracked in memory (e.g. processed by the OTS). Succeeding in tasks with large numerosities would demonstrate that the numerical nature of the stimuli is perceived *per se* and processed in an abstract manner, as the numerosities are far beyond the memory limits of the OTS. Smyth & Ansari (2020) conducted a comprehensive p-curve analysis of infant numerosity discrimination studies, highlighting the differences between unimodal and cross-modal tasks. They assess the evidential value for numerical discrimination in cross-modal tasks for both small and large numerosities, confirming the infants' ability to perform abstract numerical transfer across different sensory modalities (Smyth & Ansari, 2020). Cross-modal numerical matching has been demonstrated even in more complex contexts. For example, at 7 months infants are able to match the number of voices they hear, with the number of faces that they see (Jordan & Brannon, 2006).

Similar evidence of cross-modal numerical correspondence has been reported in rhesus monkeys (*Macaca mulatta*). When facing two different muted videos playing either two or three simultaneously vocalizing conspecific faces, the animals spontaneously oriented themselves to the video showing the number of faces corresponding to the number of calls they were in a background played audio (Jordan et al., 2005). Since all visual and auditory elements were synchronized and had identical durations, the subjects were unable to rely on amodal cues, like rate, to make a match. Another example of numerical generalization across sensory modalities in non-human animals comes from rats. The animals were trained to press one lever in response to two lights or two bursts of white noise, and a different lever in response to four lights or four bursts of white noise. When presented with two lights and two sounds simultaneously, the rats pressed the lever corresponding to "four," indicating that they spontaneously combined the quantities of light and sound and reacted to the overall sum. This response occurred even though the combined stimulus was made up of two components that, if presented separately, would have required different responses. Later, when the rats were shown a new combination of two lights and two sounds, they again responded by pressing the lever associated with "four" (Church & Meck, 1983). Not only did this demonstrate that rats have been able to generalize across two different modalities, but they were also able to manipulate them and combine the two pieces of information performing an addition. The ability to access and manipulate numerical information across different sensory modalities has also been found in rhesus macaques (Jordan et al., 2008). In a first experiment monkeys successfully matched different numerical stimuli across different presentation formats, either as sequential or simultaneous arrays, within the visual modality or across both stimulus formats and sensory modalities (visual versus auditory). In a second experiment, they successfully selected a visual array corresponding to the sum of visual and auditory stimuli previously presented (Jordan et al., 2008). It is interesting to note that cross-modal numerical integration does not seem to be limited to audio/video stimuli, but there is also evidence of successful approximate numerical judgments in cross-modal auditory/tactile presentation (Togoli & Arrighi, 2021; Tokita & Ishiguchi, 2016).

The ability to perform cross-modal numerical integration may have an evolutionary advantage (for other cross-modal correspondences see Ratcliffe et al., 2016 for a review). For example, Wang et al. (2021) revealed that at 4-month-old the co-occurrence of auditory and visual stimulation enhanced the acuity of the ANS, shifting from a 1:4 ratio to a 1:3 ratio (Wang & Feigenson, 2021). Similarly, it has been demonstrated that multisensory information can boost numerical matching abilities in young children. In a number-matching computer game, 5-year-old children were instructed to choose a picture depicting the same number of elements they saw, heard, or saw and heard in a previous moment. All the possible combinations of choice numerosities from 5 to 9 were used, in ratio ranges from 0.56 to 0.89. The children performed significantly better on the numerical tasks (i.e. better accuracy across all ratios) in trials where both visual and auditory information were provided compared to unimodal trials (Jordan & Baker, 2011). These studies suggest that the redundancy of numerical information provided in different modalities increases numerical discrimination abilities, which could be of advantage for all numerically based behaviour, such as hunting and mating, also in non-human animals.

1.1.3. Neural basis of ANS in humans

The neural bases of the ANS have been extensively studied, with converging evidence from neuroimaging, neurophysiological, and developmental research highlighting the critical role of the intraparietal sulcus (IPS) within the parietal cortex. Functional Magnetic Resonance Imaging (fMRI) studies have consistently shown that the IPS is selectively engaged during tasks that involve numerical processing, such as the rapid estimation of the number of objects in a visual array or the comparison of magnitudes between two sets of stimuli (Dehaene et al., 2003; Piazza et al., 2004). The IPS neural activation shows features similar to behavioral results previously described. Indeed, it has been demonstrated that the horizontal portion of IPS (hIPS) responds irrespectively to stimulus modality, e.g. visual or auditory (Piazza et al., 2006), or presentation mode, e.g. simultaneous or sequential (Dormal et al., 2010). Moreover, the same ratio effect found in behavioral responses has also been found in IPS activity (see Nieder & Dehaene, 2009 for a review). Through an adaption paradigm, where the distance between the deviant and adaptation numerosities was systematically manipulated, Piazza and colleagues demonstrated that the brain activation responses showed the characteristic features of Weber's law (Piazza et al., 2004). With the adoption of neurophysiological techniques, the role of specific groups of neurons in response to numerosities has been deepened, to the point that *number neurons* have been found in non-human and human animals (see Nieder 2016 for a review, Kobylkov et al., 2022; Kutter et al., 2018). However, for the scope of this thesis, we are more interested in discussing more global activation in the human brain specifically in humans (see Lorenzi et al., 2021 for a review on numerosities representation in the brain of non-human animals), and in particular to refer to lateralized responses in the two hemispheres.

Behavioral evidence of dissociable hemisphere involvement was already hypothesized in the late 90's. In one study with an aphasic and acalculia patient, Dehaene and Cohen reported a superiority for the left hemisphere in exact calculation and for the right hemisphere in approximate judgments (Dehaene & Cohen, 1991). More recent neuroimaging studies also found differences between left and right hemispheres activations. Piazza and colleagues (2004) found an overall bilateral activation of IPS in response to numerical changes. However, the authors also reported that the precision of numerical representation was generally more accurate in the left IPS compared to the right (Piazza et al., 2004). A similar study aimed to test notation-independent coding of numerical quantity in the hIPS found that when participants were presented with dot deviants among digits (i.e. non-symbolic among symbolic) there was a reduced left hIPS response. Using an adaptation paradigm, the author found that there was a significant recovery from adaptation for deviant dots across digits, but not for digits across dots (Piazza et al., 2007, Figure 1.3).

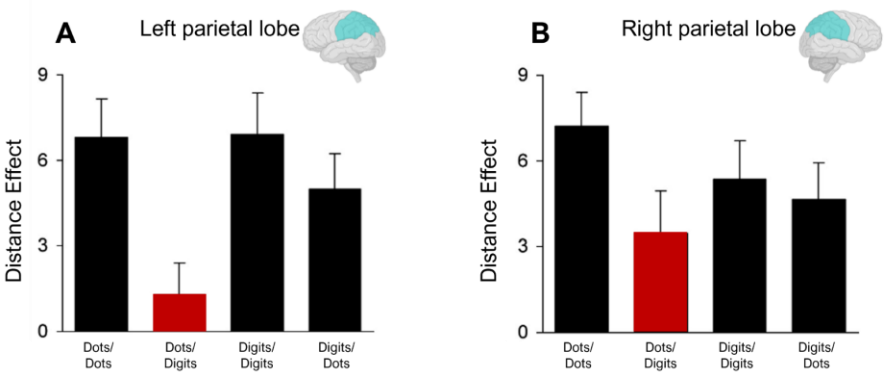


Figure 1.3: Distance effect in left and right parietal lobe as a function of numerical distance
In red condition where dot arrays were presented as deviant stimuli across digits. The left parietal lobe smaller distance effect (activation to far − close deviants) than the right lobe. According to the authors, this result suggested that the neural population responsive to dot stimuli was not adapting to digits. *Modified from Piazza et al., 2007*

The authors interpreted these results as evidence that the acquisition of number symbols affects more the left than the right parietal representation of numerosities (as proposed by Verguts & Fias, 2004). Further support for this hypothesis has also been provided by developmental studies. Rivera and colleagues (2005) observed that activation during a symbolic digit calculation task increased with age in the left parietal and left occipitotemporal cortices, but did not show a similar increase in the right parietal lobe. Cantlon and colleagues (2006) demonstrated that 4-year-old children exhibit selective activation of the right intraparietal sulcus (IPS) in response to non-symbolic numerical stimuli. In a later study it was demonstrated that children aged 3 to 6 years exhibit a strong neural tuning curve in response to changes in numerical values, similar to what is seen in adults, but primarily in the right IPS rather than the bilateral activation pattern found in adults (Kersey & Cantlon, 2017). This developmental curve, from right to bilateral, is further supported by evidence found in infants. At 6 months response to numerical changes is localized in the right inferior parietal region (Hyde et al., 2010), and at 3 months of age infants already showed a distinct activation for tasks involving the ASN in the right parietal-prefrontal cortex (Izard et al., 2008). The dominance of right-lateralized activation in young children suggests that the early stages of numerical cognition rely on non-verbal numerosity mechanisms managed predominantly by the right hemisphere. As children grow and develop language skills, there is a transition from a primarily right-lateralized pattern to a more bilateral one. This change might be associated with the integration of symbolic number representations, such as counting and arithmetic, which are linked to left hemisphere regions involved in language and semantic processing (Ansari & Dhital, 2006). As children learn to associate non-symbolic quantities with symbolic numerical concepts, the left intraparietal sulcus (IPS), which is connected to language networks in the left hemisphere, becomes increasingly active. This bilateral activation reflects the increasing role of language in numerical cognition, with the left hemisphere contributing to the processing of symbolic numbers and their meanings .

In conclusion, the combined pieces of evidence that numerosity acuity follows Weber's law, that cross-modal presentation seems to enhance performances, and that there are similar findings between literate adults, pre-verbal infants, and non-human animals, seem to suggest that the so-called *number sense* refers to mechanisms occurring at the perceptual level and therefore before any symbolic/linguistic processing start. Nevertheless, there are several pieces of evidence that cultural influences can affect certain aspects of numerical cognition, especially in response to symbolic number representation. In this thesis, we will especially focus on the Space-Number Association phenomenon.

1.2. Spatial Numerical Association

One of the aspects that has been long debated about numerical cognition and ANS's features is the Spatial Numerical Association (SNA) phenomenon. Since the first findings of Francis Galton our representation of numbers has been recognized to be spatially organized (Galton, 1880). In the framework of SNA, the Space Number Association Response Code (SNARC) effect has been identified as one of the strongest evidence of an existing mapping between numbers and space. The SNARC effect, in its first and pure definition, is described as the phenomenon that produces faster reaction times in response to small numbers using the left hand and to larger numbers using the right hand (Dehaene et al., 1993). This effect has been replicated in several studies, with different paradigms and stimuli, reinforcing the idea of the existence of a Mental Number Line (MNL), an internally ordered representation of numbers spatially organized in a horizontal plane (Dehaene et al., 1993; Restle, 1970). One of the current major debates about SNA and SNARC effects revolves around their origins. In the first years after their discovery, SNA have been strongly linked to cultural influences: left-to-right oriented SNA for left-to-right readers and reversed (or absent) SNA for right-to-left readers (see Göbel et al., 2011 and Toomarian & Hubbard, 2018 for a review). However, an independent growing body of evidence suggests that also non-human animals and humans at a pre-verbal stage of development exhibit SNA, in a consistent left-to-right direction. Toomarian and Hubbard (2018) exhaustively discussed how evolutionary and cultural factors might originate the MNL and give rise to the SNARC effect. The apparent contrasting findings of strong cultural influences on SNA with SNA being nonetheless present in subjects lacking possible cultural inputs, keep the debate about its origin still open.

Here we will review how the direction of SNA can be influenced not only by cultural habits but also by the task demands. Indeed, as we will discuss in the next chapter, we suggest that different SNA might emerge in relation to different cognitive mechanisms that are activated by explicit versus implicit tasks.

1.2.1. Spatial Numerical Association under cultural influences

In this paragraph, we will review how different cultural backgrounds (specifically writing/riding habits) influence SNA phenomena. The majority of these studies test the response to symbolic numerosities, such as digits and words, and therefore they involve mainly participants that have some degree of education, as the instructions per se require knowing and understanding their meaning. This also implies that these participants are likely able to read and write. Writing and reading are special actions in terms of spatial mapping. They require strong eye-hand coordination resulting in both visual and motor movements in space: we orient our eyes and move our hands systematically in the same direction each time we want to read or write something (Bergen & Chan Lau, 2012). This means that literate people of different cultures are constantly exposed to specific spatial direction routines: Romance language speakers automatically scan pages from top to bottom and left to right, while Arabic speakers follow a right-to-left direction (Fuhrman & Boroditsky, 2010). Here we briefly report SNA evidence as a function of cultural influences in three major groups of literate adults: left-to-right and right-to-left readers, and multilingual readers that speak different languages going in both directions. Moreover, we will also discuss how formal education is crucial in determining the emergence of some forms of SNA.

Left-to-right Readers

As already mentioned, the SNARC effect was first reported in 1993 by Dehaene and colleagues. While studying how parity is mentally represented they found that French-speaking participants were faster to respond with the left hand to small numbers and with the right hand to large numbers (Dehaene et al., 1993). Since then, the same kind of small-left and large-right association has been vastly reported in literature irrespective of the types of numerical stimuli used, including positive and negative integers, single and multi-digit numbers, and number words (Fischer & Rottmann, 2005; Nuerk et al., 2004). In a meta-analysis of the SNARC effect, more than 40 studies were reported, all conducted in Western populations, analyzing responses coming from left-to-right readers. In all these studies the direction of the reported SNARC effects was indeed left-to-right: small numerosities are associated with the left hemispace while large numerosities are associated with the right hemispace (Wood et al., 2008). Of particular interest for future discussion is the study from Nuerk and colleagues where they demonstrated that the SNARC effect was present regardless of the stimuli modality presented in Western adults. In this study, German students were required to express a parity judgment for numbers presented in different modalities, both visual and acoustical. Stimuli were Arabic numbers, written number words, spoken number words, and dice patterns. Participants were required to press a right or left response key according to the numbers being odd or even. The authors not only found a SNARC effect both for reaction times (RTs) and error rates in all the conditions but they also reported no significant interaction between number format, number magnitude, and hand response. Interestingly, also when comparing the individual SNARC slopes they found no differences between conditions (Nuerk et al., 2005).

Right-to-Left Readers

There is evidence of several lateralized behaviors in right-to-left-oriented cultures. For instance, Arab and Hebrew subjects showed a right-to-left directionality in various tasks such as perceptual exploration, drawing, and even aesthetic preference (Nachshon, 1985; Nachson et al., 1999; Tversky et al., 1991). If even behaviors that do not seem to be closely related to written language are influenced by its directionality, it should not be surprising that the same happens also for a potential symbolic-based phenomenon such as the MNL representation. Indeed, studies investigating SNARC effects in cultures where reading and writing happen from right to left report a reverse SNARC effect: small numerosities are associated with the right side while large numerosities are associated with the left side (Shaki et al., 2009; Zebian, 2005). For example in 2009, Shaki and colleagues reported that Palestinian participants, who used only Arabic-Indi numerals (where both words and numbers are written from right to left), showed a significant reverse SNARC effect in a classical parity judgment task: they were faster in correctly classifying small numbers placed on the right and large numbers placed on the left side of the screen. In the same study, Canadian participants showed a classical SNARC effect, while Israeli students, familiar with the English language, reported no significant SNARC effect (see Figure 1.4, Shaki et al., 2009). As a side note, It is also important to notice that Chinese and Japanese speakers do not present the typical SNARC effects in response to number words, but rather a top-to-bottom bias compatible with the influence of their language (Hung et al., 2008).

Multilingual Readers

Of particular interest are people who fluently speak and read in more than one language, with at least one of them being right-to-left oriented. They should not be subjected to a specific direction being more prominent or perceived as correct in their daily life, as in their cultural system, the two directions have the same value. In the original paper of Dehaene and colleagues, researchers found that Iranian subjects performing the same task had different results compared to only French-speaking subjects. Despite not showing any SNARC effect at the group level, the individual analysis identified different behaviors among participants as a function of the duration of their stay in France. In particular, participants who were there for longer, hence being more exposed to the Western language, showed a behavioral response more similar to the French group, while participants who had only recently moved from Iran showed a weaker or reversed SNA (Dehaene et al., 1993). A similar result has been replicated by Shaki and colleagues when they compared the responses of Canadians, Israelis, and Palestinians in a parity judgment task. Crucially, while Israelis write Hebrew words from right to left but Arabic numbers from left to right, Palestinians write both words and numbers from right to left. The direct comparison of the three groups demonstrated how the direction of the SNARC effect was strongly affected by their reading/writing habits: English speakers had a typical SNARC effect, Palestinians had a complete reverse SNARC effect, and Israelis, who were familiar with both the direction, had no SNARC effect in neither the two directions (see Figure 1.4, Shaki et al., 2009).

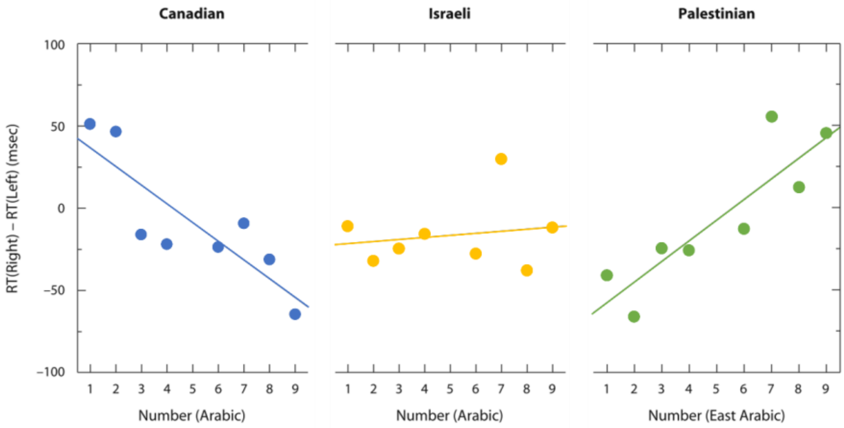


Figure 1.4: SNARC effect across populations with different reading/writing habits
The three graphs show SNARC effect as mean difference between $RT_{(right)} – RT_{(left)}$ for each presented digit. Negative slope (in blue, Canadian participants) indicates the classical left-to-right oriented SNARC, no slope (in yellow, Israeli participants) indicates the absence of a systematic association of numbers with space, and positive slope (in green, Palestinian participants) indicates the reverse right-to-left SNARC. *Modified from Shaki et al., 2019*

Besides being influenced by reading habits at the population level, SNAs look flexible within participants. In two different experiments, Russian-Hebrew readers were given a parity task with varying contexts designed to alter the cultural influences they responded to. In one study, they had to read a short Cyrillic Russian script (left-to-right direction) or a short Hebrew script (right-to-left direction) before performing a parity task (Shaki & Fischer, 2008), while in the second study, they alternated responses to Arabic numbers and number words, alternatively Russian or Hebrew (Fischer et al., 2009). In both scenarios, a reduced SNARC effect was found in the Hebrew-dependent context compared to the typical response in the Russian one. Although it does not take into account bilingual readers per se, it is worth mentioning that in 2010 Fischer and colleagues implemented an interesting study where they manipulated the position of some digits of a written recipe to modulate the SNARC effect in English speaker subjects. In the congruent SNARC condition, a group of participants saw digits from 1 to 4 displayed on the left side of the page and digits from 6 to 9 displayed on the right side. Conversely in the incongruent SNARC condition, another group of subjects was exposed to the opposite condition. The authors noted that although both groups exhibited a similar SNARC effect before reading the recipe, in the incongruent SNARC group this effect was absent afterward (Fischer et al., 2010). Taking together all this evidence seems to support the claim the direction of reading/writing habits influences the MNL representation and the relative SNA.

On the contrary, other several studies have reported marginal to no effect of language on the SNARC effect (Cipora, Loenneker, et al., 2019; Cipora, Soltanlou, et al., 2019; Dehaene et al., 1993; Heubner et al., 2018; Hochman et al., 2024; Zohar-Shai et al., 2017), therefore it has been hypothesized that other kind of cultural practice can influence SNA. For example, in a recent study, Bulut and colleagues used the unique context of Turkish culture, where reading is left-to-right, but cultural directional preferences generally lean right-to-left, to separate the influence of reading habits from other cultural practices. When comparing German and Turkish participants across various tasks and scores, they found notable differences between the two groups, despite their shared left-to-right reading direction (Bulut et al., 2024). The fact that reading/writing direction might not be the sole responsible factor of SNA direction, might also explain why SNA and specifically SNARC effect appear to be quite variable across and within participants. Indeed, the SNARC effect is widely recognized for its significant variability among participants, even among those who speak the same language and share the same reading and writing direction (Cipora, van Dijck, et al., 2019; Wood et al., 2008). For example, Roth and colleagues (2024) revealed that while the SNARC, MARC (Linguistic Markedness of Response Codes), and Odd effects (i.e. faster responses to even numbers; Hines, 1990) are robust at the group level, they exhibit significant variability at the individual level. Specifically, only a minority of participants consistently demonstrated these effects over 30 days. This substantial variability at the individual level not only explains the difficulties in replicating stable findings in SNA studies but also accounts for the wide distribution of results across samples (as we will see in our studies).

SNA as a function of formal education

If it is true that SNA effects are strongly determined by reading and writing habits, the passage from kindergarten to school should be a key point in its emergence. Indeed, despite the cognitive challenges that children might encounter in kindergarten, it is only from their first day of primary school that they are daily and constantly taught to read and write. Additionally, beginning in the first year of elementary school, visual number lines are introduced to help children learn counting and basic arithmetic. This is one of the first times that children are systematically taught that symbolic numbers follow a particular sequence in a specific direction. Several developmental studies have been carried on to verify the effect that formal education has on SNARC effects. One of the first, done in 1999, demonstrated that in classical parity judgment task (the same task used by Dehaene and colleagues in 1993) the SNARC effect emerges only at age 9, while it is absent at age 7, suggesting that 2 years of exposition to formal numerical knowledge was necessary to show an SNA congruent with the directionality of their cultural system (Berch et al., 1999). However, subsequent studies challenged the finding demonstrating instead that when investigated through a magnitude-relevant task it is possible to find typical SNARC effects in Western children as young as 7 years old, but not at age 6 (Gibson & Maurer, 2016; van Galen & Reitsma, 2008). Studies using the size congruity effect (i.e. faster and more accurate numerical comparisons when the physical size of numerals corresponds with their numerical magnitude; Henik & Tzelgov, 1982) provided also indirect evidence that explicit SNA effects emerge only after a certain level of knowledge about the symbolic number system. In adults, the numerical size and physical font interact during the task: a larger font facilitates numerical judgment and a larger number facilitates also physical judgment. During the development instead, while the influence of physical size over the numerical judgment task is observed also in children as young as 5 years old, the influence of numerical size over physical size emerges only later in school age (Gebuis et al., 2016; Girelli et al., 2000; Rubinsten et al., 2002). These studies provide shreds of evidence that a general association of magnitude with symbolic numbers is dependent on the level of formal education. As a consequence, also some types of SNA effects may depend on it, while other forms of SNA might emerge earlier in life (see next paragraph). Anyway, conflicting results about the exact age SNARC effects emerge might be caused by external factors such as different approaches in education across different school' systems and cultures. Nevertheless, a positive link between the effect size of SNARC effects and reading skills and age has been found in the already mentioned meta-analysis by Wood and colleagues in 2008 (Wood et al., 2008).

In conclusion, it is possible to claim that some forms of SNARC effects are dependent on cultural factors: the direction of the association is congruent with the cultural context (or almost absent when two different directions need to compete) and its emergence is strongly based on the knowledge of symbolic numbers and their order.

1.2.2. Spatial Number Association not under cultural influences

As we already mentioned, in the first decade after its discovery SNA has been strongly linked to cultural factors. Probably for this reason, for years the majority of the study investigated the phenomenon not only in literate adults but also by employing mainly digits and number words. In the meta-analysis of Wood and colleagues, among 46 studies and more than 180 experiments in total only 12 experiments were run with non-symbolic stimuli, i.e. no digits and no number words (Wood et al., 2008). Among them, 7 implied fingers counting, which we can consider to be also a culturally influenced practice, as it is normally taught and passed on from parents to children. As a consequence, the collected data were by definition strongly biased by the explicit knowledge of numbers. To better appreciate a not culturally biased manifestation of SNA, we believe it is worth shifting our attention, once again, to those subjects that are per definition lacking any possible cultural influence.

Spatial Numerical Association in Infancy

Studying preverbal children represents the possibility of accessing SNARC effects before the influence of reading/writing habits and any explicit formal knowledge of the numeric symbolic system. Evidence of SNA has been found even at the earliest stage of human life. Di Giorgio and colleagues (2019) demonstrated that newborns preferred to shift their attention according to a left-right-oriented SNA. Two-day-old newborns were habituated to specific numerosity (i.e., 12 dots on a screen) and then presented at the test phase with novel numerosities, either smaller (i.e., 4 dots) or larger (i.e., 36 dots). Both during the habituation and test phases, the same stimulus was simultaneously presented on the left and the right side of the monitor. Analysis performed on newborns' gaze behavior revealed that they preferred to look longer at the smaller numerosity placed on the right side of the screen and the larger numerosity placed on the left side of the screen. Similarly, seven-month-old infants presented with an increasing sequence of dots preferred the condition where they were moving from left to right rather than right to left (De Hevia et al., 2014). Furthermore, eight-month-old infants showed the same oriented SNA as they oriented their attention to the left after being cued with a small number and to the right after being cued with a large one (Bulf et al., 2016). Even more interestingly, Hoffmann and colleagues reported how the explicit or implicit reference to the numerical aspect can determine the emergence of the SNARC effect at a young age. When asked to perform a classical magnitude-judgment task, 5-to-6-year-old children failed to show a coherent SNARC. Instead, the same group of children showed a small/left and large/right association when they were presented with colored numbers for a color discrimination task (Hoffmann et al., 2013). It is important to notice that this study was conducted with children enrolled in kindergarten, thus having an education level that likely allowed them to have enough familiarity with numbers to succeed in the task (i.e., recognizing digits) but not enough exposition to the directionality of reading habits to develop a SNARC effect in the explicit numerical task. Similarly, 5-year-old pre-schoolers showed a left-to-right bias when asked to identify the middle point on a line flanked by a small versus large set of dots (de Hevia & Spelke, 2009). These results support the idea that a spontaneous preference for a left-to-right order of numerosities might be available before being significantly exposed to the written system of our own culture when investigated through proper, i.e. more implicit, tasks.

Nonetheless, it needs to be considered that during infancy humans are still subjected to cultural input that is not directly related to reading/writing habits. For example, counting, drawing preferences of objects, and ordering preferences are spatially organized behaviors that can still be observed during daily life by preverbal infants or preschoolers. For example, when drawing a line to the left, a mother from a culture where reading flows from right to left implicitly teaches her child that moving leftward extends the line, while a mother from a left-to-right reading culture would convey the opposite (Patro et al., 2016). As already discussed, factors beyond reading direction may play a role in the development of SNA, potentially even influencing preschoolers and preverbal infants, though further research is needed to confirm this. However, the notion of an innate predisposition for SNA and the idea that cultural practices can influence them from an early age are not mutually exclusive. As argued by McCrink and de Hevia (2018) toddlerhood may be a time of flexibility regarding the directional aspects of spatial associations, as innate left-to-right scanning biases diminish while children start to absorb socially transmitted information about the spatial organization of their surroundings. Early tendencies to associate initial information with the left side of space and final information with the right will emerge only when the specific context of numbers is involved. To better understand this complex interplay of various factors, comparative studies across different species, cultures, and age groups are necessary.

Spatial Numerical Association in non-human animals

Another way to investigate SNA in the absence of cultural influence is to study non-human animals' spontaneous and learned behavior. One of the first evidence of lateralized SNA in non-human animals was found in avian species. Clark's nutcrackers (*Nucifraga Columbiana)* and domestic chicks were trained to choose the fourth or sixth element from the bottom along a sagittal line of sixteen items. At the test, the animals were presented with the same number of items horizontally distributed and were left free to choose one of them. Results showed that the birds chose significantly more often the fourth and the sixth item from the left, demonstrating a left-to-right counting strategy (Rugani, Kelly, et al., 2010). In 2015 Rugani and colleagues (2015) further investigated SNA in 3-days-old chicks, demonstrating that after being trained to circumnavigate a panel with a "middle" value numerosity (e.g. 5 dots), at the test phase the animals spontaneously preferred to approach the left-placed panel when presented with smaller numerosity (e.g. 2 dots) and the right placed panel when presented with larger numerosity (e.g. 8 dots). They also demonstrate that the SNA did not depend on the absolute value of the test numerosity, but rather depended on the relative magnitude change from habituation to test (see Figure 1.5, Rugani et al., 2015). The same findings have been replicated in a more recent study that further demonstrated that SNA emerged in response to the numerical magnitudes themselves, but not individual lateral biases (Rugani et al., 2020).

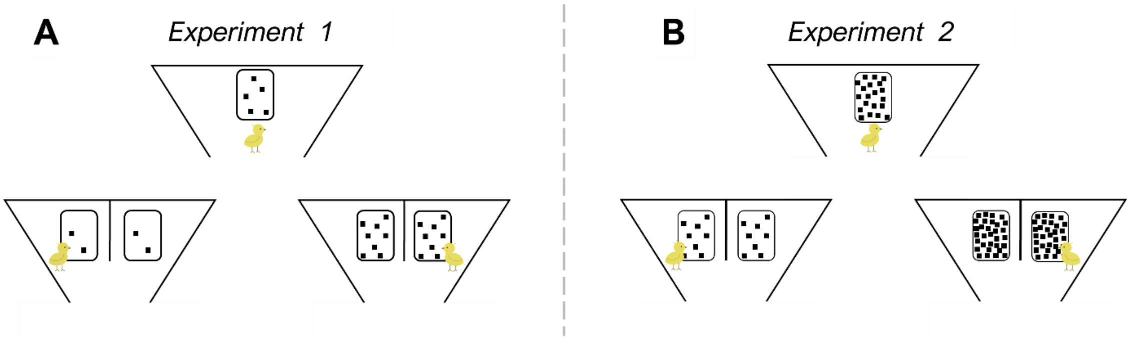


Figure 1.5: SNA behaviour in domestic chicks
Schematic representation of paradigm procedure in Rugani et al., 2015. Chicks were trained to circumnavigate a panel depicting a certain numerosity. At test, animals faced two identical panels on left and right sides of the apparatus, and they were free to spontaneously approach one or the other.
Panel A) Experiment 1 – Chicks were trained with 5 dots (top row), and presented at test with 2 dots or 8 dots (bottom row). Chicks facing the smaller numerosity spontaneously approached the left panel, while chicks facing the larger numerosity spontaneously approached the right panel.
Panel B) Experiment 2 – Chicks were trained with 12 dots (top row), and presented at test with 8 dots or 20 dots (bottom row). Chicks facing the smaller numerosity spontaneously approached the left panel, while chicks facing the larger numerosity spontaneously approached the right panel.
The second experiment demonstrated the *relative* and not *absolute* value of SNA, as chicks showed opposite behaviors in response to the same (8 dots) numerosity in the function of the perceived magnitude change.

Evidence of SNA following the left-to-right direction was found also in primates. Using a procedure similar to the one used with chicks, Drucker and Brannon (2014) demonstrated that four adult male rhesus macaques showed a left-to-right counting strategy on a horizontal line at the test after being trained to choose the fourth item on a vertical line. In another study, chimpanzees learned a number sequence task consisting of touching Arabic numerals from small to large. Later at the test, monkeys were presented with only 1 and 9 numerals, either presented as 1 on the left and 9 on the right (i.e., left-to-right condition) or as 9 on the left and 1 on the right (i.e., right-to-left condition). Results showed that they were faster to respond in the left-to-right condition, thus suggesting a privileged left-to-right direction of SNA (Adachi, 2014). More recently, Rugani and colleagues (2024) found additional evidence of magnitude-space mapping in rhesus macaques, as they demonstrated that monkeys recall better target position on the left when presented with a small array of horizontal dots, and on the right when presented with large arrays. More recently evidence of spontaneous left-to-right SNA has been found also in honeybees. By adopting the above-described procedure, in 2022, Giurfa and colleagues demonstrated the first evidence of SNA in invertebrates (Giurfa et al., 2022), supporting thus the hypothesis that the space-number mapping might rely on an ancient neural mechanism deeply rooted in the brain.

Despite the growing body of evidence that demonstrates SNA in non-human animals, caution is still needed in making direct comparisons. As argued by Patro and Nuerk (2017) in various species, lateralized behavior may be shaped by differing hemispheric organization, its developmental trajectory, or by active engagement in cultural environments with a high sensitivity to social learning. Therefore, directly applying spatial-numerical mapping mechanisms from non-human to humans is unwarranted and requires proper testing.

In summary, both animals and preschool children consistently show SNA, proving that a predisposition to associate numbers with space exists beyond cultural inputs. Moreover, they all show SNA in the same direction, suggesting that other reasons than cultural biases might be responsible for this phenomenon. Indeed, when examining these forms of SNA, only a left-to-right association has been observed, as no spontaneous right-to-left association has been reported in preverbal children or non-human animals, at least to our knowledge.

1.3. Domestic chicks as animal models

There is undoubtedly a strong connection between developmental studies and comparative studies. For instance, research on behavior and cognition in developing humans offers a foundation for comprehending basic intellectual abilities and how representation works in animal minds. On the other hand, comparative studies of animal cognition help to tackle challenging questions about the influence of innate abilities, culture, and language in the evolution of the more complex adult human cognition (Platt & Spelke, 2009). Moreover, they both help to track the phylogenetic and ontogenetic development of the neural mechanism that underlies the cognitive abilities of human adults, helping to understand their ecological values and their evolutionary paths. As already mentioned in the previous paragraphs these kinds of studies are powerful tools, especially for investigating how different phenomena manifest themselves in the absence of strong cultural inputs. However, regarding the human species, assessing cognitive abilities at birth presents both practical and ethical challenges. Humans are very immature at birth from both a sensory and a motor point of view, posing considerable limitations to performing behavioural experiments. Moreover, it is almost impossible to have fine control over their early experiences (e.g. visual experience, auditory experience, and social interactions). The same limitations hold for other altricial species (Sugita, 2008). On the contrary, precocial species are born with advanced motor and sensory abilities, enabling them to stand, walk, or interact with their environment shortly after birth. Among them, newborn chickens have been greatly used to investigate innate predisposition and abilities (see Versace & Vallortigara, 2015).

Domestic chicks are particularly useful as animal models because they are fully developed at birth, allowing for detailed behavioral studies soon after hatching. Moreover, as they fully develop *in ovo* it is possible to control and assess the impact of sensory experiences from the hatching to the test phase. Therefore domestic chicks have been used as a very successful animal model for studying various aspects of cognition (see Marino, 2017 for a review). In the last decades, the domain of numerical cognition has also been greatly investigated in avian species, specifically domestic chicks. We already discussed several studies in the previous paragraphs and other studies will be discussed in the details in Study 2 and Study 3 introduction. In the following two paragraphs, we will discuss two key methodologies used to study domestic chicks, which were employed in Study 2 and Study 3 of this thesis.

1.3.1. Spontaneous Preferences

Domestic chicks have been demonstrated to have several predispositions to approach preferentially some kind of stimuli. Generally speaking, these predispositions guide the precocial bird's attention towards features shared by living creatures: animacy features (Lorenzi & Vallortigara, 2021; Rosa Salva et al., 2015). An important function of these early preferences may be to guide the attention of young animals toward animate objects that provide care, facilitating the development of affiliative behaviors and later, sexual responses (Di Giorgio et al., 2017; Rosa-Salva et al., 2021; Vallortigara, 2012). For example, chicks developed a spontaneous preference for the red color (Ham & Osorio, 2007; Kovach, 1971; Salzen et al., 1971; Schaefer & Hess, 1959), given its association with salient features of hen's head and therefore its high ecological value. Similarly, they also have striking preferences for face-like patterns (Kobylkov et al., 2022; Rosa-Salva et al., 2010). Additionally, it has also been demonstrated, that predispositions in domestic chicks often manifest after they have undergone activating experiences, such as exposure to visual patterns, motor activity, handling, or acoustic stimulation (J. J. Bolhuis et al., 1985; Miura & Matsushima, 2016).

In research, spontaneous preferences and predispositions are often studied using the so-called *spontaneous choice paradigms.* These paradigms rely on observing the choices animals make when presented with different options, without prior conditioning or reinforcement. In this kind of test naïve chicks are individually placed inside an apparatus where they are presented with two different stimuli. The stimuli are normally placed at the two opposite ends of the apparatus, in the case of a longitudinal corridor, or at the end of two arms in a Y-maze-like apparatus. The animals enter in central area, the *starting area,* and then their movements are recorded for the whole duration of the test. It is typically measured by the time spent near each stimulus (in a pre-defined *choice area*), the frequency of approaches, or the dichotomic choice made by the animal (e.g., which arm of a Y-maze it enters). Those variables are used to determine preferences or perceptual biases. This type of paradigm is a powerful tool to study innate predispositions as they do not require prior training of the animals, they are easily adaptable to the presentation of a variety of stimuli (e.g. stimuli displayed on a screen, 3D objects, or sounds) and, in the case of precocial species as chicks, allow to manifest naturalistic behaviors (e.g. moving, pecking or even freezing).

1.3.2. Filial imprinting

Filial imprinting, is a form of early learning where a young animal forms an attachment to a particular stimulus, typically its parent or a surrogate (Bolhuis, 1991; Hess, 1959; Mccabe, 2013; Vallortigara & Versace, 2018). Imprinting is most famously observed in precocial birds, like chicks, ducklings, and goslings, which are mobile and visually active soon after hatching, and it occurs during a specific time window shortly after hatching, known as the sensitive period. During this time, chicks are particularly receptive to forming strong associations with visual, auditory, or tactile stimuli that are available to them, meaning that they are able to learn their characteristics through mere exposure. In just a few minutes, imprinting takes place, causing the animals to develop a preference for the imprinting stimulus afterward. Although imprinting is directed toward the mother hen in natural conditions, chicks can imprint on a plethora of stimuli (see Mccabe, 2013 for a review). Interestingly, researchers demonstrated that chicks can also be imprinted to stimuli flashing lights and virtual stimuli displayed on monitor screens (Bateson & Jaeckel, 1976; Versace, Spierings, et al., 2017) as well as static and moving 3D objects (see Mccabe, 2013 for a review). Therefore the imprinting phenomenon can be exploited in controlled laboratory settings for studying both learning abilities and attachment predispositions in chicks.

The typical imprinting-based study is normally divided into two main phases: the *imprinting phase* and the *testing phase*. During the *imprinting phase*, naïve newborn chicks are exposed to the so-called *imprinting stimulus.* Duration and structure of the imprinting phase can vary from less than one hour to a couple of days depending on the experimental design (the possible correlation between exposure duration and imprinting strength is still under debate, see Lemaire & Vallortigara, 2021). During the *test phase,* animals are normally placed in the same apparatus used for the *spontaneous choice paradigms* (see above) and they are then presented with a pair of stimuli (Figure 1.6). This pair can consist of the original imprinting stimulus and a completely new stimulus to assess recognition of the exact imprinting stimulus, or two entirely new stimuli, where one shares some features with the imprinting stimulus and the other does not, to evaluate the ability to generalize the features learned during imprinting.

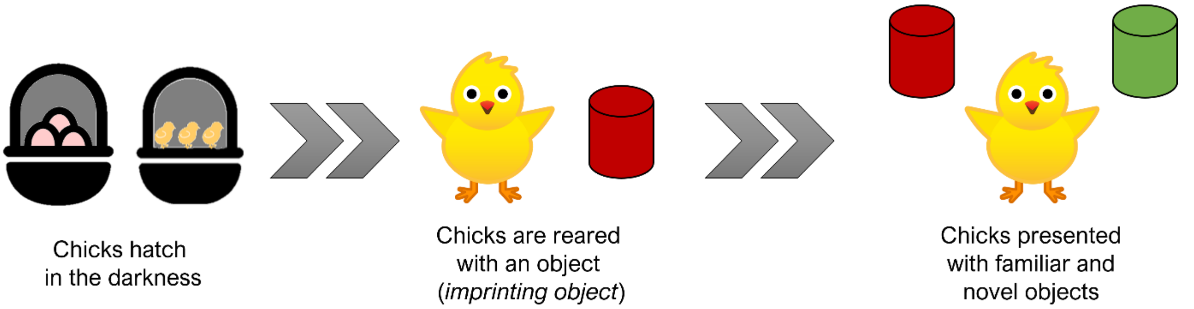


Figure 1.6: Schematic representation of imprinting procedure
Chicks hatch in the dark to ensure minimal exposure to the external environment. They are later reared with an object of interest for investigating specific perceptual features. During the test phase, chicks are presented with different stimuli to evaluate learning through imprinting.

1.4. Aim of the thesis

The present work aims to study some aspects typical of the ANS in different populations, in particular in the absence of cultural inputs.

In the first study of this thesis, we investigate SNA as a function of culture, age, and stimuli format. In Experiment 1 we tested SNA in three different populations: Italian adults, Italian preschoolers, and Himba adults. We investigated if even in the absence of formal numeric knowledge Italian preschoolers and Himba Adults could manifest SNA behaviours. Particularly we were interested in demonstrating that different kinds of tasks, explicit versus implicit, could lead to different patterns of SNA respect to Italian Adults. In the following study, we performed a follow-up study to better understand the role of stimuli modality presentation, i.e. symbolic versus non-symbolic, in a group of literate Italian adults. The aim was to shed light on the results of the previous study and to further support the discussed hypothesis of tasks' and stimuli's specific features' relevance in determining SNA.

In the second study, we investigated whether domestic chicks show numerical abilities when presented with auditory stimuli. The aim of the study was to test if the behaviors found in the visual domain, are present also in the auditory domain. In Experiment 1 we investigated the spontaneous preference for numerosities, to test whether also in response to auditory stimuli, chicks show a preference for larger numerosities. We ran two different experiments, one where extensive variables correlated to the presented numerosity (1a) and another one where they were controlled (1b). This has been made to better understand which features the animal responds to. In Experiment 2 we investigated whether chicks could be imprinted on a train of sounds of specific numerosities.

In the third study of this thesis, we studied chicks' ability to make cross-modal transfer of numerical information. In Experiment 3 we investigated whether chicks show a spontaneous preference for a congruent situation of numerosities seen and heard with respect to an incongruent situation. In Experiment 4 we tested whether chicks could be imprinted on numerosity acoustically and then recognize it in a visual test.

All three studies were conducted with the overarching goal of identifying which aspects of the ANS can be considered "innate predispositions," thus potentially providing primary evidence for a shared mechanism that is evolutionarily conserved. All the results will be discussed in the final section in the context of the current state of the art.

Chapter 2

Study 1

Influence of Culture, Age, and Stimuli Format on Spatial Numerical Association in Humans

2.1. General Introduction

One of the major debates about SNA regards its origin. There is now some emerging evidence that SNA behaviors could be divided into two different categories: a cultural-driven one and an innate one. As discussed, the first one would be strongly related to both the explicit nature of the task (e.g. consciously mapping numerosities into space) and to the symbolic nature of the stimuli used (e.g. digits and number words) and thus influenced by cultural contexts. The second one instead would emerge as the product of the biological predisposition to map numerosities into space depending on their valence value (Vallortigara, 2018, see discussion and conclusion paragraphs). The idea of the typical SNARC effect as a mapping between small/left and large/right might have emerged as a consequence of the Western culture being more widely represented in scientific research (Henrich et al., 2010). Therefore, it is possible that the dissociation of these two types of SNA phenomena has not yet been deepened enough as the cultural Western SNARC goes in the same direction as the innate one. This makes it difficult to discern them in human adults (at least in most countries) and requires shifting the attention to a population meeting specific criteria (e.g. tribes or preverbal children) or to animals, that are however more difficult to test. Another challenge for these hypotheses is also to explain why cultural context influences differently the responses to explicit and implicit tasks. A possible explanation is that the explicit direction of a specific language is recalled only when an equivalent level of explicitness is required by the task. According to this hypothesis, the presence of digits and number words should be enough to influence the level of perceived explicitness of the tasks, even though the paradigm itself is implicit, as they are symbols that are learned with a specific intrinsic order since elementary school. As consequence these kind of tasks could more easily lead to a cultural-dependent type of SNA, either in one direction or the other according to the spoken language.

To shed light on these issues we conducted two experiments. In the first study, we aimed to dissociate the culturally-determined form of SNA from the biologically-predisposed one. As already mentioned, to make it possible we had to shift our focus from the Western educated population only (where we suppose the explicit and implicit SNA would coincide in the left-to-right direction), to other populations not subjected to possible cultural biases. Thus, in Study 1a we investigated explicit and implicit forms of SNA in Italian preschoolers and Himba adults compared to Italian adults. In Study 1b, instead, we deepen the role of format stimuli presentation to understand its relevance in determining SNA phenomena.

2.2. Study 1a

A universal left-to-right bias in number-space mapping across ages and cultures[1]

2.2.1. Introduction

Number and space are inherently related in everyday life in cultures with formal education systems. From calendars to measuring instruments we are used to seeing numbers mapped onto space. Since Galton's initial systematic description of individuals reporting vivid associations between numbers and space (Galton, 1880), subsequent studies have consistently demonstrated that a majority of Western-educated adults exhibit a robust tendency to associate small numbers with the left side and larger numbers with the right side of space. Among others, one of the most studied and reported evidence of this Spatial Numerical Association (SNA) phenomenon is the *Spatial-Numerical Association of Response Codes* (known as SNARC). Being faster and more accurate to associate small numbers with the left hand/side of the space and large numbers with the right hand/side of the space (Dehaene et al., 1993). The strength and the direction of the SNARC effect, however, have been shown to be malleable to reading habits and contexts: while left-to-right readers systematically show a left-to-right SNARC (Fischer & Rottmann, 2005; Göbel et al., 2011; H. C. Nuerk et al., 2004; Wood et al., 2008), participants raised in cultures with writing systems that are organized from right-to-left or top-to-bottom tend to show no, inverted, or top-to-bottom SNARC (Fischer et al., 2009; Hung et al., 2008; Zebian, 2005; Zohar-Shai et al., 2017).

---

[1] This study is currently *under review* in an international peer-review journal and a preprint version has already been deposited (https://doi.org/10.31234/osf.io/w2st6). The text here report verbatim the methods section of the preprint version, while Introduction, Discussion and Results have been partially revised.

While the precise contribution of reading and writing direction per se is still under debate (Pitt & Casasanto, 2020), other cultural and contextual features seem to influence SNARC effects. For example, adding a clockface in the background and making time salient for the task(Bächtold et al., 1998; Mingolo et al., 2024) can reverse SNARC (i.e., participants associate small numbers with the right and large numbers with the left, as depicted on the clock), and performing a short session of finger counting in a non-canonical order can influence SNARC (Pitt & Casasanto, 2014). It has also been demonstrated that children are influenced by the cultural norms of their surroundings, even if they are not directly involved in these activities. For instance, children with differing reading directions for words and numbers develop number–space associations that tend to weaken with age due to a growing directional conflict. At first, young children without reading experience naturally adopt a counting direction in line with the reading direction they observe in others (right to left for Israeli children). However, as they begin learning to read, this initial bias diminishes, unlike the stable spatial associations seen in British and Palestinian children, because Israeli children learn to read text from right to left but are taught to read numbers from left to right in school (Shaki, Fischer, et al., 2012). Additional evidence in favor of the important influence of cultural practices like reading on the direction of the number-to-space mapping comes from the study of illiterate or pre-literate populations. In a recent study, Pitt et al. (2021) examined a group of adult Tsimanes, an indigenous culture with an oral tradition from Bolivia, and a group of U.S. preschoolers, to investigate the directional bias when organizing cards depicting 1 to 5 dots (or 2, 4, 8, 16, 32 in a second experiment) based on their numerical value. Both populations exhibited a lack of consistent directional preference, with individuals equally likely to arrange the cards in either a left-to-right or a right-to-left order. These outcomes suggest that the direction of the number-to-space mapping is largely determined by cultural inputs and that in the absence of cultural influence, mental mappings are "direction-agnostic," as concluded by the authors (Pitt et al., 2021).

This conclusion, however, seems to stand in stark contrast with an independent growing body of evidence from both human infant and non-human animal studies indicating an early and strong culture-independent direction-specific (left-to-right) bias in associating number to space. Starting from birth, humans preferentially orient towards the left when experiencing a decrease and towards the right when experiencing an increase in numerosity (de Hevia et al., 2017; Di Giorgio et al., 2019). Similarly, in their first year of life, infants prefer increasing sequences of dots moving from left to right (7 months old, De Hevia et al., 2014) and orient their attention to the left after being cued with a small numerosity and to the right after a large one (8 months old, Bulf et al., 2016). Studies with non-human animals (for example, Clark's nutcrackers (Rugani, Kelly, et al., 2010), domestic chicks (Rugani et al., 2015; Rugani, Kelly, et al., 2010; Rugani & Regolin, 2020) and honeybees (Giurfa et al., 2022)) also indicate an association between small quantities to the left side of space and large quantities to the right side. A simple interpretation of these findings is the existence of a universal biologically determined and possibly innate mechanism that predisposes to a left-to-right oriented SNA. How can we reconcile these findings with the previously reported absence of specific directional SNA in preschool children and populations with no reading-writing systems?

One possibility is that, in humans, two dissociable kinds of SNAs co-exist: one that is biologically determined, universal, and potentially innate and has a left-to-right directional bias, and another that is learned and that is shaped by the directionality of specific cultural input, such as reading and writing. According to this hypothesis, while the biologically determined SNA should manifest itself very early in life and persist throughout the lifespan, independent of both maturation and cultural influences, the culturally determined SNA only emerges in the presence of strong and well-integrated cultural practices such as reading/writing (Guida et al., 2018; Van Dijck et al., 2020). Additionally to this hypothesis, to account for the contrasting findings in the literature we also that while the biologically determined SNA, being reflexive, mostly manifests itself in tasks that do not require overt decision-making on SNA, tasks that overtly and explicitly require order numbers in space are mostly influenced by culturally-driven SNA. Therefore, the behavioral paradigms and/or tasks used to test the directional biases of SNAs might be key in differentiating among SNAs. Animal and infant studies necessarily rely on implicit procedures to examine associations between numbers and space. On the contrary, studies on children and adults typically require explicit numerical judgments (Pitt et al., 2021). The idea that different kinds of SNA manifest themselves in different tasks is not new in the literature. Patro and colleagues initially proposed a taxonomy classification of different SNA kinds, specifically referring to preschooler studies (Patro et al., 2014), later expanded by Cipora and colleagues by including SNAs observed also in adults (Cipora et al., 2018). In particular, they distinguish SNA phenomena into two main categories, extension and directional, with the latter specifically referring to tasks where the directional processing of numerical orders is relevant, rather than the magnitude value per se. Within this category, the authors report the existence of implicit coding SNA and explicit coding SNA based on the type of tasks required to study them (see Cipora et al., 2018 for extended definitions). Similarly, in the present work, we will use the terms "explicit" and "implicit" types of SNA indicating the explicit/implicit reference to space in the instruction of the task.

Our main aim is therefore to understand if SNAs emerge in implicit and explicit tasks and if the two can be dissociated. In particular, we expect that in implicit tasks, left-to-right SNAs emerge independent from age and cultural factors, while in explicit tasks, requiring a conscious ordering of numbers in space, left-to-right SNAs only emerge in those populations that have received a clear cultural left-to-right spatial ordering input. To put these hypotheses under scrutiny, we tested space-number mapping, using both explicit and implicit tasks, in three different populations: Italian adults, Italian preschoolers, and Himba adults, a population with an oral culture (no reading/writing) from Northern Namibia, with limited (orally transmitted) mathematical knowledge and no formal schooling, tested in two separate field trips, one carried out in 2021 and the other in 2022 (see Material and Methods). While Italian and Himba adults are equivalent in the overall level of maturation (unlike the Italian kindergarteners), Italian kindergarteners and Himba adults are equivalent in terms of their lack of, or extremely reduced, formal education. According to our hypothesis, we should find a similar SNA pattern across the three populations in the covert-implicit task, and a different pattern of spatial-numerical associations between Italian adults and the other groups in the overt-explicit task. Specifically, we expect a left-to-right SNA irrespective of age and culture (thus present in all three groups) in the implicit task, and a systematic left-to-right oriented SNA only in the adult literate group in the explicit task as in Pitt et al., 2021.

2.2.2. Material and Methods

The study has been conducted following the Declaration of Helsinki and has been approved by the local University of Trento Research Ethics Committee (for Italian preschoolers and adults' study) and the ethics committee of Inserm (for Himba studies; opinion number 21-855). The Italian adults gave their written consent after being informed about the study's purpose and procedures. The Himba participants, who were unable to read or write, provided oral consent following ethical guidelines. For the Italian preschoolers, we obtained informed consent from their legal guardians. Italian adults received a small monetary reimbursement for their participation, while Himba participants were given a set of four gift items including maize meal, sugar, vaseline, and soap (the approximate value of the items was equivalent to 5 euros).

2.2.2.1. Participants

- 191 Himba adults, an indigenous group with an oral culture from northern Namibia, with little or no formal education, were recruited in small villages. All were monolingual native speakers of Otjihimba, a dialect of the Otjiherero language. Data has been collected in two different missions, one in October/November 2021 (90 Himba people tested in 4 villages) and the other one year later November 2022 (101 Himba people tested in 2 other villages). Here, we will refer to the first group as Himba 2021 and to the second one as Himba 2022. Among these, we report data for 130 participants (60 from the first, 44 females, mean age = 33 ± 16 years, mean grade in school=0.5 ± 1.3; 70 from the second mission, 37 females, mean age = 33 ± 11 years, mean grade= 0.1 ± 0.3 years) who self-reported as being unable to read/write (note that a few participants have attended one or two years of school but were still unable to read/write). The same translator and experimenter conducted the two missions, and the experimental conditions were relatively similar. To our knowledge, none of the participants had prior experience with experimental research. It is important to mention that there is some uncertainty regarding the age of some participants, as it is culturally rare to track count of age. When necessary, we relied on the translator to provide an estimated assessment of the participants' ages. All experiments were conducted on an outdoor table, in a shaded area.
- 45 Italian preschoolers, recruited in kindergartens in the area of Rovereto (TN), took part in the study. Among them, 2 were discarded as they did not complete all three tasks. 43 children (21 females; mean age = 5.1 ± 0.3 years) were included in the final sample, and they were all reported by their parents as being unable to read/write.
- 47 young adults (36 female; mean age = 21.7 ± 2.3 years), recruited through the social media group of the University of Trento, took part in the study.

2.2.2.2. Stimuli

- *Ordering Card Task*

Stimuli were ten white cards (4 cm in diameter). On each card, black dots from 1 to 10 were printed. The dots were 0.5 cm in diameter, and they were displayed in a random configuration (Figure 2.1)

Figure 2.1: Stimuli used in the Ordering Cards Task
A picture taken on the field of one Himba participant during the Ordering Card Task. In this task, participants were asked to order 10 cards, depicting 1 to 10 dots, on the table in front of them, such that they would look "in order". No other instructions were given.

- *Numerosity comparison task*

For this computerized task, the stimuli were visual arrays of 4, 12, or 36 black squares (1.3 x 1.3 degrees of visual angle each, distance from the screen was approximately 50 cm) presented on a white background (17 x 17 degrees of visual angle; Figure 2.2).

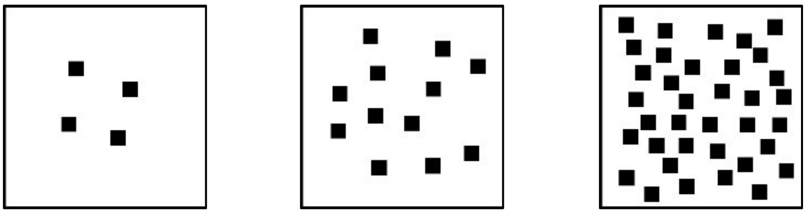


Figure 2.2: Stimuli used in the Numerosity Comparison task
Stimuli used in the Numerosity Comparison Task to represent 4, 12, and 36 dots. The stimuli are similar to the one used in Di Giorgio et al., 2019,

2.2.2.3. Experimental protocol

The first data collection for this study took place with the Himba population in 2021 and has been subsequently replicated with the Himba population, Italian Adults, and Italian preschoolers in 2022. The three populations ran the exact same three experiments (Ordering cards experiments and the increasing/decreasing version of the Numerosity Comparison experiment), albeit in slightly different settings: Italian adults were tested in a semi-dark room of the University laboratory, Italian pre-schoolers in a quiet room in their kindergarten, and Himba adults were tested outside, sitting at a shaded table located nearby their local villages. While all participants received oral instructions from the experimenter (helped by a translator in the case of Himba people), Italian adults also read the instructions on the computer screen (for the computerized experiment) or on a paper sheet (for the ordering card experiments). All the participants performed first the ordering cards task and then the numerosity comparison task. Before the experimental tasks, we interviewed both Himba and Italian adults on their literacy (i.e. their ability to read and write) and their schooling level. The same information about reading/writing ability was collected from Italian preschoolers' parents through a written questionnaire.

- *Ordering card tasks*

For the Ordering card task, participants were asked to manually order 10 cards depicting 1 to 10 dots on the table in front of them, with the only constraint that they would be considered to be “in order”. No other instruction was given. Participants were presented with the 10 cards piled up in random order and headed down to the table in front of them.

- *Numerosity comparison task*

Stimuli were presented on a laptop computer screen and participants sat at approximately 50 cm from it. Trials started with a fixation cross at the center of the screen that lasted for 1 s, followed by a set of dots presented centrally for 500 ms. Then, a black screen appeared for 200 ms and was followed by a second set of dots (hereafter “test stimulus”), which was presented on the left or on the right side of the screen. The test stimulus remained on the screen until the subject pressed the response key or for a maximum of 3 seconds (Figure 2.3). Participants were instructed to press a central key as fast as possible (with their dominant hand) only when the test stimulus was less numerous compared to the previous stimulus (hereafter “decreasing condition”), or only when the test stimulus was more numerous (hereafter “increasing condition”). Feedback was always provided (a green happy smiley for correct responses and a red sad smiley for incorrect responses were presented for 1 s after the participants’ response). The experiments started with 18 trials to familiarize the participants with the task, followed by 48 experimental trials. Of those, 30 were targets (i.e. trials that required a response; 15 presented on the left and 15 on the right) and 18 were distractors (i.e. trials that did not require a response: 9 presented on the left and 9 on the right). We analyzed the responses to the target trials only. All the experimental groups did the task in both conditions (with the order counterbalanced across participants), except for Himba 2021 who performed the task in a between-subject design. Both stimuli presentation and data collection were performed with Psychopy software (Peirce et al., 2019).

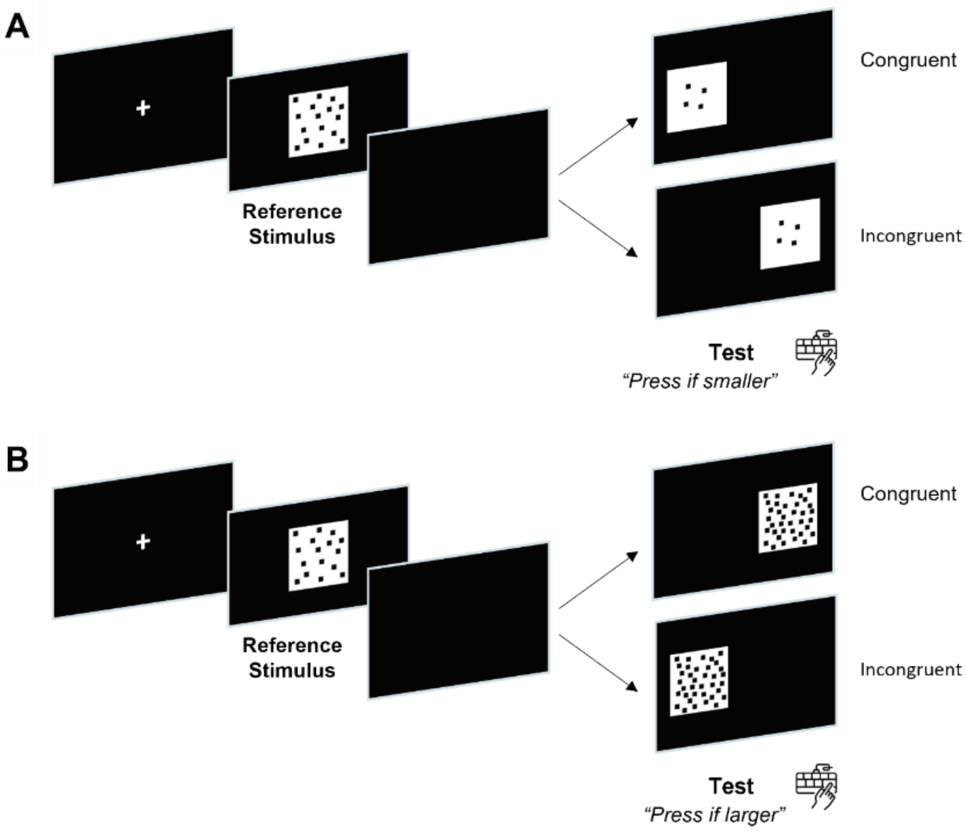


Figure 2.3: Schematic representation of the Numerical Comparison Task
Congruent and incongruent positions have been defined according to left-to-right SNA.
Panel A) Example trial for the *Decreasing Task instruction*, participants were instructed to press a central key as fast as possible only when the test numerosity was smaller than the prime numerosity.
Panel B) Example trial for the *Increasing Task instruction*, participants were instructed to press a central key as fast as possible only when the test numerosity was larger than the prime numerosity.

2.2.2.4. Statistical analysis

Data analysis and plot generation have been performed with R software (R Core Team, 2022), R version 4.1.3 (2022-03-10) in the RStudio environment (RStudio Team, 2020). For all statistical analyses alpha = .05 significance level was chosen. The effect size was measured according to Cohen's formula d = $\frac{M_1 - M_2}{SD_{pooled}}$ for parametric t-test and with the formula r = $\frac{Z_{statistic}}{\sqrt{N}}$ recommended for a non-parametric test; partial eta squared is reported for ANOVA effect size(Tomczak & Tomczak, 2014). Whenever a non-significant effect was determined (i.e. pvalue>.05), Bayesian analyses were performed to allow meaningful interpretation of the null results using the JASP software (version 0.18.3.0, JASP Team, 2024). Interpretation of the Log10$_{\text{(Bayesian Factor)}}$ is provided according to the scale proposed by Jeffreys in 1961 (Jeffreys, 1961).

- *Ordering cards task*

For the Free Ordering task, we started by performing a χ^2 analysis, on the frequency with which each type of spatial configuration (lateral line, sagittal line, two-dimensional, and random shape) was chosen by the participants in each group (Fig. 1B, Table S1). Then, for those participants who used a lateral linear configuration, we performed the quantitative analysis, which, following Pitt et al., 2021, proceeded in two steps.

First, we tested the *systematicity* of the lateral line mapping. To do so, for each participant we correlated the chosen order to that of the ideal left-to-right ideal order using Kendall's Tau, yielding a score between -1 and 1. The absolute value of the score indexes the systematicity of the mapping (i.e. how orderly: a score of +/-1 corresponds to a mapping where numerosity increases monotonically in one direction across all ten positions). Intermediate scores reflect imperfectly ordered mappings (Fig. S1). To determine whether participants performed the task by ordering the cards with a certain systematicity, we compared the distribution of mapping scores they produced to the distribution of mapping scores that would be expected by chance (i.e. as a result of random arrangements of the 10 stimuli on the line). The chance distribution was generated by performing all the possible permutations of the 10 elements sequences without repetitions (10! = 3,628,800 permutations) and it was compared with the real data distributions using the Two sample Kolmogorov-Smirnov goodness-of-fit test.

Second, we tested the *direction* of the mapping. The sign of Kendall's Tau correlation indexes the direction of the mapping: positive scores indicate generally rightward mappings, while negative scores indicate a leftward one. To determine whether a preferential direction of mapping emerged at the population level we tested whether the group average score differed significantly from 0, using a non-parametric two-tailed Wilcoxon signed-rank test to account for the non-normal distribution of the scores.

- <u>*Numerosity comparison task*</u>

We started the analysis by excluding participants who responded at or below chance (62.5% accuracy, i.e. the resulting accuracy of participants always responding to all the stimuli), resulting in the removal of 5 Himba 2021, 6 Himba 2022, and 5 pre-schoolers. To condense speed and accuracy in a single measure, for each subject and condition we computed the inverse efficiency score (IES), defined as $\frac{RT_{mean}}{Accuracy}$ (Townsend et al., 1983) (i.e. lower scores mean a better performance). The Shapiro-Wilk test was performed to test for the normal distribution of data, and data was not normally distributed. However, due to the necessity of testing for both repeated measure factors and interaction effects, we decided to perform an ANOVA test as it has been demonstrated to be robust against Type 1 Error also in case of violation of normality assumption (Blanca et al., 2017).

We first analyzed the Himba adults 2021 data with a two-way mixed ANOVA setting Congruency condition as a within-subjects variable and Task instruction as a between-subjects variable (Table S1). As we found a significant interaction between the Congruency condition and Task instruction, in the subsequent data collection and data analysis Task instruction has been considered as a relevant independent variable to insert into the model. We performed a Three-way mixed ANOVA, with Group (Himba adults 2022 vs Italian adults vs Italian preschooler) as a between-subjects variable, Congruency condition (congruent vs incongruent), and Task instruction (decreasing vs increasing) as a within-subjects variable (Table S2). Planned t-tests were performed to compare differences between congruent and incongruent conditions for each Task for each Group T-test comparisons were performed one-tailed with *less* alternative, as for priori defined hypothesis of congruent condition eliciting better performance (i.e. lower scores) than incongruent condition. (Table S4, two-tailed-tests also reported in Table S5).

To account for the average difference in the inverse efficiency across groups (see Table S1) mostly because children were slower and Italian adults were more accurate compared with the other groups, for each subject and Task instruction we also computed a congruency effect by normalizing the difference in IE for incongruent and congruent conditions by their sum: $(\frac{IES_{incongruent}-IES_{congruent}}{IES_{incongruent}+IES_{congruent}})$ Shapiro-Wilk test was performed to test for normal distribution of data (Tables S6, Fig. S2) and a 3x2 ANOVA was also implemented to test for differences across Group and Task instruction (Himba adults tested in 2021 were excluded from the model as they did not perform the Task in a within design fashion, Table 2). Then to investigate whether this congruency effect was significant we tested it using a two-tailed one-sample t-test against zero (alternative hypothesis $\mu \neq 0$) for normally distributed data and a two-tailed Wilcoxon signed-rank test against 0 (alternative hypothesis $\mu \neq 0$) for not normally distributed data (Fig. 2C, Table 2.3).

Mixed Model analyses were carried out to analyze Reaction Times, Errors, and Inverse Efficiency Scores using the R -package "*lme4*" (Bates et al., 2015). Given the data distributions, we modelled Reaction times and Inverse Efficiency Scores (IES) with a Gamma distribution and Errors with a binomial distribution. Here we reported the description of the models used for the three variables (see Appendix for the detailed model outputs and models comparison).

Inverse Efficiency Score model

For IES we tested the fixed effect of Congruency, Task, Group, and their full interaction. We also added the random effect of participants and the random slopes for the variable of interest (i.e. Congruency). It was not possible to add the random slope for type of testing as the IES is an aggregate value that collapses over different reference-test stimulus pairings. We also used the "*bobyqa*" optimizer from the "*glmerControl"* R package.

Model formula: IES ~ congruency * task * group + (1 + congruency | part_id)

Errors model

Italian Adults were not included in the model as they made no mistakes. For Errors, we tested the fixed effect of Congruency, Task, Group, and their full interaction. We also added the random effect of participants and type of test (i.e. the reference-test stimulus pairing). It was not possible to add random slopes for the variable of interest (i.e. Congruency) as it consistently led the model to fail to converge. We also used the "*bobyqa*" optimizer from the "*glmerControl"* R package.

Model formula: Errors ~ congruency * task * group + (1 | part_id) + (1 | type_test)

Reaction Times model

For RT we tested the fixed effect of Congruency, Task, Group, and their full interaction. We also added the random subject slopes for the variable of interest (i.e. Congruency) and the random effect of type of test (i.e. the reference-test stimulus pairing). We also used the "*bobyqa*" optimizer from the "*glmerContro*l" R package.

Model formula: RT~ congruency * task * group + (1 + congruency | part_id) + (1 | type_test)

2.2.3. Results

Explicit Spatial Numerical Associations

To investigate how number and space are associated in an explicit task, we asked participants to manually order 10 cards depicting 1 to 10 dots on the table in front of them, with the only constraint that they should arrange the cards so that they were "well ordered" (Fig 1A). In all three groups, the majority of participants spontaneously chose a lateral linear spatial arrangement (left/right), while other spatial layouts chosen were non-lateral linear arrangements (diagonal or vertical), 2-D geometric configurations (grids or circles), and no geometric shape (random arrangements). Figure 2.4 reports the percentages of choice per group (see also Table 1 in the Appendix for detailed contingency table and χ2 analysis).

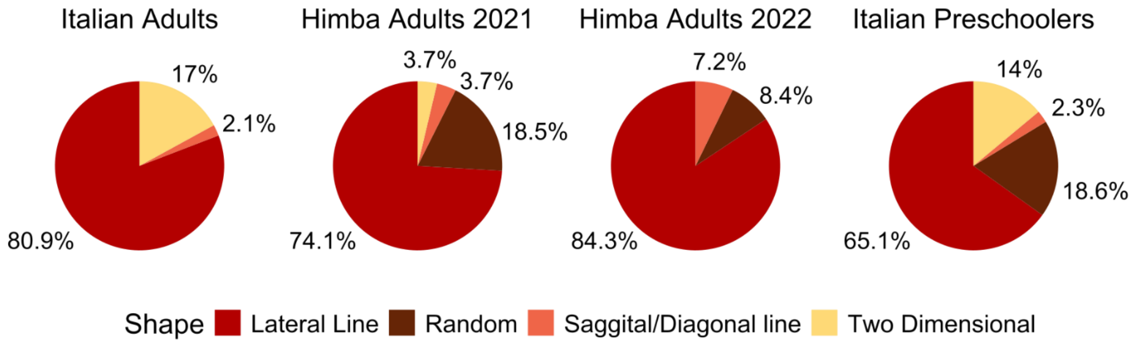


Figure 2.4: Ordering Cards task results
Percentage of choice for each type of spatial arrangement in the four groups. Lateral line (red) refers to a lateral disposition, Random (brown) to an absence of identifiable configuration (e.g. random scatterplot), Sagittal/Diagonal (orange) to non-lateral lines, and Two dimensional (yellow) to a configuration which was identifiable but not linear(i.e. grid, square, circle; see Table S1).

We then focussed on the data from those participants who organized the cards laterally (left/right direction) and we analyzed their mapping scores: for all groups, the distribution of their mapping differed from chance, indicating that participants in all groups spontaneously understood the task and arranged the cards based on their numerosity and not randomly (Two samples Kolmogorov-Smirnov goodness-of-fit hypothesis test; Italian adults: N = 38, D = 0.957, pvalue <.0001, Himba 2021: N = 60, D = 0.913, pvalue <.0001, Himba 2022: N = 70, D = 0.913, pvalue <.0001, Italian Preschooler: N = 28, D = 0.957, pvalue <.0001; Figure 2.5). However, only Italian adults systematically organized numerosities monotonically from left to right. By contrast, neither Himba adults nor Italian preschoolers showed such a systematic rightward bias, in that they were equally likely to order numbers from right to left and from left to right, as also confirmed by Bayesian analyses that reported moderate evidence against the alternative hypothesis (μ ≠ 0) for all populations, except for Italian Preschoolers that reported anecdotal evidence (Table 2.1, Figure 2.5). This result nicely replicates the one reported by Pitt et al. in 2021 and confirms that, for this explicit mapping task, in the absence of strong cultural bias, the mapping between number and space is arbitrary and inconsistent across individuals.

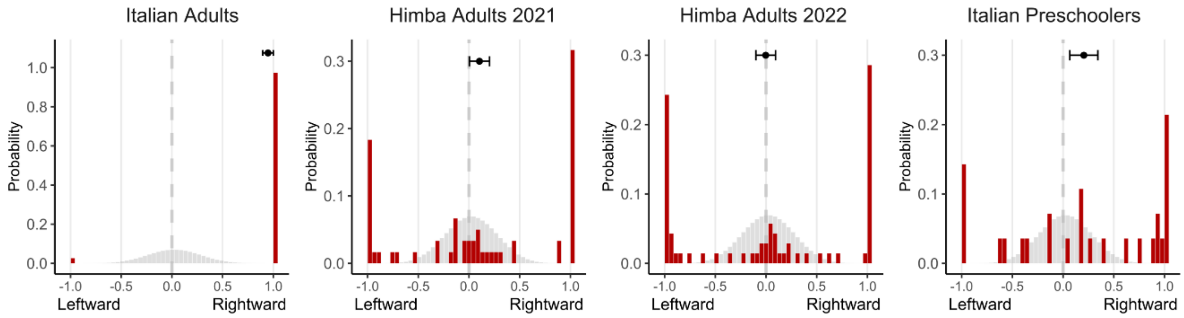


Figure 2.5: Ordering Cards task results
Kendall Tau's correlation between ideal left-to-right order and participants' disposition was calculated: negative values indicate leftward bias, positive values indicate rightward bias and values close to 0 indicate not-ordered mappings (see Methods section for a detailed explanation of the analysis). For all groups, the mapping distributions (red) are different from the chance distribution (grey; Two-tailed two-sample Kolmogorov-Smirnov goodness-of-fit hypothesis test; Italian adults: N = 38, p <.001, Himba adults 2021: N = 60, p <.001, Himba adults 2022: N = 70, p <.001, Italian preschooler: N = 28, p <.001). Black dots and whiskers show the mean mapping scores and the standard error of the mean. Only Italian Adults showed a consistent left-to-right preference at group level (Two-tailed one-sample Wilcoxon signed rank test; Italian adults: N = 38, p <.001, Himba adults 2021: N = 60, p = .257, Himba adults 2022: N = 70, p = .805, Italian preschooler: N = 28, p = .188; see Table 2.1).

Two-tailed one-sample Wilcoxon signed rank test against chance level (μ = 0)

Group	*Descriptive statistic*			*Test statistic*				
	N	*Median*	*IQR*	*Df*	*W value*	*pvalue*	*95% CI*	*Log10(BF)*
Italian Adults	38	1	0	37	721.5	<.0001	[1 1]	>2
Himba Adults 2021	60	0.067	1.6	59	1067	.257	[-0.022 0.040]	-0.564
Himba Adults 2022	70	0.022	1.96	69	1284.5	.805	[-0.156 0.111]	-0.842
Italian Preschoolers	28	0.020	1.31	27	261	.188	[-0.044 0.556]	-0.374

Table 2.1
Two-tailed one-sample Wilcoxon signed rank test against chance level (μ = 0) for mapping distributions

Implicit Spatial Numerical Associations

To investigate implicit number to space mapping we employed a computerized numerosity comparison task using a go-no-go paradigm. After the presentation of a first set of dots (reference stimulus), participants were required to press a central button when a second set (test stimulus) was smaller (decreasing task) or larger (increasing task), in different blocks. Critically, while the reference stimulus was presented in the center, the test stimulus was presented either on the left or on the right side of the screen. We defined targets as *Congruent* or *Incongruent* based on their quantitative relation with the reference and their spatial location: targets that were smaller than the reference and appeared on the left, as well as targets that were larger and appeared on the right, were defined as congruent, whereas smaller right targets and larger left targets were defined as incongruent. Congruency was thus defined with respect to the canonical left-to-right orientation (Figure 2.3). Using inverse efficiency as a composite summary measure of reaction times and accuracy, we found that contrary to what was found in the explicit task, where Himba adults and Italian preschoolers behaved differently from Italian adults, in the implicit task all groups exhibited a consistent congruency effect (better performance when smaller numerosities were presented on the left than on the right side of the screen) but only for the decreasing Task instruction (Table S3 and S4). These conclusions were also statistically supported by a 3x2 mixed ANOVA (Group x Task instruction) on the congruency effect, that showed a main effect of Task instruction ($F_{(1,160)} = 21.164$, $p_{value} < .001$, η2p = 0.117, see Table 2.2), no main effect of Group nor a Group x Task instruction interaction.

*Two-way ANOVA (Himba Adults 2021 excluded)**								
Factor	*Dfn*	*Dfd*	*SSn*	*SSd*	*F-value*	*p_{value}*	*$\eta^2_{partial}$*	*η^2_p CI 90%*
Group	2	160	0.007	0.928	0.576	.563	0.007	[0 0.034]
Task instruction	1	160	0.122	0.924	21.164	**<.001**	0.117	[0.049 0.196]
Group x Task instruction	2	160	0.007	0.924	0.599	.551	0.007	[0 0.035]
*One-way ANOVA (Himba Adults 2021 only)**								
Factor	*Df_n*	*Df_d*	*SS_n*	*SS_d*	*F-value*	*p_{value}*	*$\eta^2_{partial}$*	*η^2_p CI 90%*
Task instruction	1	57	0.048	0.541	5.011	**.029**	0.081	[0.005 0.206]

Table 2.2
Statistical analysis on congruency effect for Population and Task instruction
*Himba 2021 was analyzed separately due to differences in the paradigm design (see Methods section for details).
Significant p_{values} for .05 significance level in bold

In all groups, the congruency effect was significantly greater than zero only for the decreasing Task instruction, and not for increasing Task instruction as confirmed by Bayesian analyses that reported moderate evidence against the alternative hypothesis (μ ≠ 0) for all groups but the Himba adults 2022 for whom the evidence was anecdotal (see Table 2.3, Figure 2.6).

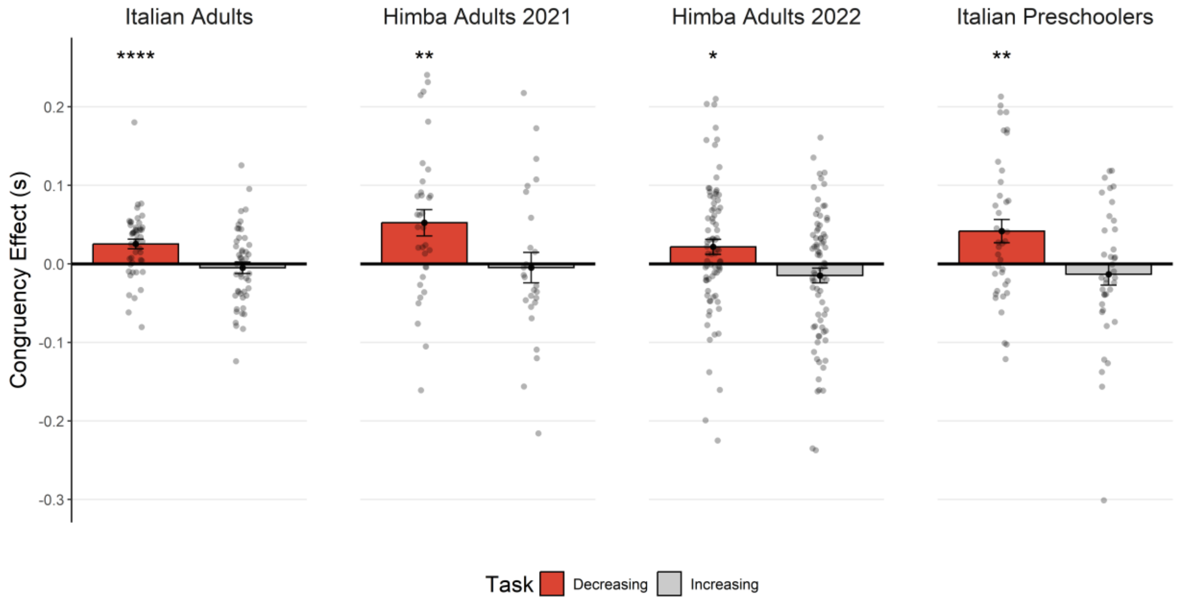


Figure 2.6: Numerosity comparison task results
We calculated the Congruency Effect for each Task instruction (see Methods, and Fig S2). In the graph mean value, standard error of the mean, and significant pvalues are reported for each population and Task instruction (Two-tailed one sample t-test and two-tailed Wilcoxon signed rank test; Decreasing condition (red): Italian adults: N = 47, p <.001, Himba adults 2021: N = 33, p = .004, Himba adults 2022: N = 78, p = .026, Italian preschooler: N = 38, p = .008; Increasing condition (grey): Italian adults: N = 47, p = .505, Himba adults 2021: N = 33, p = .805, Himba adults 2022: N = 78, p = .119, Italian preschooler: N = 38, p = .421; see Table 2.3). Significance levels are defined as follows: * = pvalue <.05, ** = pvalue <.01, **** = pvalue <.0001).

Two-tailed one-sample T-test against chance level (μ = 0)

Group	*Task*	*Descriptive statistic*			*Test statistic*				
		N	Mean	SD	Df	t-value	p_{value}	95% CI	Effect size**
Italian Adults	Increasing	47	-0.005	0.05	46	-0.67	.505	[-0.02 0.01]	-0.708
Himba Adults 2021	Decreasing	33	0.052	0.096	25	3.13	**.004**	[0.018 0.086]	0.545
	Increasing	26	-0.005	0.099	25	-0.249	.805	[-0.045 0.035]	-0.672
Himba Adults 2022	Decreasing	78	0.022	0.083	77	2.267	**.026**	[0.003 0.041]	0.257
	Increasing	78	-0.014	0.083	77	-1.575	.119	[-0.033 0.039]	-0.391
Italian Preschoolers	Decreasing	38	0.042	0.091	37	2.824	**.008**	[0.012 0.072]	0.458

*Two-tailed one-sample Wilcoxon signed rank test against chance level (μ = 0)**

Group	*Task*	*Descriptive statistic*			*Test statistic*				
		N	Median	IQR	Df	W value	p_{value}	95% CI	Effect size**
Italian Adults	Decreasing	47	0.034	0.045	46	918	**<.0001**	[0.016 0.038]	0.55
Italian Preschoolers	Increasing	38	-0.018	0.092	37	314	.421	[-0.036 0.018]	-0.564

Table 2.3
Statistical analysis against chance level for congruency effect for each population in each Task instruction.
* Non-parametric analysis was conducted for data not normally distributed.
**Cohen's D effect size is reported for Decreasing Task instruction, while Log_{10}(BF) as a measure of null effect is reported for Increasing Task instruction.
Significant pvalue for .05 significance level in bold

In order to verify the strength of our conclusions we also run a second set of analyses on the raw inverse efficiency data without pre-computing the congruency effect, using mixed-effect models. For each subject and each condition, we tested the inverse efficiency scores against three predictors: Congruency condition (congruent vs incongruent), Task instruction (increasing vs decreasing), and Group (Himba adults 2022, Himba adults 2021, Italian preschoolers, Italian adults), and their interactions. The results indicated a significant main effect of Congruency ($\beta_{Congruency\ condition}$= -0.107, SE = 0.445, z = 2.403, p_{value} = .0163) together with a trend for a Congruency condition and Task instruction interaction ($\beta_{Congruency\ condition\ *Task\ Instructions}$= 0.111, SE = 0.060, z = 1.846, p_{value} = .0648), indicating a stronger congruency effect for the decreasing Task instruction. No other interactions appeared. This analysis confirms and strengthens the previous results indicating that in all groups, participants were more efficient in detecting increasing numerosity when they appeared on the right side and decreasing numerosities when they appeared on the left side and that this effect tended to be more prominent when the task was to detect a decreasing sequence. Group factor also was significant as a consequence of the notable difference in accuracy and reaction times between the Italian adults and the other groups (see the Appendix for detailed models' output). Finally, to remain closer to the raw data, as the use of the IES may be sometimes debated, we also performed mixed-effect analyses on the reaction time and error rate separately (see Methods for detailed descriptions of the models). Error rate results consistently highlighted a significant main effect of Congruency condition ($\beta_{Congruency\ condition}$= 0.587, SE = 0.278, z = 2.107, p_{value} = .035) but this effect was not present in RTs ($\beta_{Congruency\ condition}$= -0.062, SE = 0.0391, t = -1.575, p_{value} = 0.115) and no Congruency Condition per Group interaction was found significant in any model for any variable of interest (see Appendix for detailed models' output), suggesting that the congruency effect is not affected by age or culture exposure. On the contrary, for RTs the only significant variable was Group, as the Italian adults were faster compared to all other groups. Taken together, these results indicate that when subjects are required to process numerosities, independently of age and education they show a congruency effect, being more efficient in processing increasing numerosity when presented on the right and decreasing numerosities when presented on the left side of space. This effect appears stronger for the decreasing task and more prominent in error rate.

Explicit and Implicit tasks are not correlated

As a final analysis, to further probe the different nature of implicit and explicit behaviors, and to further demonstrate the independence of one behavior from the other, we tested whether the tendency to map numbers to space of Himbas adults and Italian preschoolers in one task predicted tendency to map number to space in the other task (Italian Adults have been excluded from the analysis due to the extreme skewness of their data, see Figure 2.5). We correlated the explicit mapping scores with the implicit congruency effect for each Task instruction and each participant (Figure 2.7). No correlation was found between the two measures, as also confirmed by Bayesian analyses that reported moderate evidence against the alternative hypothesis ($\tau \neq 0$) (Two-tailed Kendall Tau correlation; for decreasing task instructions, N = 110, $\tau = -.073$, $p_{value} = .285$, 95% CI = [-0.209 0.064], log10(BF)= -0.633; for increasing task instructions, N = 110, $\tau = .017$, $p_{value} = .799$, 95% CI = [0.122 0.157], log10(BF)= -0.889). These results further suggest that the explicit and implicit performance are unlikely to be related, as they might be the result of two qualitatively different forms of SNAs.

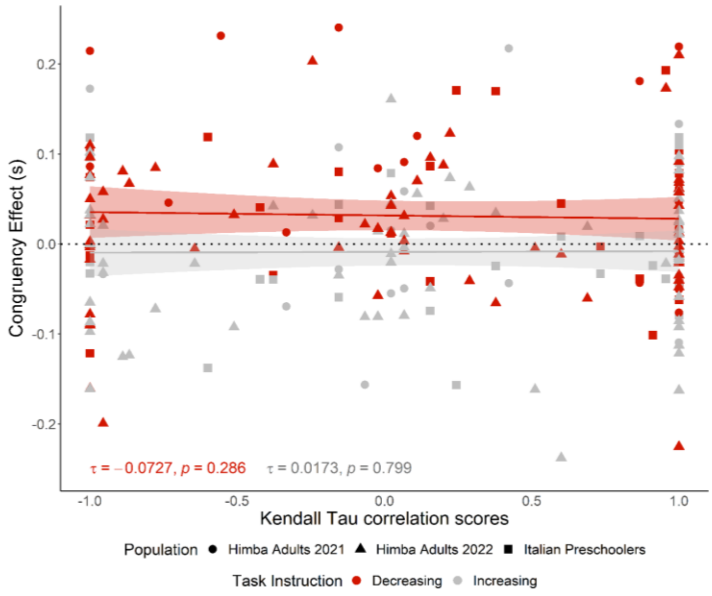


Figure 2.7: Correlation between explicit and implicit tasks
The correlation between Mapping Scores (x-axis) and Congruency Effect (y-axis) is not significant for either of the two Task instruction (Two-tailed Kendall Tau correlation; Decreasing task instructions (red), N = 110, p = .285; for Increasing task instructions (grey), N = 110, p = .799). The graph displays the linear regression with 95% confidence intervals, and individual points for each participant who completed both tasks (circles refer to Himba Adults 2021, triangles to Himba Adults 2022, and squares to Italian preschoolers). This analysis included only the participants for which it was possible to calculate the Kendall tau scores (i.e. those who ordered the cards on a lateral line, see Figure 1C), moreover, Italian adults were excluded from the analysis due to the skewness in their data (98% of the scores were perfectly equal to 1, see Fig 2.5).

2.2.4. Discussion

The origins of SNA and SNARC-like effects are currently a major topic of debate. One view holds that all SNA phenomena are determined by cultural input, in the absence of which the number to space mappings are "direction agnostic". This view, however, clashes with growing evidence that non-human animals and pre-verbal humans do display a clearly left-to-right directional SNA. Thus, the specific role of culture in the emergence of SNA phenomena remains to be understood. In the past, conflicting findings have been reported regarding SNA during the critical age when children begin schooling. Notably, while some studies suggest that children at this age do not consistently exhibit left-to-right oriented space-number mapping or SNARC-like effects before starting primary school (Berch et al., 1999; Gibson & Maurer, 2016; van Galen & Reitsma, 2008), others demonstrate that SNA can emerge when small adjustments are made to the task features. For instance, using non-symbolic instead of symbolic stimuli (de Hevia & Spelke, 2009) or indirectly accessing magnitude information through a color discrimination task (Hoffmann et al., 2013) has been shown to bring out these effects, highlighting how task design can significantly highlight or hinder SNA effects within the same age. We think that our results shed light on this debate, the nature of the tasks themselves might induce different components of SNA to emerge. We found that an implicit task can reveal a clear left-to-right oriented SNA in subjects who, when given an explicit task, either show no SNA or exhibit an SNA with an opposite orientation. Specifically, a task requiring explicit ordering of numbers in space elicited a systematically left-to-right oriented SNA only in Italian educated adults, while in a task where space was implicitly elaborated, we found evidence for a consistent left-to-right oriented SNA in all three tested populations. The experiment using the explicit task not only nicely replicated a previous study similarly conducted in another Indigenous population (Pitt et al., 2021), but it extended those findings. Indeed, while in Pitt and colleagues' experiment participants were instructed to order the numerosity in line and specifically on the lateral axis, in our experiment there were no instructions about the spatial disposition to use. Our results show that in all the populations the majority of the participants spontaneously chose a lateral linear arrangement. This result seems to suggest that the lateral linear organization of numerosities might not be strictly related to cultural factors. Further experiments are needed to investigate which biological or environmental factors might support the emergence of this spontaneous choice of a lateral arrangement of numerosities and if and to what extent this explicit behavior is consistent within the participants over different testing sessions.

The main scope of our study was to conduct two different experiments, one with an explicit spatial component and one with an implicit one, to investigate the influence of cultural exposure and conscious decision-making on SNA. While cultural background influenced performance only in the explicit spatial mapping task, it did not influence the response patterns in the implicit spatial mapping task. Indeed, illiterate Himba adults, and Italian preschoolers both mapped numerosities into space when explicitly asked, but without a systematic directional bias at the group level. As suggested by other researchers, this could mean that consistent directional SNA requires specific cultural practices or experiences to emerge (Cooperrider et al., 2017), but this is true only when the request is explicit and participants need to make conscious decisions on the mapping. In the context of implicit tasks instead, we believe that humans, as well as many animals, may be endowed with a biologically grounded mechanism by which the perception of changes in numerousness (and possibly other quantities) triggers specific lateralized spatial attentional shifts, thus showing an implicit form of SNA that might be not consistent with an explicit form of SNA even within the same individual. Indeed, the directionality of the SNA in the two tasks did not correlate across subjects, further supporting the idea that the two tasks investigate different phenomena.

What is the origin of this biologically grounded, culture-independent mechanism that links numbers to space? Two main hypotheses have been postulated. One is the brain's asymmetric frequency tuning (BAFT) hypothesis, where the hemispheric specialization for different spatial frequency (SF) bands is used to explain SNA. According to this theory, as SFs are present when fewer bigger elements are presented, an attentional bias towards the left side could occur as the right hemisphere predominantly selects low SFs (and vice versa for larger smaller elements with high SFs selected by the left hemisphere) (Felisatti et al., 2020). However, in our opinion, this hypothesis cannot fully explain SNA phenomena. Indeed, while as the authors describe in their work the BAFT hypothesis can explain the results of Di Giorgio and colleagues in 2019 obtained during the test phase of the experiment, it cannot explain the results obtained during the habituation phase. In the habituation phase when presented with two identical stimuli (i.e. sets of 4 dots or sets of 36 dots, the same used in Experiment 2, see Figure 2.2) on the left and on the right side of the screen newborns looked equal time to both the stimuli (Di Giorgio et al., 2019). According to the BAFT hypothesis, however, newborns should have looked longer to the stimulus on the right when there were 4 dots and to the left when there were 36 dots given the respective low and high values of SFs in the stimuli.

A more convincing hypothesis, instead, is that the biological predisposition for a left-to-right oriented SNA could originate in the lateralized organization of the bilaterian nervous system with the left side of the brain attending to stimuli with positive valence and the right side to stimuli with negative valence (Davidson, 2004). It has been argued that changes in numerosities towards larger or smaller numbers are associated with differential hemispheric activation and consequently with contralateral hemispace biased attention (Vallortigara, 2018). In an appetitive context, this would result in an attentional shift to the left for a change from larger to smaller numerosities (e.g., less food, perceived as a negative valence stimulus), and to the right for a change from smaller to larger numerosities (e.g., more food, perceived as a positive valence stimulus). In contrast, an aversive context would produce the opposite pattern (e.g., fewer predators, perceived as a positive valence stimulus, see also Figure 5.2 in the Conclusion paragraph). An advantage of this hypothesis is that it could explain SNAs as the results of overall increasing and decreasing change situations, being not limited only to some specific visual features as SFs bands.

We also found an asymmetry in the implicit task, showing a significant congruency effect for smaller stimuli presented on the left side, but a reduced one for larger stimuli presented on the right. Similar asymmetries have been reported also in other studies (Patro et al., 2015; Patro & Haman, 2011; Shaki, Petrusic, et al., 2012). For example, in 2012 Shaki and colleagues found that instruction-dependent effects (i.e. "*press for larger*" or *"press for smaller"*) occur in non-numerical magnitude tasks like judging size and height (Shaki, Petrusic, et al., 2012). It has also been demonstrated that at 4 months of age, infants can discriminate changes in ordinal relationships after being habituated to increasing sequences, but did not demonstrate discrimination following habituation to decreasing sequences (Macchi Cassia et al., 2012). These results indicate that infants exhibit an early advantage in processing ordinal relationships involving increasing magnitudes compared to decreasing ones. A general bias toward increasing changes in numerosity may also partially account for our findings. SNA effects may be subtler and secondary to other more automatic mechanisms; thus, in our study, they may have been masked in the increasing task instruction but were detectable in the decreasing task instruction. The precise dissociation of SNA in response to increasing or decreasing chance in numerosities should be better investigated in the future.

We also propose that the non-symbolic nature of our stimuli, as opposed to the symbolic stimuli often used in SNARC-like studies may explain this asymmetrical response pattern. The Approximate Number System (ANS, or number sense) is the ability to perceive quantities in a nonverbal and non-symbolic manner (Piazza & Izard, 2009), which is relevant to the stimuli used in Experiment 2. Research indicates that the neural basis of the ANS first develops in the right parietal cortex (Cantlon et al., 2006; Hyde et al., 2010; Izard et al., 2008; Kersey & Cantlon, 2017), extending later to the left parietal cortex (Ansari & Dhital, 2006; Cohen Kadosh et al., 2008; Piazza et al., 2004). This development is likely linked to language acquisition: the initial right hemisphere-based perception of quantities is later accompanied by left hemisphere-based semantic knowledge of numbers. In our task, which involves non-symbolic numerosities, the right hemisphere might likely play a more relevant role, enhancing performance in the left visual field (i.e. congruent condition in the decreasing task). However, the possible right hemisphere dominance alone cannot fully explain our results. If it did, we would also expect better performance for the incongruent condition (left visual field) in the increasing task, which we did not observe at a significant level. Thus, the most likely explanation for the observed asymmetrical effect between increasing and decreasing tasks is the interaction between increased right hemisphere activation and the left-to-right spatial-numerical association (SNA) in our task. Thus, the most likely explanation for the observed asymmetrical effect between increasing and decreasing tasks is the interaction between increased right hemisphere activation and the left-to-right spatial-numerical association (SNA) in our task.

Although we found evidence for left-to-right SNA in the implicit task in all the populations we tested, further research is needed to overcome some limits of the present study. For example, even if we only selected Himba adults who reported little to no literacy competence and Italian preschoolers who were reported to be illiterate by their parents, we cannot exclude that minimal contamination from the cultural context partially permeated the participants' minds. We also cannot exclude that participants used strategies based on non-numerical features to solve the tasks as in the present study stimuli were not controlled for other continuous variables, e.g. total colored area and convex hull. However, the significant difference in the distribution of mapping scores from chance suggests that not only Italian adults but also most Himba adults and Italian preschoolers accurately ordered the cards by their magnitude. The numerosities depicted exceeded the capacity of the subitizing system; therefore, participants had to actively count the dots to successfully perform the task. This is further supported by the absence of significant clusters of data around intermediate correlation values (e.g., -0.8 or 0.8), which would occur if a substantial number of participants ordered only the first half of the cards by numerosity while randomizing the rest. This suggests that the numerosities of the stimuli, rather than other dimensions, were considered and that the majority of participants could accurately count at least up to 10. Moreover, as we observed a difference between the two tasks, it remains clear that the dissociation between implicit and explicit processing was present regardless. Future studies will be needed to better disentangle the precise role of numerical features versus other non-numerical dimensions.

Another limitation of the current study could be found in the lateralized presentation of the test stimuli. Indeed, in 2018, Shaki and Fischer discussed how explicit magnitude processing and explicit spatial-directional processing might bias the SNA effects per se, as *"they may have artificially imposed spatial-numerical associations"* (see Shaki & Fischer, 2018). According to their description in our numerosity comparison task, both the magnitude processing and the spatial component are explicitly activated, as we ask participants to respond to laterally presented stimuli based on their magnitude value (i.e. "*press for less*" or "*press for more*"). However, we believe that since the stimuli are presented alternately on the left or right side of the screen (rather than simultaneously on both sides), any potential spatial attentional shift would equally affect both congruent and incongruent conditions, as the magnitude/side pairs are meticulously counterbalanced. Nevertheless, we found a congruent effect coherent only with the left-to-right SNA. Moreover, through the use of a central response key, instead of two lateralized response keys, we further reduced the possible effect of explicit spatial mapping happening at the hand-motor level. It is worth noticing, that given the unique challenges of working with Himbas and preschoolers (i.e. participants who are unfamiliar with computers and tasks of this nature, unlike Italian adults), we opted for a paradigm that required instructions as simple as possible (i.e. "*press one key for less/more*"). Additionally, classical parity judgment tasks were unsuitable, as neither Italian preschoolers nor the Himba possesses sufficient symbolic numerical knowledge to perform them. Furthermore, our implicit task was relatively brief, which may have introduced noise in estimating the individual subject's bias. Overall SNA effects have been demonstrated to be quite variable across individuals (see paragraph 1.2.1) and one possibility is for example that the hemispheric bias, even if biologically determined may exhibit minimally variable across subjects (see, e.g., the effect of the left hemisphere superiority in language processing in right-handed that is nevertheless not devoid of individual exceptions (Knecht et al., 2000)). As preliminary evidence of left-to-right oriented SNA has been now established with a simpler paradigm, future studies should explore the stability of the effect over other kinds of tasks across different tasks possibly overcoming the above-mentioned limitations.

In conclusion, this study provides valuable insights into the relationship between SNA, formal education, and cultural influences. Indeed, the findings highlight the importance of task instructions in revealing different forms of SNA and provide the first demonstration of dissociability between covert/implicit and overt/explicit behaviors within the same population lacking formal cultural biases to organize information in space. However, the debate about the precise foundations of these two dissociable types of SNA is still open. We suggest that on one side we are all endowed with a biologically predisposed SNA mechanism, rooted in general brain hemispheric asymmetry (Davidson, 2004), that is present from birth and remains constant in its left-to-right orientation over time, regardless of cultural influences. On the other hand, in explicit tasks requiring a conscious response, culturally influenced mechanisms support the final overt behavior. This work did not aim to investigate the nature of these latter mechanisms and how they are influenced by culture. Now that an initial dissociation between different forms of SNAs has been found, further studies are needed to understand which mechanisms support each form. Indeed, it is not yet clear how the culturally determined form of SNA emerges as a parallel independent process from the biologically predisposed one or whether it is its evolution. One possibility is that the direction of cultural biases in humans is purely conventional, thus not directly issuing from the inborn SNA mechanism. Indeed left-to-right and right-to-left writing/reading systems both exist, even though overall the right-to-left systems are much less frequent than the left-to-right ones (according to the ISO codes list, there have been only 40 languages written from right-to-left throughout all history, among more than 200 see https://unicode.org/iso15924/codelists.html). In sum, the biologically predisposed form of SNA may rely on different mechanisms from the ones that support forms of SNA that humans developed culturally, allowing for the co-existence of different types of SNA, even within the same individuals. Further research is needed to explore the intricate interplay between cultural practices, brain lateralization, and the development of SNAs, to shed more light on these intriguing phenomena.

2.3. Study 1b

Format-dependent SNA in Italian adults

2.3.1. Introduction

In the previous experiment, we found an asymmetrical effect in the implicit numerosity comparison task: participants showed a significant congruency effect in the decreasing condition only, meaning that they had a better performance for smaller stimuli on the left but not for larger stimuli on the right. As mentioned in the previous study discussion this effect might have been related to the stimuli modality, i.e. non-symbolic, or to the difficulty of the two task conditions themselves. To test these hypotheses we carried out a second experiment, with a similar procedure but different stimuli. Here we present some preliminary data.

2.3.2. Material and Methods

The experimental design was the same as the *Numerosity Comparison Task* in the previous study (see Figure 2.3), with minimal differences in the procedure (as described below).

2.3.2.1. Participants

Forty-one healthy Italian adults participated in the study (28 female; mean age = 23,4 ± 4,9 years), recruited through the social media group of the University of Trento. All of them completed the *Dots* experiment, while twenty-nine participants took part in the *Digits* and *Words* experiments. Written informed consent was collected before starting the test session.

2.3.2.2. Stimuli

Stimuli presentation modality was a key point of the experiment to primarily test the role of symbolic versus non-symbolic processing of numerosity. We decided to test one non-symbolic modality, as in the original experiment, and two symbolic modalities. In all modalities, we used numerosities from 12 to 20, with 16 as a constant reference stimulus.

- *Dots*

Non-symbolic stimuli were random configurations of black squares, designed based on the original experiment: 1.3 x 1.3 degrees of visual angle for each square on a 17 x 17 degrees white background (distance from the screen was approximately 50 cm).

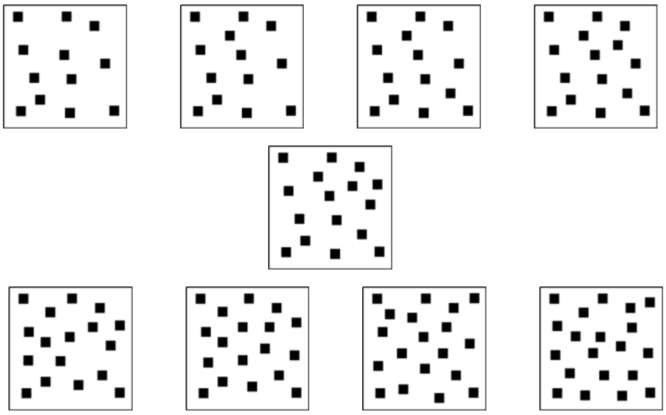


Figure 2.8: Stimuli used in the Dots experiment
Stimuli used as non-symbolic numerosities. Top row stimuli displaying numerosities from 12 to 15, central row 16 dots (always reference stimulus), bottom row numerosities from 17 to 20.

- *Digits*

The digit stimuli were Arabic numbers from 12 to 20. They were displayed on the screen in “Arial” font, 85 pt sized, white color on black screen.

- *Words*

Words stimuli were written number words from 12 to 20. They were displayed on the screen, all capital letters, in “Arial” font, 57 pt sized, white color on a black screen.

2.3.2.3. Experimental protocol

The participants were tested in a semi-dark room of the University laboratory where stimuli were presented on a computer monitor at a 30 cm distance. The whole test session lasted 40 minutes approximately. The experiment consisted of a total of six different tasks: three presentation modalities (Dots, Digits, and Words), each of them in both increasing and decreasing conditions. To avoid the priming effects of the symbolic presentation, all participants performed the non-symbolic task (i.e. dots) first, both conditions, and then the symbolic ones (i.e. Words and Digits, order counterbalanced across participants). The starting condition, i.e. increasing or decreasing, was also counterbalanced across participants. Differently from the original experiment, the reference stimulus was kept constant at “16” for all trials in all modalities in both conditions. Each condition, increasing and decreasing, consisted of 50 trials: 32 Go trials (target stimuli, 8 trials for each numerosity pair) and 18 No-Go trials (distractor stimuli).

2.3.2.4. Statistical analysis

For each participant, we collected Reaction Times (RT) in *congruent* and *incongruent* stimuli positions. As the first step of the analysis, we discarded outlier trials for each Experiment (i.e Dots, Digits, and Words), removing reaction times that were lower than the first quartile minus 1.5*IQR (interquartile range), or higher than the third quartile 1.5*IQR. Then, similarly to what has been done in the previous study, we condense speed and accuracy in the inverse efficiency score (IES), defined as $\frac{RT_{mean}}{Accuracy}$ (Townsend et al., 1983), for each Congruency Condition, Task instruction, and Experiment. We performed a 3x2x2 repeated measures ANOVA, with Experiment (Dots, Digits, and Words), Task instruction (Increasing versus Increasing), and Congruency Condition (Congruent versus Incongruent) as factors.

To parallel, the analysis performed in the first study we also computed the congruency effect as ($\frac{IES_{incongruent}-IES_{congruent}}{IES_{incongruent}+IES_{congruent}}$). A 3x2 repeated measures ANOVA was implemented to test for Experiment and Task instruction effects. Then we tested the congruency effect against zero (two-tailed Wilcoxon signed-rank test against 0) for not normally distributed data. We then conducted the same analysis by also adding the Test stimulus factor, to measure performance in response to the different numerosities presented. In case the ANOVA analysis revealed that data has violated the assumption of sphericity (i.e., Mauchly's test, $p_{value} \leq .05$), the Greenhouse-Geisser sphericity correction was applied to only factors violating the sphericity assumption. Corrected results are reported. Please note that, since not all participants completed every experiment, the ANOVA was conducted only on the subset of participants who participated in all experiments (i.e., 29 participants). However, the post-hoc analyses included the entire available sample for each experiment (as reflected in the degrees of freedom reported in the tables).

Then, we collapsed the data obtained by the two symbolic conditions, i.e. digits and words stimuli, aiming to further present the data as a function of the symbolic versus non-symbolic nature of the stimuli (Stimulus Type). We then performed the same analysis above described.

2.3.3. Results

All three Experiments

We first analyzed the data by keeping all the experiments separate, and we measured the mean Inverse Efficiency Score for each condition in each Experiment and Task instruction (Figure 2.9).

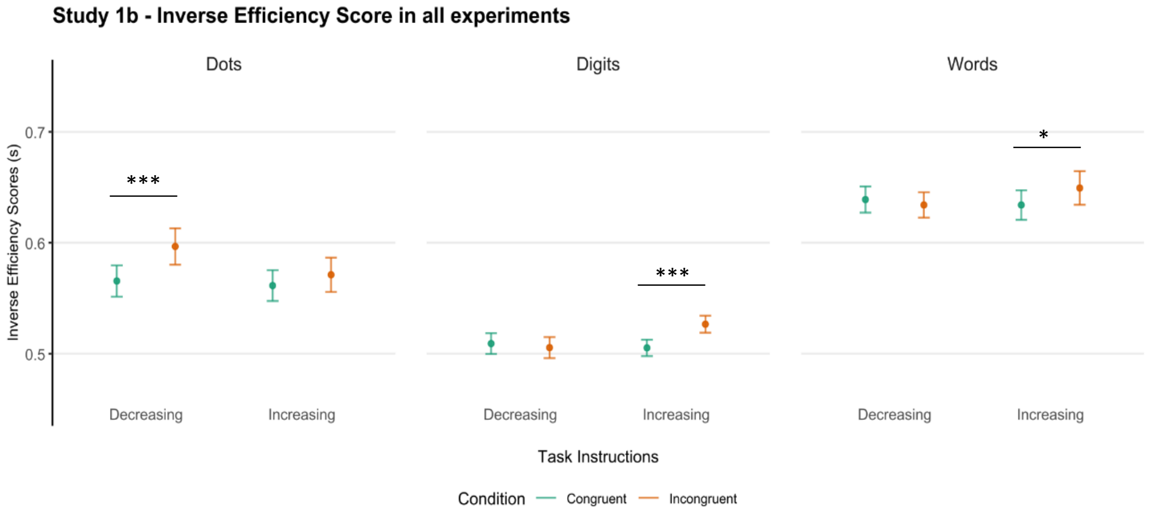


Figure 2.9: Inverse Efficiency Scores for all stimulus modalities
Inverse Efficiency Scores were measured for each Congruency condition for each Experiment and each Task instruction. In the graph mean value, standard error of the mean, and significant pvalues are reported for each Experiment, Task instruction, and Congruency condition (two-tailed Wilcoxon rank sum test; see Table 2.5).
Significance levels are defined as follows: * = pvalue <.05, *** = pvalue <.001)

The 3x2x2 Anova on the Inverse Efficiency Score revealed a main effect of Experiment, Congruency condition, and a significant interaction between the Experiment and Congruency condition (Table 2.4).

Three-Way Analysis of Variance of Inverse Efficiency Scores

Factor	Df_n	Df_d	SS_n	SS_d	*F-value*	p_{value}	$\eta^2_{partial}$	η^2_p *CI 90%*
Experiment	1.53	42.74	0. 943	0.386	68.347	**<.0001**	0.709	[0.567 0.778]
Task instruction	1	28	0.0005	0.109	0.138	.713	0.005	[0 0.109]
Congruency Condition	1	28	0.010	0.009	31.092	**<.0001**	0.526	[0.289 0.657]
Experiment *x* Task instruction	2	56	0.016	0.135	3.231	**.047**	0.103	[0.001 0.219]
Experiment *x* Congruency Condition	2	56	0.003	0.034	2.270	.113	0.75	[0 0.182]
Task instruction *x* Congruency Condition	1	28	0.002	0.022	3.136	.087	0.101	[0 0.284]
Experiment *x* Task instruction *x* Congruency Condition	1.52	42.47	0.006	0.052	3.402	0.55	0.108	[0 .25]

Table 2.4
Three-Way Analysis of Variance of Inverse Efficiency Scores for each Experiment by Task instruction and Congruency condition. Greenhouse-Geisser correction for the violation of the sphericity assumption has been applied. Significant pvalue for .05 significance level in bold

Given the marginal significant triple interaction of Experiment*Task instruction*Congruency Condition, after the Greenhouse-Geisser correction method, we run explorative pairwise comparisons between congruent and incongruent congruency conditions for each Experiment and Task instruction (Table 2.5).

Explorative pairwise comparison - Two-tailed two-sample paired Wilcoxon Signed rank test

Group	*Task*	*Condition 1*	*Condition 2*	*Df*	*t-value*	*95% CI*	p_{value}	*Effect size*
Dots	Decreasing	Congruent	Incongruent	40	180	[-0.046 -0.012]	**<.001**	0.507
	Increasing	Congruent	Incongruent	40	314	[-0.023 0.003]	.134	0.236
Digits	Decreasing	Congruent	Incongruent	28	278	[-0.004 0.015]	.198	0.243
	Increasing	Congruent	Incongruent	28	60	[-0.033 -0.011]	**<.001**	0.632
Words	Decreasing	Congruent	Incongruent	28	248	[-0.01 0.019]	.522	0.122
	Increasing	Congruent	Incongruent	28	114	[-0.025 -0.003]	**.024**	0.416

Table 2.5
Inverse efficiency scores comparison for differences between Congruency condition for Task instruction for each Experiment. Significant pvalue for .05 significance level in bold

As for the previous study we also tested the congruency effect. The 3x2 Anova revealed no main effects of Task ($F_{(1,28)} = 2.776$, $p_{value} = .107$, $\eta^2_p = 0.090$, 90% CI = [0 0.271]) or Experiment ($F_{(2,56)} = 2.230$, $p_{value} = .117$, $\eta^2_p = 0.074$, 90% CI = [0 0.18]), while a significant interaction between the two factors emerged ($F_{(1.52,\ 42.55)} = 4.026$, $p_{value} = .035$, $\eta_p^2 = 0.126$, 90% CI = [0.005 0.27]). Results are shown in Figure 2.10.

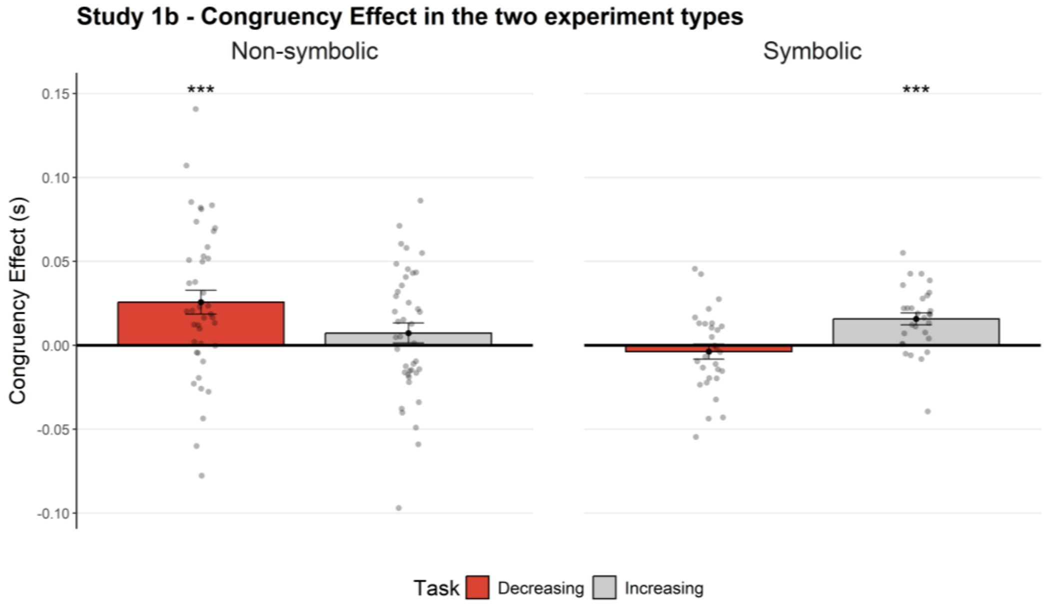


Figure 2.10: Congruency effect for each Task instruction for all stimulus modalities
We measured participants' performance in terms of IES, and we calculated the Congruency Effect for each Task instruction. In the graph mean value, standard error of the mean, and significant pvalues are reported for each Experiment and Task instruction (Two-tailed Wilcoxon signed rank test; Decreasing condition (red): Dots: N = 41, p <.001, Digits: N = 29, p <.001, Words: N = 29, p = .022; Increasing condition (grey): Dots: N = 41, p = .193, Digits: N = 29, p = .205, Words: N = 29, p = .455; see Table 2.6). Significance levels are defined as follows: * = pvalue <.05, *** = pvalue <.001)

Comparison against chance level revealed a significant congruency effect for the decreasing Task instruction in the Dots experiment only. On the other hand, a significant congruency effect for the increasing Task instruction was found in the Digits and Words experiment (Table 2.6).

*Two-tailed one-sample Wilcoxon signed rank test against chance level (μ = 0)**

Experiment	*Task*	*Descriptive statistic*			*Test statistic*				
		N	*Mean*	*SD*	*Df*	*t-value*	*pvalue*	*95% CI*	*Effect size*
Dots	Decreasing	41	0.026	0.007	40	681	**<.001**	[0.012 0.04]	0.507
	Increasing	41	0.007	0.006	40	532	.193	[-0.005 0.02]	0.205
Digits	Decreasing	29	-0.004	0.005	28	158	.205	[-0.015 0.005]	0.239
	Increasing	29	0.021	0.005	28	376	**<.001**	[0.011 0.032]	0.636
Words	Decreasing	29	-0.004	0.006	28	182	.455	[-0.015 0.007]	0.143
	Increasing	29	0.011	0.005	28	323	**.022**	[0.003 0.019]	0.424

Table 2.6
Statistical test against chance level for each Experiment for both Task instructions

We then analyzed the congruency effect as a function of the test stimulus, instead of general Task instruction. The 3x8 Anova revealed a main effect of Test stimulus ($F_{(7,154)} = 5.961$,

p_{value} = .001, η_p^2 = 0.213, η_p^2 90% CI = [0.094 0.272]), but no significant effect of Experiment ($F_{(2,44)}$ = 1.862, p_{value} = .167, η_p^2 = 0.078, ηp^2 90% CI = [0 0.198]). A significant interaction was found ($F_{(6.33, 139.29)}$ = 2.087, p_{value} = .012, η_p^2 = 0.087, ηp^2 90% CI = [0 0.129]). Results are shown in Figure 2.11.

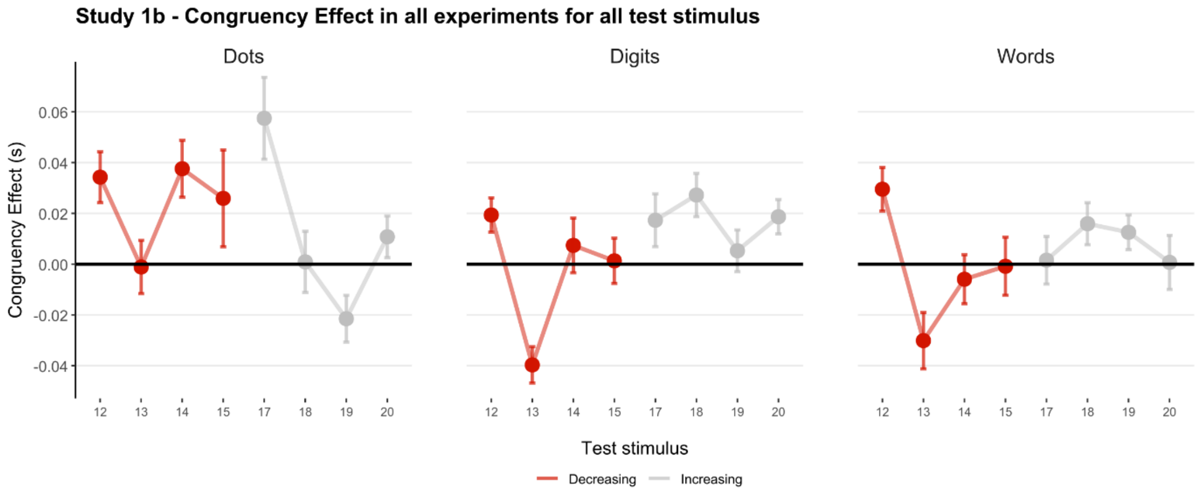


Figure 2.11: Congruency effect for all test stimuli
Congruency Effect for each Test stimulus numerosities for Each experiment. The means and standard errors of the mean are reported.

Symbolic vs Non-Symbolic

We collapsed the data from Digits and Words experiments and measured the Inverse Efficiency scores for symbolic and non-symbolic experiments (i.e. Type) (Figure 2.12)

Figure 2.12: Inverse Efficiency Score for congruent vs incongruent conditions in each Task instruction for symbolic and non-symbolic experiments
Inverse Efficiency Scores were measured for each Congruency condition for each Task instruction for each Type of experiment. In the graph mean value, standard error of the mean, and significant pvalues are reported for each Experiment, Task instruction and Congruency condition (two-tailed Wilcoxon rank sum test; see Table 2.8).
Significance levels are defined as follows: * = pvalue <.05, *** = pvalue <.001)

When collapsing the Inverse Efficiency scores of Digits and Word experiments, the 3x2x8 Anova revealed a main effect of congruency condition, a significant interaction between Type and Task instruction, and a significant interaction between Type, Task instruction, and Congruency condition (Table 2.7).

Three-Way Analysis of Variance on Inverse Efficiency Score

Factor	*Df_n*	*Df_d*	*SS_n*	*SS_d*	*F-value*	*p_{value}*	*$\eta^2_{partial}$*	η^2_p CI 90%
Type	1	28	0.001	0.221	0.061	.807	0.002	[0 0.084]
Task instruction	1	28	0.003	0.103	0.817	.713	0.028	[0 0.178]
Congruency Condition	1	28	0.010	0.011	24.535	**<.001**	0.467	[0.225 0.613]
Type *x* Task instruction	1	28	0.012	0.055	5.905	**.022**	0.174	[0.014 0.363]
Type *x* Congruency Condition	1	28	0.002	0.017	3.250	.822	0.174	[0 0.287]
Task instruction *x* Congruency Condition	1	28	0.001	0.017	0.495	.487	0.017	[0 0.154]
Type *x* Task instruction *x* Congruency Condition	1	28	0.005	0.029	4.500	**.043**	0.138	[0.002 0.326]

Table 2.7
Three-Way Analysis of Variance of Inverse Efficiency Scores for each Type by Task instruction and Congruency Condition. Significant pvalue for .05 significance level in bold

The pairwise comparisons revealed a significant difference between congruent and incongruent conditions in decreasing task instructions for the non-symbolic stimuli and between congruent and incongruent conditions in the increasing task instructions for the symbolic stimuli (Table 2.8).

Two-tailed two-sample paired Wilcoxon Signed rank test

Type	*Task*	*Condition 1*	*Condition 2*	*Df*	*t-value*	*95% CI*	*p_{value}*	*Effect size*
Non-symbolic	Decreasing	Congruent	Incongruent	40	180	[-0.046 -0.012]	**<.001**	0.507
	Increasing	Congruent	Incongruent	40	314	[-0.023 0.003]	.134	0.236
Symbolic	Decreasing	Congruent	Incongruent	28	255	[-0.007 0.016]	.429	0.151
	Increasing	Congruent	Incongruent	28	50	[-0.027 -0.011]	**<.001**	0.673

Table 2.8
Comparison for differences between Congruency Condition for Task instruction for no-symbolic and symbolic stimuli. Significant pvalue for .05 significance level in bold

The 2x2 Anova on the congruency effect revealed no main effects of Type ($F_{(1,28)} = 3.038$, $p_{value} = .092$, $\eta_p^2 = 0.098$, η_p^2 90% CI = [0 0.280]) or Task ($F_{(1,28)} = 0.29$, $p_{value} = .589$, $\eta_p^2 = 0.011$, η_p^2 90% CI = [0 0.133]), while a significant interaction between the two factors emerged ($F_{(1,28)} = 4.736$, $p_{value} = .038$, $\eta_p^2 = 0.145$, η_p^2 90% CI = [0.004 0.033]). Results are shown in Figure 2.13.

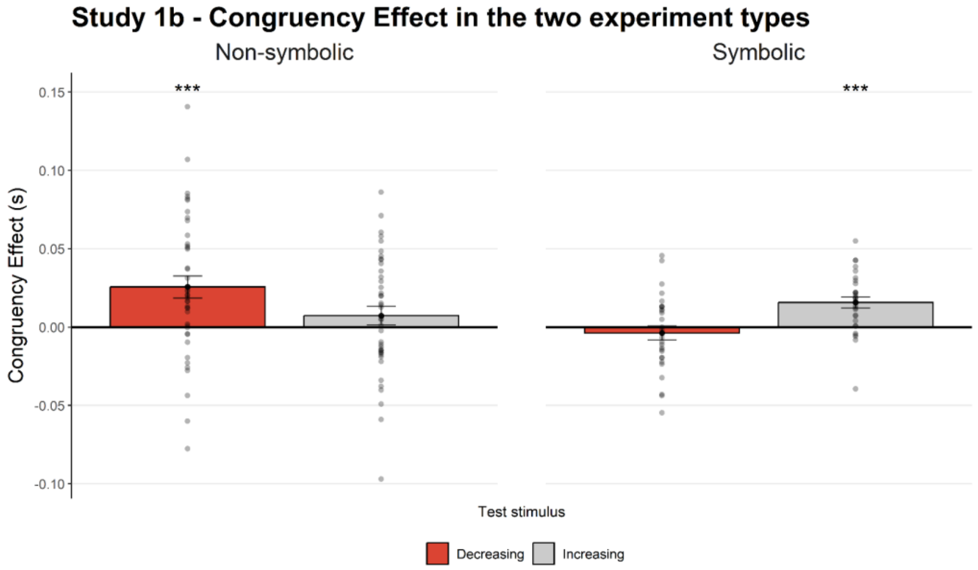


Figure 2.13: Congruency effect for Non-symbolic and Symbolic numerosities
We measured participants' performance in terms of IES, and we calculated the Congruency Effect for each Task instruction. In the graph mean value, standard error of the mean, and significant pvalues are reported for each Type and Task instruction (Two-tailed Wilcoxon signed rank test; Decreasing condition (red): Non-symbolic: N = 41, p <.001, Symbolic: N = 29, =.393; Increasing condition (grey): Non-symbolic: N = 41, p =.193, Symbolic: N = 29, p <.001; see Table 2.9). Significance levels are defined as follows: *** = pvalue <.001)

Comparison against chance level revealed a significant congruency effect for the decreasing Task instruction for non-symbolic stimuli. A significant congruency effect was found for the increasing Task instruction for symbolic stimuli instead (Table 2.9).

Two-tailed one-sample T-test against chance level (μ = 0)

Type	*Task*	*Descriptive statistic*			*Test statistic*				
		N	*Mean*	*SD*	*Df*	*t-value*	*95% CI*	*p_{value}*	*Effect size*
Non-symbolic	Decreasing	41	0.026	0.007	40	681	[0.012 0.040]	**<.001**	0.507
	Increasing	41	0.007	0.006	40	532	[-0.005 0.020]	.193	0.205
Symbolic	Decreasing	29	- 0.004	0.006	28	177	[-0.013 .006]	.393	0.163
	Increasing	29	0.011	0.005	28	385	[0.009 .023]	**<.001**	0.673

Table 2.9
Statistical test against chance level for each Type of experiments for each Task instruction.
Significant pvalue for .05 significance level in bold

We then analyzed the congruency effect as a function of the test stimulus, instead of general Task instruction. The 2x8 Anova revealed a main effect of Test Stimulus ($F_{(4.44, 97.63)} = 4.609$, $p_{value} < .001$, $\eta_p^2 = 0.173$, η_p^2 90% CI = [0.046 0.253]), but no significant effect of Type ($F_{(1,22)} = 2.682$, $p_{value} = .115$, $\eta_p^2 = 0.109$, η_p^2 90% CI = [0 0.314]). A significant interaction was also found ($F_{(3.79, 83,33)} = 2.826$, $p_{value} = .032$, $\eta_p^2 = 0.114$, η_p^2 90% CI = [0.005 0.194]). The results are shown in Figure 2.14.

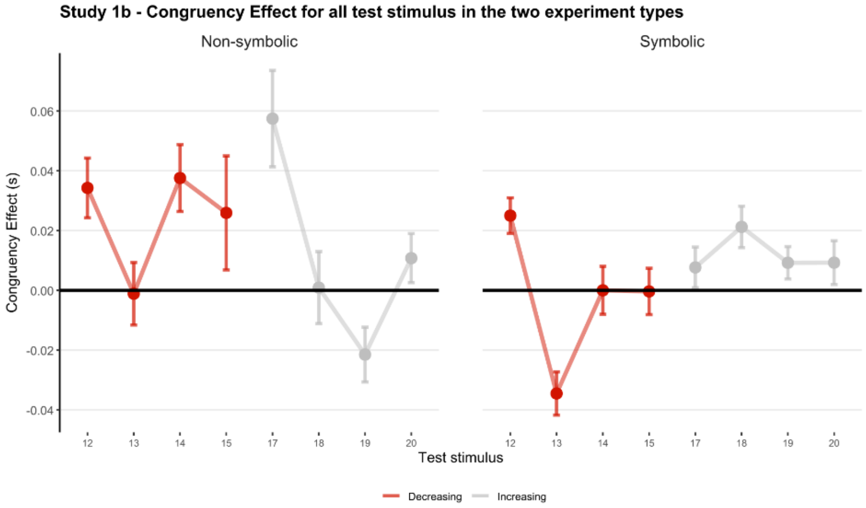


Figure 2.14: Congruency Effect for all test stimuli for Non-Symbolic and Symbolic numerosities
Congruency Effect for each Test stimulus numerosities for non-symbolic and symbolic modality.
The means and standard errors of the mean are reported.

2.3.4. Discussion

In this study (comprising three experiments), we further investigated the role of stimuli nature, symbolic vs non-symbolic, on SNAs that emerge in a Numerosity comparison task. Literate Italian adults performed the same experiment with stimuli presented in three different visual modalities, arrays of dots, Arabic format, and numeric words. Moreover, we manipulate the numerical distance between reference and test stimulus to further explore the role of task difficulty. The currently unbalanced samples of the study, 40 participants in the Dots experiment and 29 participants in the Digits and Words experiments, are to all probability responsible for some of the results being only marginally significant. Therefore, we will now discuss these preliminary results with caution, although the overall pattern shows strong internal coherence (as demonstrated by linear result patterns when symbolic experiments are collapsed or not).

In this experiment, we found different patterns of responses according to stimulus modality. Participants had better performance for decreasing change in numerosities on the left in the non-symbolic task and better performance for increasing change in numerosities on the right in the symbolic tasks. Overall, the SNA appears to be modulated by the nature of the stimuli. The development and the hemisphere specialization of the ANS, see paragraph 1.1.3, could explain these findings. As already mentioned, the non-symbolic versus symbolic nature of the stimuli seems to engage differently the brain hemispheres, as they process different kinds of numerical representation. Coherently with the results of the Numerosity comparison task in the first study (Figure 2.6), we found that when presented with a set of dots participants have a greater advantage in the decreasing task. Instead when presented with digits and words they present the opposite pattern, i.e. a greater advantage in the increasing task (Figure 2.10). These results seem to suggest that in the two task instructions*,* the two hemispaces are monitored differently according to the nature of the stimuli used. The left hemispace seems to stand in a privileged position in the decreasing task for non-symbolic stimuli only, while the right hemispace seems to stand in a privileged position in the increasing task for symbolic numerosities only. As we already saw, it is not new in literature to find a left hemispace advantage for non-symbolic stimuli and a right hemispace advantage for symbolic stimuli. This could be determined by the different hemispheric involvement of the IPSs (Figure 1.3). Given the high presence of fibers that interconnect frontal, parietal, and temporal areas involved in language and semantic processing (for example, the Arcuate fasciculus; Catani & Thiebaut de Schotten, 2008), a symbolic representation may lead to a generally increased activation of the left IPS (given the left-lateralized location of symbolic/language mechanisms), which in turn would cause greater activation of the left hemisphere as a whole, resulting in activation of the right hemispace (due to the fibers decussation happening at the level of the optic chiasm). Similarly, a non-symbolic representation might induce the same effect in the right hemisphere (given the right-lateralized location of numerosities processing), leading to subsequent activation of the left hemispace.

Additionally, the difficulty of the task, determined by the numerical distance between the reference stimulus and the test stimulus, could influence responses differently depending on the task and the modality of stimulus presentation (in line with the known *distance effect*). Unlike in Study 1a, the non-symbolic task in this case produced a congruence effect across both task instructions. We hypothesize that this may be due to the perceived difficulty of the task. As shown in Figure 2.11, in the Dots experiment the condition of greatest difficulty in the increasing task seems to elicit a pronounced congruency effect, which gradually diminished in subsequent comparisons. This suggests that SNA might also be related to the ratio effect and distance effect, which are typical of ANS processing: in Study 1a, the stimuli pairs 4-12 and 12-36 were much more distant and separated by a larger ratio compared to the pair 16-17 tested in Study 1b. However, it seems that except for the 16 vs 17 comparisons, the congruency effect is less pronounced in the increasing task instruction overall. To our knowledge, no studies have yet systematically investigated a potential dissociative effect between increasing and decreasing tasks in relation to the SNA. If there are evolutionary advantages behind the development of the ANS, it could be speculated that our minds have evolved to be more sensitive to increases rather than decreases as suggested by the study conducted by Macchi Cassia and colleagues on 4-month-old infants (discussed in paragraph 2.2.4). This kind of asymmetrical sensitivity could be evolutionarily advantageous, enabling quicker responses to situations where there is an increase in food, social and sexual mates, or even predators. As we will deeply discuss in the next study domestic chicks for example show stronger predispositions to approach larger groups of social companions (Pulliam, 1973; Roberts, 1996). We hypothesize that a more general sensibility to increasing change in numerosities might overrule SNA in conditions where numerical differences are easier to perceive, for example as a possible consequence of faster reaction times that could be anticipating the manifestations of SNA. As first reported in 2006 by Gerves and colleagues, and then replicated and revised by Cipora and colleagues in 2019, the strength of the SNARC effect seems to correlate with the length of the RTs (Cipora, Soltanlou, et al., 2019; Gevers et al., 2006). This could also explain why only the hardest comparison in the increasing task elicits a congruency effect in our task. However, the results from the symbolic experiments seem to contradict the hypothesis of a direct correlation between RTs and SNA strength. Indeed, when Arabic digits were used, participants showed at the same time the best performance and a strong SNA behavior in the increasing task. Results of the Words experiment, instead, show the opposite pattern: participants had the worst performance and the weaker SNA behavior, despite being the one with the longest RTs. In our studies, we used an extended response time window, initially to facilitate data collection among Himba adults and Italian preschoolers, and later maintained also for Italian adults for replication consistency. This may have led to the collection of longer response times than those typically observed in other paradigms. It is then possible that the relationship between task difficulty (and RTs as consequence) and SNA is not linear, but rather follows a U-shaped curve, similar to the arousal-performance relationship. At low arousal levels, individuals may experience diminished motivation and engagement, resulting in suboptimal performance. As arousal increases to a moderate, optimal level, performance generally improves, supported by heightened alertness and focus. However, when arousal exceeds this optimal threshold performance tends to deteriorate. This U-shaped curve suggests that both insufficient and excessive arousal can impair performance, with optimal functioning occurring within a specific intermediate arousal range (Yerkes & Dodson, 1908). Similarly, there may be a specific window of time within which the response needs to occur to let SNA phenomena emerge. Moreover, digits have a stronger intrinsic culturally-dependent association with space, given the explicit teaching of a “number line” at school. This intrinsic acquired mapping that could support culturally-driven SNA even in the presence of short reaction times. Number words, instead, are less often associated with space in daily life, as it is less common to represent numbers in words format. Therefore, it is possible that association between number words and space requires additional cognitive steps, that however might fall outside the ideal time window for SNARC effect to emerge. Future research is needed to shed light on this hypothesis and to disentangle the complex relationship that stimuli features and tasks' demands have on SNA. For example, studies exploiting a Delayed Matching-to-Sample could provide insightful information on the timing occurrence of SNA phenomena. Additionally, for stimulus "13," there appears to be no SNA effect in the Dots task, and a reversed SNA effect in both the Digits and Words tasks. In the Digits task, this reversal could be attributed to the perceptual similarity between "13" and "18," combined with the previously discussed response advantage for increasing numerosity. This resemblance may have led participants to misinterpret "13" as "18," thus consistently favouring right-side responses, resulting in better performance on that side. Additional studies and further investigation are needed to explain why this effect appears to occur even in the absence of perceptual similarity.

In conclusion, the preliminary results of this study seem to suggest, once again, that SNAs are the intricate result of both culturally dependent and biologically dependent influences. The left-to-right SNA in Italian adults seems indeed to manifest differently in response to increasing and decreasing changes in numerosities and as a function of the stimuli format used in the experiment. Moreover, we found possible evidence of more complex features of SNA as a dependency on specific task difficulty or time response windows.

2.4. General Discussion

Study 1 aimed to shed light on the influences that both the cultural background context of participants and task features can have on the emergence of SNA responses. In Study 1a we demonstrated that when we tested different populations subjected to different levels of cultural exposure it was possible to dissociate two different SNA patterns within the same group. Italian adults indeed showed left-to-right oriented forms of SNA in both explicit and implicit tasks, as we were expecting. On the other hand, both Himba adults and Italian preschoolers manifested different SNAs between the two tasks: at the group level, they had no preferred SNA in the explicit task, while they showed a significant left-to-right bias in the implicit task. As we already discussed in paragraph 2.2.4, we believe that this difference between the two populations is determined by their relative cultural background. When explicitly asked to map numerosities into space Italian adults rely on explicit, formal, and symbolic knowledge that is easily available to them: numbers start on the left and they grow on the right, as it is taught in school. Instead, this type of knowledge is not available to Himba adults and Italian preschoolers and therefore they do not show a consistent bias in direction. It is important to note, however, that they spontaneously choose to order the numerosities based on their magnitude values. These results prove that sequential ordering of numerosities on a lateral disposition is a possible innate predisposition, while the specific order is culturally determined, at least on this kind of task. Instead, when implicitly tested, all the populations can rely on other neural mechanisms routed in the brain with no apparent difference between them. Those mechanisms are spontaneously activated by the processing of non-symbolic numerosities in an implicit task, leaving no need to recruit external information or activate a parallel mechanism as the one related to symbolic knowledge.

The results from Study 1b seem to further support this hypothesis. As already mentioned, we designed Study 1b because we were initially surprised by the apparent asymmetrical effect between the two task instruction conditions. As possible to observe in Figure 2.6, the SNA results are prominent only in the decreasing tasks, i.e. there is a stronger effect for smaller numerosities on the left than for larger numerosities on the right. As we discussed we believe that this could be related to the non-symbolic nature of the study and the hemispheric specialization of the ANS. The results from Study 1b seem to further support this hypothesis. Indeed, we demonstrated that when manipulating one of the levels of implicitness of the task (i.e. symbolic vs non-symbolic nature of the stimuli) there is a shift in the asymmetrical effect. The symbolic nature of the stimuli behaves in the opposite way of the non-symbolic one: there is a plausible advantage of the right hemispace as a result of a more prominent engagement of the left hemisphere. Interestingly enough, this does not overrule the spontaneous left-to-right direction of innate SNA, as we do not find an overall advantage of one hemispace over the either. These findings about a spontaneous left-to-right SNA are in line with the evidence coming from the non-human studies already discussed (see paragraph 1.2.2). It is then plausible that SNA phenomena in humans are the complex result of more mechanisms happening in parallel. As mentioned in the introduction, digits and number words may introduce varying levels of "explicitness," even within an implicit task, thereby activating additional cognitive processes. Further research is needed to clarify if and how a continuum might exist between explicit and implicit task characteristics, considering not only the paradigm but also the types of stimuli used. The results are also in line with symbolic and non-symbolic stimuli led to variables SNA effect across different tasks. While some studies found clear evidence of SNA effects in response to non-symbolic stimuli (Nemeh et al., 2018; H.-C. Nuerk et al., 2005) other seems to contradict those findings (Cleland et al., 2020; Prpic et al., 2023). As reviewed by Buijsman and Tirado (2019), SNARC has been used as a proxy to try to understand whether the representation underling the ANS is shared between symbolic and non-symbolic processing of numbers (as suggested by the Triple-code model; (Dehaene, 1992) or not (as suggested by Sensory Integration System; Gebuis et al., 2016). In the review, the authors conclude that neither of the two positions can be strongly supported by the behavioral data available at that time. As we mentioned, contrasting findings are still present at this date and we believe that our studies could help to further shed light on this intricated phenomenon. The dissociation between symbolic and non-symbolic experiments performed with the same task could for example be in future investigated with neuroimaging techniques (as also suggested by Buijsman and Tirado) to better understand the neural networks involved in each task.

But then why the spontaneous implicit (non-symbolic) association should be specifically left-to-right oriented? The general lateralization of the two brain hemispheres could stand at the origins of the direction of this type of SNA. Indeed, SNA is not the only lateralized behavior in human and non-human animals: several asymmetries between left and right have been reported both in the two hemispheres of the brain and behaviors (see Rogers et al., 2013 for a review). Among these, there are the approach and withdrawal behaviors: approach behavior appears to be associated with the left hemisphere while withdrawal behavior is with the right one (Davidson, 1992). This existent brain asymmetry leads to the so-called Valence hypothesis: there is general control of functions related to motivation and emotion that is differentiated in the two hemispheres (Davidson, 2004). According to this theory, the hemispheres of the brain are activated in response to stimuli of opposite value: positive stimuli are preferentially processed by the left hemisphere, and negative stimuli by the right hemisphere. Given the decussation of the nerves, the result is a spontaneous association between positive stimuli and the right hemispace and negative stimuli and the left hemispace. Evidence in favor of this theory has been found both in human and non-human animals (Rogers et al., 2013) so it has been hypothesized that the implicit SNA small/right and large/left may have emerged as a consequence of a more general cognitive mechanism that associates a change towards more numerous stimuli with positive valence, (i.e. larger is better), which is processed by the left hemisphere of the brain and hence results in association with the right hemispace, and a change towards less numerous stimuli with negative valence, which is processed by the right hemisphere of the brain and hence results in association with the left hemispace, (Vallortigara, 2018). Therefore one possibility is that the valence of the stimuli (positive/neutral or negative) might influence this implicit SNA effect. Indeed, generally speaking, the majority of the studies present in the literature have been carried on with stimuli with a clear positive valence (e.g. the association with the reinforcement used for animals' studies) or at best a neutral value. This means that till now it has been virtually impossible to discern the association of both decreasing/bad with the left hemispace and increasing/good with the right hemispace. This hypothesis holds also an adaptive meaning for SNA. It could be that a spontaneous association between large and right, due to a left hemisphere activation, is the result of the evolutionary need to quickly approach the best possible situation (e.g. more food, more social companions, etc) in the presence of positive stimuli. The same is true if the reverse direction of SNA is proven in response to negative stimuli: this high adaptive value would help to rapidly adapt and make the association "more is bad" resulting in large numerosities associated with the left hemispace, due to a right hemisphere activation.

In conclusion, attentive dissociation when analyzing explicit and implicit SNA phenomena can lead to a better understanding of the origins of the phenomenon. There is no doubt that explicit forms of SNA effects are influenced by the cultural context, especially in relation to the direction of reading and writing habits. However, it seems that for implicit forms of SNA, the same assumption does not hold. Indeed, evidence coming from pre-verbal children and animal studies seems to suggest a more general shared system for SNA. In this study, we demonstrated that also humans show different SNA patterns as a function of their culture exposure (Study 1a) and that the symbolic versus non-symbolic nature of the stimuli significantly influences SNA responses (Study 1b). New experiments with novel designs are needed to better explore the intriguing relationship of all these cognitive mechanisms, with particular interest in the role of stimuli valence. Moreover, in line with the arguments presented in the introduction (paragraph 1.2.1), it is important to consider that these studies alone do not definitively establish the extent to which reading and writing direction specifically influences SNA. In Italian culture, both reading and writing, as well as other cultural preferences, predominantly favor a left-to-right orientation. This overlapping of factors suggests that further research is necessary to disentangle the specific contributions of reading and writing direction from other cultural influences on SNA. Future studies should aim to clarify these relationships to enhance our understanding of how various factors shape spatial-numerical processing across different cultural contexts.

Chapter 3

Study 2

Investigating numerical cognition through auditory stimulation in domestic chicks

3.1. General Introduction

Domestic chicks (*Gallus gallus domesticus*) are able to represent and manipulate numerosities. As discussed in the thesis introduction, there are several pieces of evidence that numerical competencies are present in non-human animals, specifically in domestic chicks (see Rugani, 2018 for a review). One of the abilities at the basis of numerical cognition is to perform numerosity discriminations, such as "*larger/smaller than*" comparisons (Vallortigara et al., 2010). Domestic chicks have been shown to perform numerosity discriminations both when presented with "small" numerosities, for example, 2 versus 3 elements (Rugani et al., 2008), and when presented with "large" numerosities, for example up to 8 versus 12 (Rugani et al., 2014). Rugani and colleagues (2016) demonstrated that chicks' ability to discriminate numerosities might also be supported by a precocious ability to represent proportional information, learning to generalize over comparisons of equal ratios. This discrimination ability does not require specific training to be acquired; rather it seems that chicks can spontaneously learn numerical information and use it to solve subsequently presented tasks. This has been shown, for instance in paradigms exploiting imprinting, in which chicks reared with a specific number of objects show, at test, preference for the same numerosities (Lemaire et al., 2020; Rugani, Regolin, et al., 2010). Young domestic chicks can also manipulate numerosity representations to the point that they solve tasks by adding or subtracting elements (Rugani et al., 2009; Rugani, Regolin, et al., 2011). For example, in a study by Rugani and colleagues, newly hatched chicks were raised with five identical objects. On the third day of life, chicks were placed in an arena and watched as two objects were hidden behind one opaque panel and three objects behind another one. The chicks approached the panel that concealed the larger number of objects. This ability was maintained also when chicks saw sequences of events representing arithmetic operations, i.e. movements of the objects from one panel to the other, thus subtracting or adding objects to each set (Rugani et al., 2009). Another aspect of numerical cognition mastered by chicks is the ability to identify a target element based on its ordinal position (Rugani, Kelly, et al., 2010; Rugani, Vallortigara, et al., 2011, see also paragraph 1.2.2).

However, despite having been the object of a relatively large number of studies, until now chicks' numerical abilities have been investigated only in the visual domain. There is evidence that other animal species can discriminate auditory numerical information. For example, untrained cotton-top tamarin monkeys (*Saguinus oedipus*) discriminated discriminate sequences composed by different numbers of syllables even when controlled for continuous variables confounds, such as duration, inter-stimulus interval, and overall energy (Hauser et al., 2003). Moreover, they showed a ratio effect in numerosity discrimination similar to the human infants: as tamarins, also 9-month-old infants discriminate auditory sequences with a ratio threshold between 1.5 and 1.25 (Lipton & Spelke, 2004). The ability to process numerical information in an auditory format does not seem limited to primates. For example, rats that were trained to respond to the presentation of three bursts of white noise, at the test showed a stronger response to that specific numerosity and reduced responses to two or four bursts (Davis & Albert, 1986). This indicates a precise numerosity discrimination, more sophisticated than a basic “many-versus-few” comparison. The ability to infer numerical information from sounds has been associated with aggressive behaviors, as animals must evaluate the potential advantages of attacking another group. For instance, lions have been shown to adjust their aggressive behavior based on the number of roars they hear, using these sounds as a proxy for numerical estimation of the rival group (McComb et al., 1994).

Auditory discrimination has been investigated in other cognitive domains in chicks. For example, it has been demonstrated that chicks can be imprinted on and later recognize different tones already during the prenatal stage (Grier et al., 1967), or show a spontaneous preference for consonant over dissonant melodies (Chiandetti & Vallortigara, 2011). However, to our knowledge, there are no studies on chicks' ability to differentiate numerosities presented through sound. Therefore in this study, we aimed to test whether newborn chicks can discriminate auditory sequences of different numerosities (Experiment 1) and if they can learn numerical information through an imprinting paradigm (Experiment 2).

3.1.1. Ethics and Animals

The study reported in this thesis has been carried out in compliance with the European Union and Italian law on the treatment of animals. The experiments and the experimental procedures were approved by the Ethical Committee of the University of Trento and licensed by the Italian Health Ministry (permit number 831/2021). We utilized domestic chicks from the Aviagen Ross 308 strain, chosen for their sexual dimorphism evident in feather patterns at birth, facilitating sex determination. Sexing was done immediately after testing, by observation of the wing feather patterns, as there are reported differences in behavioral responses between male and female chicks across several studies (Lemaire et al., 2020; Rosa-Salva et al., 2023; Santolin et al., 2020; Vallortigara & Zanforlin, 1988). Eggs were sourced from a commercial hatchery (CRESCENTI Societa` Agricola S.r.l. –AllevamentoTrepola– cod. Allevamento 127BS105/2) and were subsequently incubated in complete darkness in our laboratory. From day 1 to day 19, eggs were maintained at 37.7°C and 40% humidity, then increased to 60% humidity until day 21 (hatching day). All subjects hatched in the dark and were kept in the incubators until the test (all chicks were visually naïve at the moment of the test). In all the experiments described, each animal completed a single trial and they were returned to the animal house facility at the end of the test day.

3.2. General Methods

The next paragraphs will describe the setup and the procedure employed in the first two experiments, which differed only for the stimuli features (see Stimuli paragraph for each study).

3.2.1. Apparatus

The setup used to test animals was a “Y-maze”, whose walls were covered with sound-absorbing material. The apparatus was virtually divided into three different areas: the central area, the left corridor, and the right corridor. On the floor, a fine pencil line delimited the choice areas, in the two arms of the Y-maze (Figure 3.1). The short wall at the end of each choice area had a built-in speaker (Z130 Stereo Speakers) for audio reproduction. The shape of this apparatus allowed us to present two stimuli frontally to a chick standing in the central area. A webcam camera (Logitech, C922 Pro, HD stream webcam) was placed above the apparatus to record the chicks’ movement for the whole test duration. Two led strips were placed above each end of the apparatus, illuminating a small red cylinder placed at about the chick’s head height. The red cylinder was hung on the ceiling of the apparatus by a translucent thread and it acted as an attractor for the chicks, which during the first days after hatching are attracted by small visually salient objects.

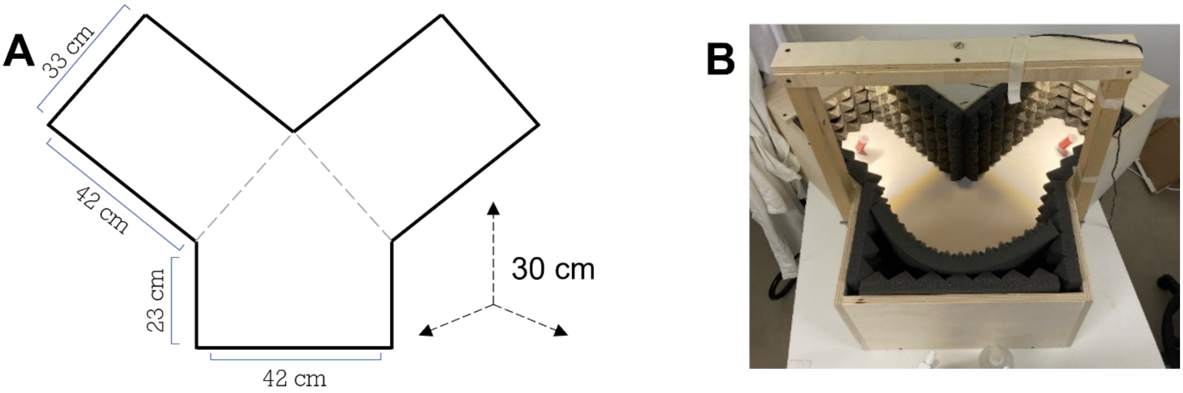


Figure 3.1: Apparatus used for all experiments 1,2, and 3
Panel A) Schematic representation of the apparatus with measures. The dotted lines represent the fine pencil lines delimiting the choice area.
Panel B) Picture of the apparatus, aerial view with stimuli used in experiments 1, 2 and 3

3.2.2. Stimuli Generation

Acoustic stimuli[2] consisted of multiple repetitions of identical notes/tones (see specific Stimuli paragraph for each experiment). Each sequence has been generated in version 2.0.5 of Audacity® software and then exported in a .mp3 or.wav file reproducible with VLC software (VideoLan, 200, VLC media player). Audio sequences were retrieved from VLC software during the test phase.

3.2.3. Test procedure

In the following experiments, different chicks were tested either on the same day of hatching (from now, P0) or on the day after hatching (P1). Crucially each chick was tested once, so the design is fully between subjects, and irrespective of the day of the test the procedure was the same. On the day of the test, chicks were individually carried from the dark incubator to the test apparatus inside a closed, opaque box ensuring that they were visually naïve until the beginning of the experiment. The test room was kept dark, except for the light coming from the apparatus itself. Each animal was singly placed in the central area of the apparatus, where it was confined

---

[2] Audio stimuli are available at the following GitHub repository https://github.com/elenaeccher/NumericalCognitionChicks.git

in a small black plastic cage (Figure 3.2). The front of the cage was made of a metal wire mesh, allowing the chick to see the attractor stimuli and ensuring that the sound could easily reach the animal. At the beginning of the test, the animal was confined for 1 minute in the cage (exposition phase), where it could observe the apparatus in which the stimuli were played and listen to the auditory sequences. This was done to ensure that the chick could attend to each of the two auditory stimuli before making a choice (e.g., to avoid an impulsive approach to the first stimulus played). After this exposition phase, the experimenter manually lifted the cage and the animal could move freely in the apparatus. Each test phase lasted 6 minutes (after the cage was lifted), no matter the response of the animal (i.e. the test did not end at the first choice made). During this phase, chicks were free to move around the apparatus and approach the end of the apparatus where the preferred stimulus was played. During both the exposition and the test phase, the animals listened to a continuous stream of the two different auditory sequences (see Stimuli paragraph for each study for a detailed description of the sequence), played asynchronously in alternation between the left and right speaker. For instance, an animal would hear the 4-sound sequence coming from the left followed by a 5 seconds silent pause, and then the 12-sound sequence coming from the right speaker, followed by another 5 seconds silence pause. This loop was repeated seamlessly for the whole duration of both the exposure and test phases (1 minute + 6 minutes, seven minutes in total). On average, given the length of the auditory sequences summed to the interstimulus pause (ca 10 s total), 42 auditory sequences were reproduced during each test session (21 per side/numerosity).

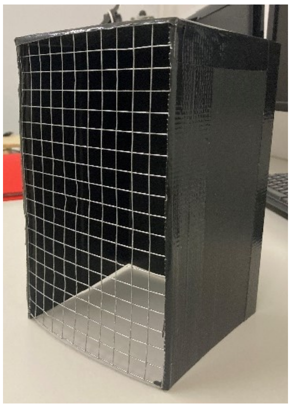


Figure 3.2: The small cage used to confine chicks
The cage measured 15 cm x 15 cm x 23 cm. It was placed inside the apparatus to allow for a 1-minute confinement of the chicks so that it could observe both stimuli before making a choice.

3.2.4. Data acquisition

We measured the time spent by the chicks in each of the two choice areas, which was used to compute an index of preference between the two acoustical stimuli (see paragraph 3.2.5). Indeed, the presence of the chick in a *choice area* was used as a proxy indicating a choice for the stimulus presented on that side of the apparatus. The central starting area was considered a *no-choice* area, and a choice was scored when the animal crossed with both feet the line delimitating the choice areas. The coding of chicks' behavior happened online. The experimenter recorded the movements of the chick in real-time, using a personalized Matlab function (MATLAB® v2020b, Inc., 2022). The Matlab function returned the total time, in seconds, spent by the animals in the different zones. Coding was performed by multiple coders which were all trained by myself. Moreover, I observed their coding for randomly selected animals to ensure consistency of data coding across coders and time.

3.2.5. Statistical analysis

All the statistical analyses were performed using R software (R Core Team, 2022) in the RStudio environment (RStudio Team, 2020). The analyses were based on the time chicks spent in each choice area. As the central area was considered a no-choice area, animals that spent the whole test duration (360 seconds) in the center of the apparatus were discarded. The raw times spent in each choice area have been also converted into indexes of preference, using the following formula: $Preference\ for\ stimulus\ A\ = \frac{Time_{spent\ close\ to\ A}}{Time_{spent\ close\ to\ A} + Time_{spent\ close\ to\ B}} * 100$

Scores higher than 50% indicated a preference for stimulus A, scores lower than 50 % indicated a preference for stimulus B, and a score of 50% indicated no preference for either stimulus. Depending on the comparison of interest, stimulus A could be, for instance, the sequence composed of the larger number of sounds (or the familiar sequence for chicks participating in an imprinting protocol, see Experiment 2). We decided to analyze both the absolute time spent close to the stimuli and the preference score, as these two measures reflect slightly different behaviors. In particular, the preference score is highly inflated by even the smallest time spent by the chick close to the stimulus, for animals that never approach the other stimulus. For example, one chick could spend only 5 seconds close to stimulus A (and the remaining 355 seconds in the no-choice area) resulting in a 100% preference score for stimulus A. Another chick instead could spend 80 seconds close to stimuli A and 20 seconds close to stimulus B, resulting in an 80% preference score for stimulus A. In this example, both chicks showed a preference for stimulus A, but their undergoing behavior was different. Even though these preference scores and absolute time are normally correlated, and tend to lead to the same conclusion, analyzing both of them provides a more comprehensive overview of the data distributions. Indeed, one more measure is more sensitive to the preference, of those chicks that make a strong choice (e.g., choosing only one stimulus), regardless of whether this choice behavior lasted for a few seconds only. In contrast, the absolute time reflects more the behavior of those chicks that have less dichotomic choice, but move more and provide longer choice times. Therefore they pick up slightly different effects (although internal coherence is expected). Moreover, the ratio value offers an easier comparison with other works conducted on domestic chicks, as absolute time is generally less reported in the literature (e.g. Rugani et al., 2015; Rugani, Regolin, et al., 2010).

More details are provided in the statistical analyses section of each experiment. In general, for all statistical analyses alpha = .05 significance level was chosen. The Z statistic and the ηp^2 of any significant result are reported as effect size, for the Wilcoxon test and ANOVA test respectively. We opted for ANOVA analysis as the ideal analysis needed to account for mixed model designs and multiple factors, even in the presence of non-normal distributed data as it has been demonstrated that this kind of analysis is robust to the violation of the normality assumption (Blanca et al., 2017). Correction for violation of sphericity has been reported when needed. Post-hoc analysis has been carried out with non-parametric tests to better take into account the non-normal distribution of the data since they provide the same information as the t-test student.

3.3. Experiment 1
Spontaneous preference for auditory discrimination

In the following two experiments, we tested the spontaneous preference of domestic chicks for auditory stimuli of different numerosity. We started with the simplest possible test, as we presented two different stimuli to the chicks and we measured whether they were significantly more inclined to approach one or the other. We tested two numerosities, 4 and 12, since a 1:3 ratio falls within the numerical discrimination acuity of chicks (Rugani et al., 2014).

One of the main challenges in studying numerical abilities using non-symbolic stimuli is disentangling the potential confound between numerosity and correlating physical variables. Indeed, continuous variables, such as size, density, and surface area (in the visual domain), often co-vary with numerical information in tasks requiring quantity discrimination. As a consequence, it might become difficult to disentangle which information is taken into account to solve the tasks. This intricate interplay between continuous and numerical variables is evident not only in non-human animals but also in human infants (Clearfield & Mix, 1999; Feigenson, Carey, & Spelke, 2002). For example, it has been demonstrated that both infants and non-human animals sometimes rely more on extensive variables rather than on numerical information. Especially when stimuli are homogeneous, continuous variables like volume may dominate the discrimination process. However, when elements are heterogeneous, subjects tend to focus on individual characteristics, leading to strategies based on numerosity (Feigenson, 2005; Rugani, Regolin, et al., 2010). Some of the studies already mentioned in the introduction, successfully demonstrated that chicks can perform numerical judgments when other quantitative information (e.g. item size, total area, volume) is controlled. For this reason, we have run two spontaneous choice experiments, one with no control over extensive variables (Experiment 1a) and one with additional manipulation to equalize extensive magnitudes between stimuli (Experiment 1b)

3.3.1. Experiment 1a – No control for extensive variables

In this experiment, the sound sequences were composed of 4 vs 12 identical sounds. Therefore, there was no control for the extensive continuous variables, as the total length of the auditory stimulation train (including both sounds and pauses) and the total amount of sound covaried with the numerosity (i.e. larger numerosity corresponded to a longer stimulation train and more sound). As a direct consequence of the sound structure being the same, the frequency (bmp, beats per minute) of stimuli was equal in both conditions (equal intensive variable).

3.3.1.1. Subjects

As preliminary pilot tests suggested a possible role of the testing day in determining the emergence of spontaneous preference, we decided to test two separate groups of chicks at P0 and P1. We tested overall 125 animals, 64 animals (30 females) at P0, and 61 animals (32 females) at P1.

3.3.1.2. Stimuli

Figure 3.3 shows a spectrogram of the acoustic stimuli used. A single 1000 Hz tone lasting 200 ms was used as the basic unit of the sequence. We concatenated 4 (or 12) tones with a 200 ms pause to generate a sequence of stimuli. Thus the resulting stimulation had a 120 bmp frequency range, similar to natural calls (De Tommaso et al., 2019; Kent, 1993). As we kept constant the inter-stimulus frequency and length of the single tones the 4-sound sequences lasted 1400 ms and the 12-sound sequence lasted 4600 ms, with a total of 800 ms and 2400 ms of played sound respectively (Figure 3.3). During the test, the two different auditory sequences were played asynchronously between the left and right speakers with a 5 s pause between each other. The starting sequence and side of the apparatus were fully counterbalanced across subjects.

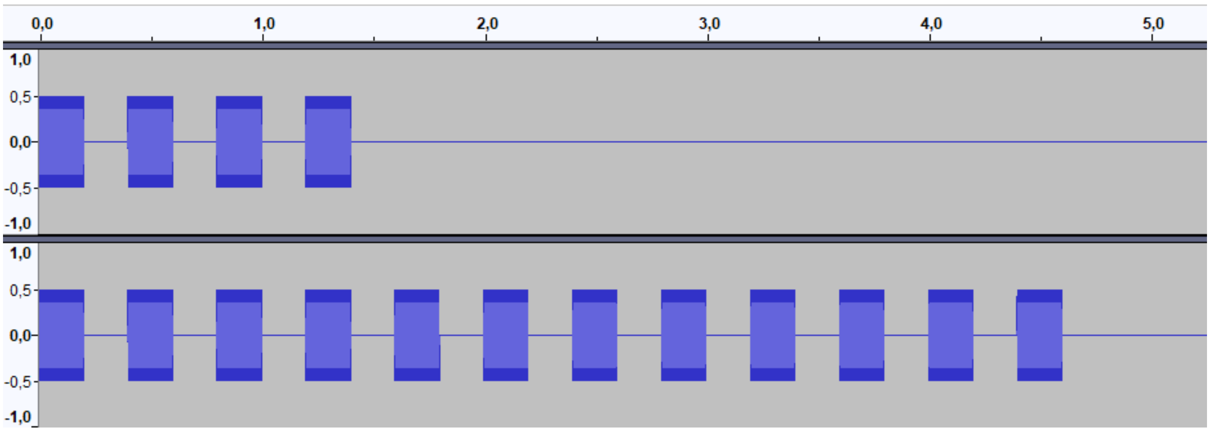


Figure 3.3: Spectrogram of Experiment 1a stimuli
Screenshot from the Audacity program window. The top spectrogram is the 4-sounds sequence, the bottom spectrogram is the 12-sounds sequence.

3.3.1.3. *Statistical analysis*

We tested separately chicks on P1 and P0. For each group, we measured the mean preference for the larger numerosity, using the index described in 3.2.4. As the majority of the chicks showed a dichotomic choice (i.e. either 100% or 0% preference for one stimulus) the data were not normally distributed, and therefore a non-parametric *two-tailed one-sample Wilcoxon signed rank test* was used to test preference against chance level *($\mu = 50$)*. A two-way analysis of variance (ANOVA) was performed to test Sex (Male vs Female) and Test Day (P0 vs P1) effects, given its robustness against violation of normality (Blanca et al., 2017), and homoscedasticity was verified with Mauchly's test: $P > 0.05$.

We also analyzed the raw time spent close to one stimulus or the other. A three-way mixed ANOVA was performed with Sex (Male vs Female), Test Day (P0 vs P1) as between factor, and Stimulus (Larger vs Smaller numerosity) as within Factor.

3.3.1.4. *Results*

Preference score

The 2x2 ANOVA revealed no significant effect of Sex ($F_{(1,120)} = 0.202$, $p_{value} = .654$), Stimulus ($F_{(1,120)} = 1.256$, $p_{value} = .256$) or interaction ($F_{(1,120)} = 0.081$, $p_{value} = .777$). Therefore we collapsed the data over both Sex and Test Day, and we found a significant preference for the larger numerosity (Mean = 59.2%, SEM = 4.32%, *two-tailed one-sample Wilcoxon signed rank test:* $W_{(124)} = 4634$, $p_{value} = .0336$, $Z = 0.191$). The results are shown in Figure 3.4.

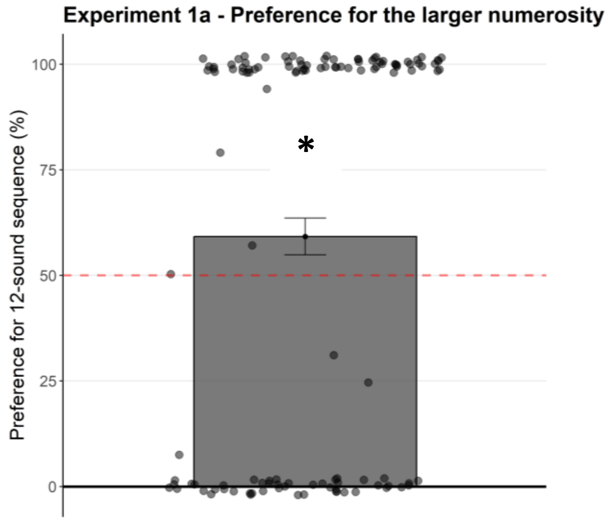


Figure 3.4: Results of Experiment 1a (Familiar preference score)
Overall preference for the 12-sound sequence in the whole sample. The red dotted line indicates the chance level.
The mean and the standard error of the mean are reported (two-tailed one sample Wilcoxon Sum rank test, * = $p_{value} < .05$).

Absolute time spent closer to stimuli

The 2x2x2 mixed ANOVA revealed no main effect of Sex ($F_{(1,120)} = 0.269$, $p_{value} = .605$), and of Test Day ($F_{(1,120)} = 1.365$, $p_{value} = .245$), but a non-significant trend was observed for Stimulus ($F_{(1,120)} = 3.467$, $p_{value} = .065$). As shown in Figure 3.5, it seems that animals spent more time closer to larger numerosity (Mean = 137 s, SEM = 11.9) than to the smaller numerosity (Mean = 95.7, SEM = 11.3). No significant interaction was found (Sex*Test Day: ($F_{(1,120)} = 0.087$, $p_{value} = .769$; Sex*Stimulus: $F_{(1,120)} =0.179$, $p_{value} = .673$, Test Day*Stimulus: $F_{(1,120)} = 0.713$, $p_{value} = .400$; Sex*Test Day*Stimulus: $F_{(1,120)} = 0.094$, $p_{value} = .760$).

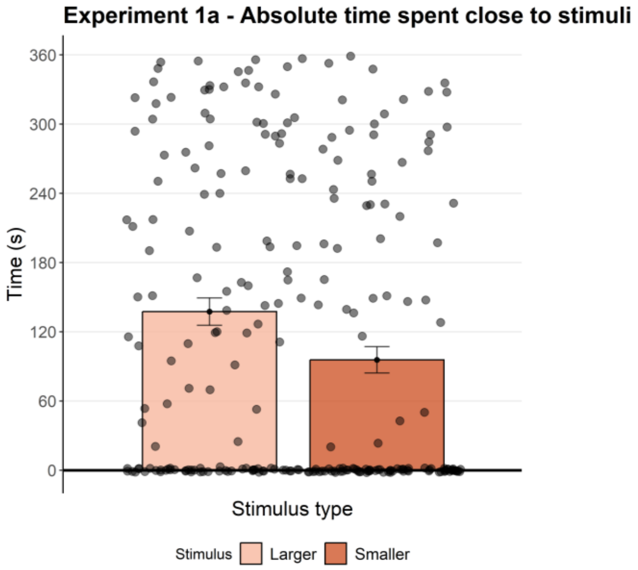


Figure 3.5: Results of Experiment 1a (Familiar vs Novel absolute time)
Absolute time spent close to each stimulus by all the chicks.
The means and the standard errors of the mean are reported.

3.3.2. Experiment 1b – Control for extensive variables

As we mentioned above, continuous variables represent one of the main confounds when investigating numerical abilities in response to non-symbolic stimuli. In Experiment 1a, we let the total length of the auditory stimulation train and the total amount of sound positively correlate with the number of discrete sounds composing the sequence. In contrast, in Experiment 1b, we lengthened the individual tones in the 4-sound sequence, matching the total length of the sequence and the total amount of sound between the two stimuli. This manipulation allows us to disentangle the difference in numerosity from the difference in these two extensive dimensions, testing whether they had a role in the preference for the 12-sound sequence observed in Experiment 1a. However, note that to match these two extensive dimensions, we had to employ two sequences with different frequency of sound-to-silence alternation frequencies.

3.3.2.1. Subjects

As for Experiment 1a, we tested two separate groups of animals on the first and the second day post-hatching. We tested overall 126 animals, 59 animals at P0 (26 females), and 67 animals at P1 (37 females).

3.3.2.2. Stimuli

Figure 3.6 shows a spectrogram of the acoustic stimuli used. The 12-sound sequence was the same as Experiment 1a, while tones in the 4-sound sequences were lengthened to 600 ms with a ca 730 ms pause. Thus, crucially the two sequences lasted both 4600 ms, with a total of 2400 ms of played sounds. As in Experiment 1a, during the test, the two different auditory sequences were played asynchronously between the left and right speakers with a 5 s pause between each other. The starting sequence and side of the apparatus were fully counterbalanced across subjects.

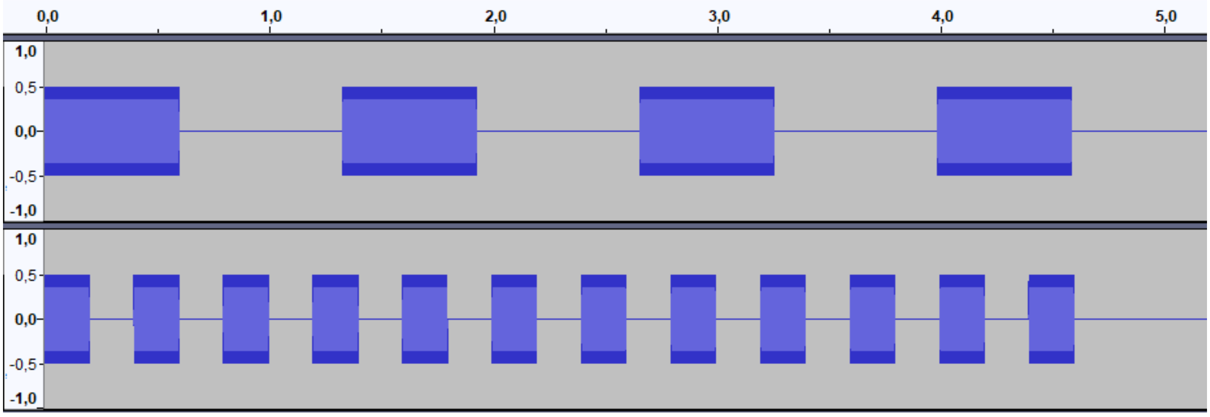


Figure 3.6: Spectrogram of Experiment 1b stimuli
Screenshot from the Audacity program window. Top spectrogram is the 4-sounds sequence, bottom spectrogram is the 12-sounds sequence.

3.3.2.3. Statistical analysis

The statistical analyses were conducted similar way to the previous experiment, see paragraph 3.3.1.3.

3.3.2.4. Results

Preference score

The 2x2 ANOVA revealed no significant main effect for Sex ($F_{(1,122)} = 0.338$, $p_{value} = .562$) and of Test Day ($F_{(1,122)} = 1.862$, $p_{value} = .175$). Instead, we found a significant interaction Sex*Test Day ($F_{(1,122)} = 4.345$, $p_{value} = .039$). Pairwise comparisons analysis revealed a difference in preference for larger numerosity in males tested at P0 (Mean = 36.4%, SEM = 8.50%) from males tested at P1 (Mean = 66.7%, SEM = 8.75%; t*wo-tailed two-sample Wilcoxon sum rank test:* $W_{(30,33)} = 345$, $p_{value} = .018$, $Z = 0.30$), while there was no difference in preference in females (P0: Mean = 49.6%, Sem = 9.44%; P1: Mean = 43.2%, SEM = 8.26%; t*wo-tailed two-sample Wilcoxon sum rank test:* $W_{(26,37)} = 518$, $p_{value} = .561$). When tested against chance level, males in P0 showed no significant difference from chance (t*wo-tailed one-sample Wilcoxon signed rank test:* $W_{(30,33)} = 204$, $p_{value} = .12$), and males tested at P1 showed a non-significant trend above chance (t*wo-tailed one-sample Wilcoxon signed rank test:* $W_{(30,33)} = 310$, $p_{value} = .067$). The results are shown in Figure 3.7.

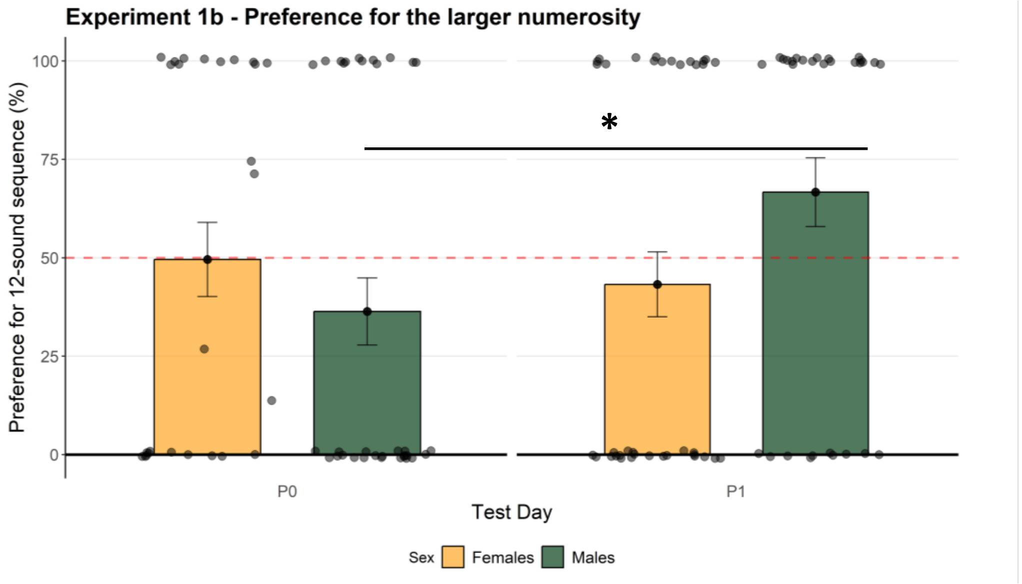


Figure 3.7: Results of Experiment 1b (Familiar preference score)
Sex preference for the 12-sound sequence in the two test days. The red dotted line indicates the chance level.
The means and the standard errors of the mean are reported (two-tailed two-sample Wilcoxon Sum rank test,* = $p_{value} < .05$).

Absolute time spent closer to stimuli

The 2x2x2 mixed ANOVA revealed only a main effect of Sex ($F_{(1,122)} = 7.837$, $p_{value} = .006$, Figure 3.8). No significant effect of Test Day ($F_{(1,122)} = 0.568$, $p_{value} = .453$), Stimulus ($F_{(1,122)} = 0.040$, p_{value}

= .843), or interaction was found significant (Sex*Test Day: ($F_{(1,122)}$ = 2.785, p_{value} = .098; Sex*Stimulus: $F_{(1,122)}$ =0.184, p_{value} = .669, Test Day*Stimulus: $F_{(1,122)}$ = 0.500, p_{value} = .481). Only a non-significant trend for Sex*Test Day*Stimulus interaction emerged ($F_{(1,122)}$ = 3.165, p_{value} = .078).

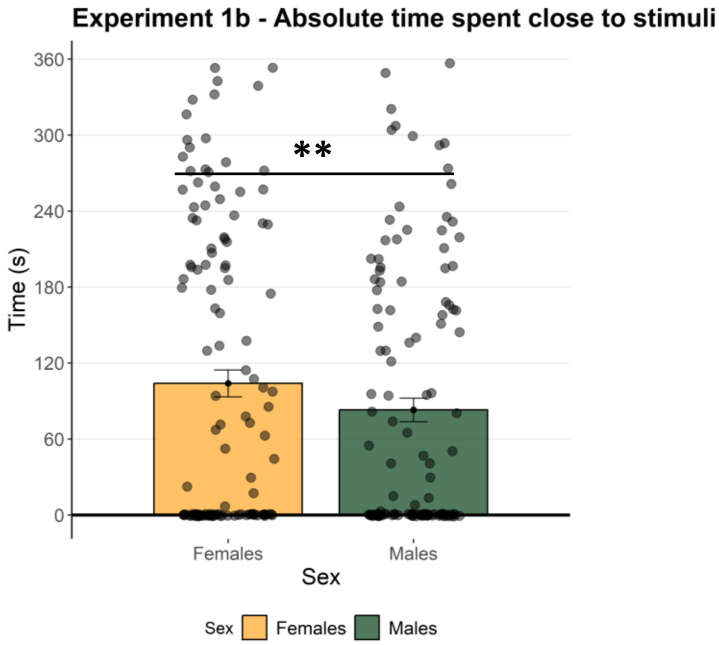


Figure 3.8: Results of Experiment 1b (Absolute time)
Absolute time spent close to each stimulus by Sex.
The means and standard errors of the mean are reported (Anova analysis,** = $p_{value} < .001$).

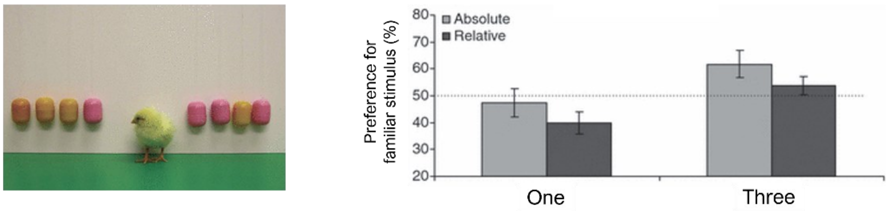
3.3.3. Discussion

In these two experiments, we investigated chicks' spontaneous preference for sound sequences of different numerosities, controlling either for extensive or intensive variables.

Experiment 1a can be considered the most ecological setting as sounds present themselves in natural conditions, with no external manipulation on other extensive dimensions. In this situation, chicks showed a preference for the larger numerosity, i.e. the 12-sound sequence. The preference for the larger set of stimuli has been already demonstrated in domestic chicks to the point that it has been exploited to test other numerical competencies. For example, chicks' arithmetic abilities have been demonstrated by exploiting their tendency to consistently inspect the screen hiding the larger set of imprinting objects (Rugani, Regolin, et al., 2011). A preference to approach the larger set of objects emerged also in other imprinting studies (Lemaire et al., 2020; Rugani, Regolin, et al., 2010). As already discussed by the authors of those studies, the tendency to approach a larger group of social companions can be motivated by different reasons, such as a preference for contexts offering a higher level of protection from predators and social interactions, or even more heat in a natural environment (Pulliam, 1973; Roberts, 1996). In our study, however, chicks did not undergo an imprinting phase, so they were not familiar with the sounds they heard. This is interesting because our results could suggest that in an auditory paradigm, chicks could spontaneously interpret sounds as an indicator of social companions, and therefore they used the perceived most numerous sound as a proxy of a larger group. In Experiment 1a, however, no control for extensive variables was made, meaning that in this case, chicks may have responded based not only on numerosity but also on other extensive dimensions.

As extensive variables may play a role in the emergence of the preference as demonstrated by other studies, in Experiment 1b we controlled for this factor. We found that in the absence of additional magnitude indications, such as total duration and amount of sound, the overall preference for the larger auditory sequence disappeared. These results suggest that, when put in conflict, extensive variables, such as the total length of the sequence and the total amount of sound, and not intensive variables, such as frequency rate, drive this preference. Indeed, in Experiment 1b the two stimuli had the same length, but different frequency rates. Nevertheless, the chicks did not show any preference for the higher frequency rate. Our results also seem to align with the current literature, as studies investigating the role of extensive variables in numerical tasks, demonstrated that the stimuli structure matters for determining which cues are used to respond. When a set of stimuli is composed of homogeneous elements, both infants and non-human animals show the tendency to use extensive variables to solve the task (Feigenson, Carey, & Hauser, 2002; Rugani, Regolin, et al., 2010). On the other hand, when sets of stimuli are composed of a heterogeneous group of elements, it is most plausible that they use numerical information (Feigenson, 2005; Rugani, Regolin, et al., 2010, see Figure 3.9).

A **Experiment 1 -** *Homogenous stimuli*



B **Experiment 4 -** *Heterogeneous stimuli*

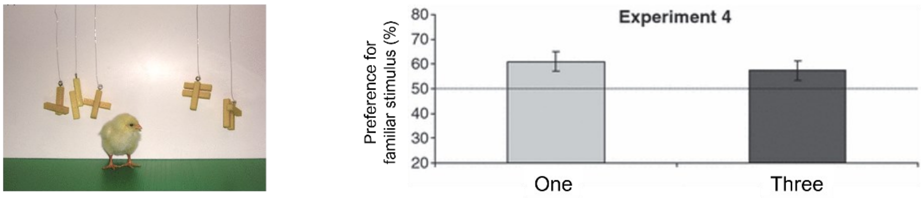


Figure 3.9: Stimuli and results from Rugani et al., 2010
Panel A) In experiment 1 the shape and the dimension of the stimuli were homogeneous, and chicks either showed a preference for the larger set of stimuli or no preference.
Panel B) In experiment 4 the shape and the dimension of the stimuli were heterogeneous, and chicks showed a preference for the familiar stimulus in both the imprinting conditions.
Modified from Rugani et al., 2010 (see the original article for other control experiments)

In our experiments, we used homogeneous stimuli, as the same sound was repeated several times. According to the results obtained in the visual domain, it is, therefore, possible that in Experiment 1a chick used the additional magnitude information to guide their behavior. While, in Experiment 1b, where that additional information was not present, they did not show any preference. Additionally, studies conducted with infants, demonstrated that the role of continuous variables is more prominent for discrimination of large numerosities (Xu & Spelke, 2000). In chicks for example, it has been demonstrated that when presented with a bigger ratio (e.g. 6 vs 9), they succeed in discrimination only when other quantitative variables were available, while this does not happen for smaller ratios (e.g. 10 vs 20), where they succeed regardless (Rugani et al., 2013). Also, when tested for arithmetical abilities, chicks were able to summate large numerosities (e.g. up to 9 in 6 vs 9 conditions) only when quantity information was not kept equal (Rugani, Regolin, et al., 2011). These results, in line with what we observed in Experiment 1b, seem to suggest that in the presence of larger sets, redundancy of numerical and extensive non-numerical information is crucial for the animals to succeed (see also Rugani, 2018 for a review). Our results suggest also that intensive variables, for example, the frequency rate, are not sufficient to guide chicks' response. Indeed, in Experiment 1b, the 12-sound sequence had a higher frequency rate but chicks did not show a preference for it. It seems that the correlation between numerosities and extensive, not intensive, variables is more crucial for numerical discrimination in chicks.

We also found differences in behavior between males and females, with males showing an increased preference for the larger numerosities from P0 to P1. Behavioral differences between sexes have been observed in chicks for various tasks in prior research (Vallortigara 1992; Miura and Matsushima 2012; Santolin et al. 2020; Rosa-Salva et al. 2023). It has been debated that motivational factors might lead to different early social behaviors. For example, males are more prone to perform tasks for food rewards (Vallortigara, 1992) and to be more aggressive (Vallortigara 1990), while females seem to have higher motivation for social reinstatement. It is possible that on the first day after hatching, in our task, males responded to unknown sounds showing explorative/aggressive behaviors. Further studies, investigating responses to more naturalistic calls should be run to better understand sex-related differences in approaching sequences of sounds of different numerosity.

In conclusion, these two experiments reveal that chicks show a preference for sequences composed of more sounds when numerical and extensive non-numerical information are both present. However, for the current results, we cannot know if chicks showed a preference for the longer sound or the most numerous one. Future research should better disentangle the role of the two features in determining the found preference. For example, in a future study, we could use one single long sound versus one single short sound, to test whether, in response to a difference in extensive magnitudes only and not numerical information, chicks would still prefer the longer sound. While chicks showed this ability without having been exposed to the test sounds before, they still had general acoustic experience previous to the test session. Indeed, while incubating chicks in darkness allows for complete visual deprivation before the test, the same does not happen for auditory stimulation, As our chicks were incubated in batches of 20-30 eggs (given the obvious practical limitation of incubating chicks individually), they could hear each other calls, which they emit to synchronize hatching since the last days of incubation. Even though chicks' calls do not particularly resemble the artificial test stimuli we used, this species-typical form of auditory stimulation might have a role in the maturation of chicks' responses to acoustical stimuli after hatching. For example, we cannot exclude the possibility that the chicks spontaneously produce call trains whose total duration or overall amount of sound is more similar to the stimulus of 12 than to that of 4. This initial evidence needs to be further supported by future studies to fully understand how chicks' numerical representation may work in the auditory domain. For example, future studies should investigate chicks' responses to auditory sequences of heterogenic sounds. If acoustic processing works similarly to the visual one, in this case, we may observe a preference for the larger set of stimuli even when extensive variables are controlled. The role of continuous variables should also be further investigated with regard to the processing of larger sets of sounds characterized by more narrow ratios. This could allow us to confirm whether the information provided by extensive continuous variables has similar selective effects also in the auditory domain.

3.4. Experiment 2
Imprinting for auditory stimulation

In this experiment, we investigated whether chicks could develop filial imprinting for and subsequently recognize a sequence composed of a specific number of sounds. To avoid potential interferences from spontaneous preferences, we decided to use the same auditory stimulation as Experiment 1b, since no spontaneous preference between the stimuli had been found in that experiment.

3.4.1. Methods

3.4.1.1. Subjects

We tested 58 animals at P1, 28 chicks (14 females) imprinted with a 4-sound sequence and 30 chicks (13 females) imprinted with a 12-sound sequence.

3.4.1.2. Stimuli and Procedure

The stimuli and the test procedure were the same as used in Experiment 1b (see paragraph 3.3.2.2 for details, Figure 4.2). The only difference from Experiment 1b, was that before the testing phase chicks were subjected to an imprinting phase, during which they were exposed to either the 4-sound sequence or the 12-sound sequence (identical stimuli were used for the test phase). Each batch of eggs was thus imprinted on one of the two numerosities, in a between-subjects design.

The imprinting phase started *in ovo*, from the 18th day of development up to and including the hatching day. Eggs were incubated in the dark, as described before, and two speakers (Z130 Stereo Speakers) inside the incubator played the auditory imprinting sequence. The sound was played starting from 8 am for 8 hours a day, alternating 2 hours of stimulation with 2 hours of silence. During the 2 hours, the sequence was reproduced every 6 minutes, mimicking the test duration. During the remaining hours, no sound was played, aiming to both reproduce an ecological situation (i.e. silence at night, since domestic chickens are diurnal animals) and to avoid transforming the sounds into a background noise that the animals may ignore. Animals were tested on P1, and on the morning of test day, the imprinting stimulus was played for 30 minutes, as a short recall session. The test session was the same as described in paragraph 3.2.3.

3.4.1.3. Statistical analysis

For each group, we measured the mean preference for the familiar numerosity (i.e. the numerosity chicks were imprinted with) with the index described in 3.2.4. As for previous experiments, a non-parametric *two*-tailed one-sample Wilcoxon signed rank test was used to test preference against chance level ($\mu = 50\%$). A two-way analysis of variance (ANOVA) was performed to test Sex (Male vs Female) and Imprinting condition (4-sound vs 12 sounds) effect, and sphericity was verified with Mauchly's test: $P > 0.05$.

We also analyzed the raw time spent close to one stimulus or the other. A three-way mixed ANOVA was performed with Sex and Imprinting condition as between factors and Stimulus (familiar vs non-familiar numerosity) as within factor.

3.4.2. Results

Preference score

The two-way 2x2 ANOVA revealed only a main effect of Imprinting condition on preference for familiar stimulus ($F_{(1,54)} = 4.343$, $p_{value} = .042$, $\eta_p^2 = 0.074$), as animals imprinted to 4-sound showed a lower, but not significant, preference for familiar stimulus (Mean = 33.8 %, SEM = 8.88 %; *Two-tailed one-sample Wilcoxon signed rank test:* $W_{(28)} = 144$, $p_{value} = 0.144$) compared to animals imprinted to 12-sound, also not significant (Mean = 58.4 %, SEM = 8.51 %; *Two-tailed one-sample Wilcoxon signed rank test:* $W_{(20)} = 276$, $p_{value} = 0.341$). No significant effect for Sex ($F_{(1,54)} = 0.358$, $p_{value} = .552$) or interaction ($F_{(1,54)} = 0.860$, $p_{value} = 0.358$) was found. The results are shown in Figure 3.10.

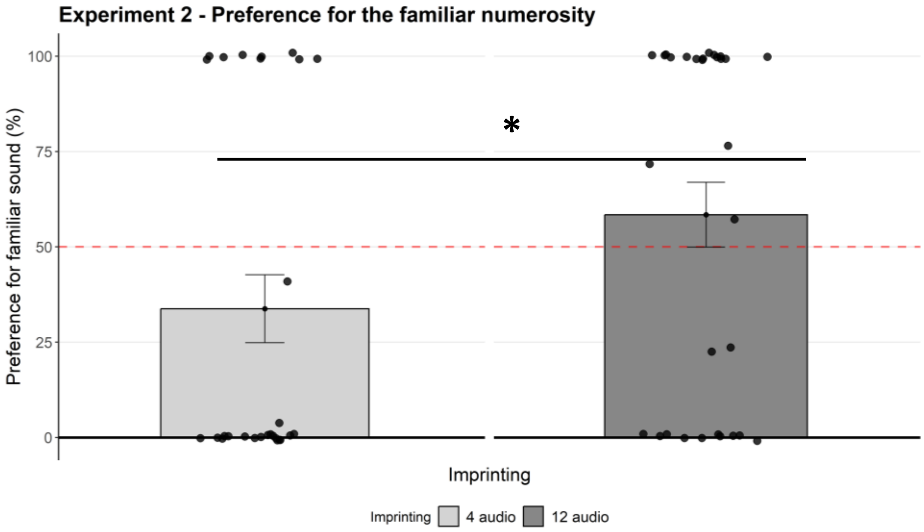


Figure 3.10: Results of Experiment 2 (Familiar preference score)
Preference for the familiar sound sequence by Imprinting conditions. The red dotted line indicates the chance level. The mean and standard errors of the mean are reported (Anova analysis, * = $p_{value} < .05$).

Given the apparent preference for larger numerosity in both groups, we performed also a two-way 2x2 ANOVA on the preference index for a larger stimulus. The ANOVA revealed no main effects (Imprinting: $F_{(1,54)} = 0.275$, $p_{value} = .602$; Sex ($F_{(1,54)} = 0.860$, $p_{value} = .358$), and no significant interaction ($F_{(1,54)} = 0.358$, $p_{value} = .552$). Therefore we collapsed the data over both Sex and Imprinting, and we found a non-significant trend in preference for larger numerosity (Mean = 62.2 %, SEM = 6.11 %; *two-tailed one-sample Wilcoxon signed rank test:* $W_{(57)} = 1064$, $p_{value} = .08$, Figure 3.11).

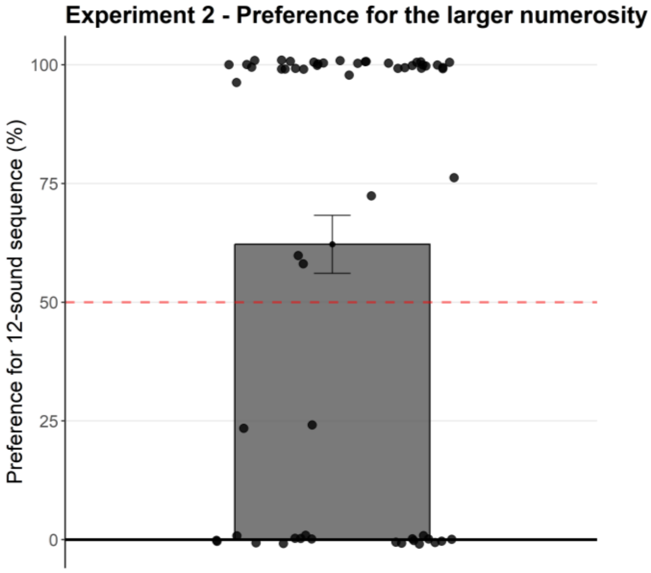


Figure 3.11: Results of Experiment 2 (Larger preference score)
Overall preference for the 12-sound sequence in the whole sample. The red dotted line indicates the chance level. The mean and the standard error of the mean are reported.

Absolute time spent closer to stimuli

We measured how much time chicks spent closer to familiar and novel stimuli. The three-way 2x2x2 ANOVA revealed no significant main effects (Sex: ($F_{(1,54)} = 1.280$, $p_{value} = .236$, Imprinting Condition: $F_{(1,54)} = 0.156$, $p_{value} = .695$, Stimulus: $F_{(1,54)} = 0.546$, $p_{value} = .463$). We found a significant interaction between the Imprinting condition and Stimulus ($F_{(1,54)} = 9.349$, $p_{value} = .003$, $\eta_p^2 = 0.148$). Animals imprinted to 4 sounds spent more time closer to the novel stimulus (Mean = 168 s, SEM =25.6 s) than to the familiar stimulus (Mean = 58 s, SEM = 18.2 s; *two-tailed two-sample Wilcoxon signed rank test:* $W_{(27)} = 90$, $p_{value} = .009$), while for animals imprinted to 12 sounds the time spent closer to the novel stimulus (Mean = 76 s, SEM = 19.5 s) was not significantly different from the time spent closer to the familiar one (Mean = 139 s, SEM = 24.5 s; *two-tailed two-sample Wilcoxon signed rank test:* $W_{(29)} = 306$, $p_{value} = .135$). No other significant interaction was found (Sex*Stimulus: $F_{(1,54)} = 0.009$, $p_{value} = .923$, Sex*Imprinting*Stimulus: $F_{(1,54)} = 1.007$, $p_{value} = .320$). The results are shown in Figure 3.12.

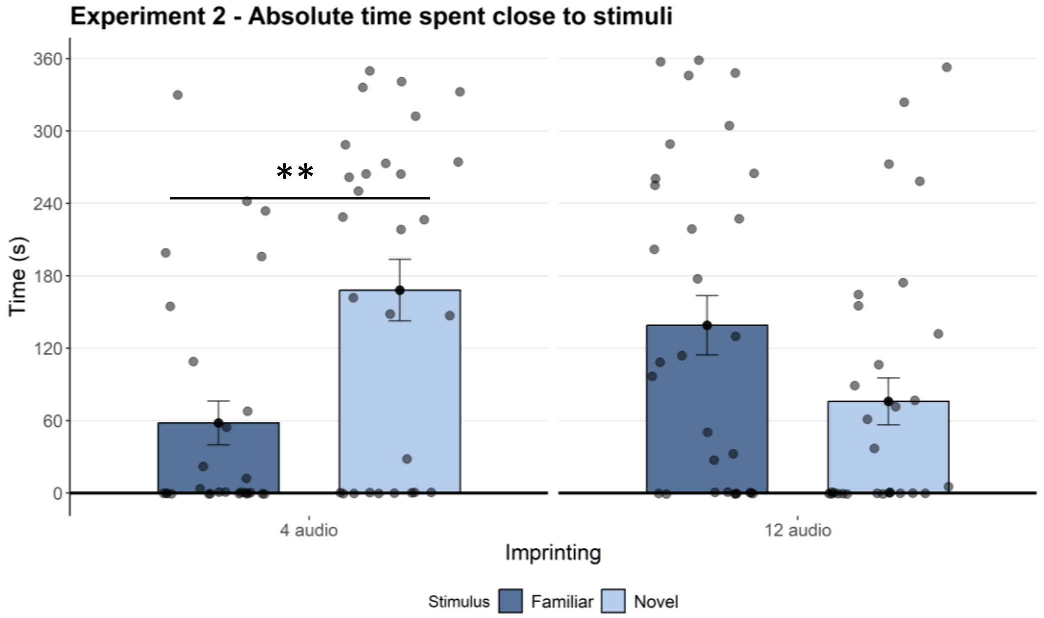


Figure 3.12: Results of Experiment 2 (Familiar vs Novel Absolute time)
Absolute time spent close to familiar and novel sound by Imprinting conditions.
The means and the standard errors of the mean are reported (two-tailed two-sample Wilcoxon signed rank test,** = p_{value} <.01).

As previously done for the other independent variable, we analyzed the time spent close to one stimulus or the other in relation to their numeric value. The three-way 2x2x2 ANOVA revealed only a main significant effect of Stimulus ($F_{(1,54)} = 9.349$, $p_{value} = .003$, $\eta_p^2 = 0.148$), as animal spent more time close to larger numerosity (Mean = 153, SEM = 17.7) compared to smaller numerosity (Mean = 67.3, SEM = 13.3). No significant effect for Sex ($F_{(1,54)} = 1.280$, $p_{value} = .236$), Imprinting Condition ($F_{(1,54)}$ = 0.156, p_{value} = .695), or interactions (Sex*Imprinting: $F_{(1,54)}$ = 0.333, p_{value} = .566; Sex*Stimulus: $F_{(1,54)}$ = 1.007, p_{value} = .320, Imprinting*Stimulus: $F_{(1,54)}$ = 0.546, p_{value} = .463; Sex*Imprinting*Stimulus: $F_{(1,54)}$ = 0.009, p_{value} = .923) was found. The results are shown in Figure 3.13.

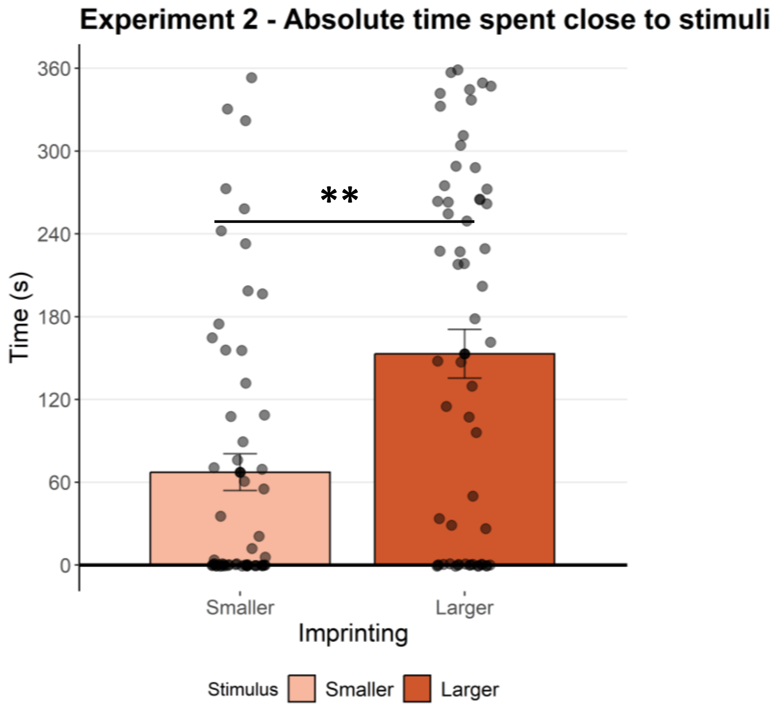


Figure 3.13: Results of Experiment 2 (Larger vs Smaller Absolute time)
Absolute time spent close to larger and smaller numerosity by all chicks.
The means and the standard errors of the mean are reported (Anova analysis,** = p_{value} <.01).

3.4.3. Discussion

The results of this experiment are in line with previous studies that investigated numerical imprinting using visual stimuli (Lemaire et al., 2020; Rugani, Regolin, et al., 2010). While chicks did not display a significant preference for the familiar numerosity, a general preference for the larger numerosity stimulus emerged. This was apparent, for instance, by the fact that, regardless of their imprinting condition, chicks spent significantly more time (absolute number of seconds) closer to the 12-sound sequence than to the 4-sound sequence. Moreover, when we analyzed the index representing the proportion of time spent near the familiar imprinting stimulus, this was higher for chicks imprinted on the 12-sound sequence (for which the familiar stimulus was also the larger numerosity) than on chicks imprinted on the 4-sound sequence.

Notably, here we used the same stimuli as in Experiment 1b, in which we did not observe any spontaneous preference for the 12-sounds sequence. The main difference between these two experiments is that here, unlike in Experiment 1b, the chicks were pre-exposed to one of the two sound sequences during the imprinting phase. Thus, it seems that this exposure may facilitate the emergence of a preference for the larger stimulus when stimuli are controlled for extensive variables. One factor that might influence the success of numerical imprinting in the acoustic modality is related to the temporal dimension along which acoustical sequences unfold. In a series of sequentially presented sounds, the time interval between individual sequence elements becomes particularly relevant to allow a comprehensive, global representation of the sequence as a whole object which is however composed of distinct individual elements (e.g., a single sequence composed of 12 sounds). For example, the strength of chicks' imprinting on colored rotating objects and flashing lights was modulated by the rate of presentation (Chantrey, 1974). In 1976, he also showed that when chicks were imprinted to two objects of alternating colors presented rapidly, the animals generalized the same response to both (Chantrey, 1976). This is considered evidence that stimuli alternating too quickly can be 'melted' together in a single representation, leading to a unified classification of the composing elements. It is therefore possible that either the structure of the sequences themselves (i.e., sounds too close together) or the structure of the entire imprinting phase (alternating every 5 seconds) was not optimal for conveying numerical information. Further research is necessary to better understand which factors can effectively facilitate numerical imprinting. For example, systematic manipulation of sounds duration, interstimulus presentation (within sequences and between sequences), and timing of imprinting procedure should be carried out.

Additionally, we also need to consider that we presented the same sounds between the imprinting and test phases. This means that, despite the difference in numerosity, both stimuli are similar to the imprinting objects. We cannot exclude that the chicks imprinted on the specific single sound, thus approaching at test the perceived larger group of social companions, similar to the results of Experiment 1a. In future studies, non-familiar sounds, e.g. different in pitch, could be used during the test phase to investigate whether, in the presence of novel stimuli, chicks are motivated to approach the stimulus that presents at least one familiar feature, namely numerosity.

3.5. General Discussion

This study is the first to investigate whether the numerical discrimination abilities observed in domestic chicks in response to visual stimuli are also present in response to auditory stimuli. Overall, our data suggest that chicks can distinguish between auditory sequences of different numerosities, at least when numerosity correlates with some extensive magnitude dimensions, as revealed by a preference for sequences composed of more sounds. However, when extensive variables are controlled, this preference seems to emerge only if chicks have been pre-exposed to the auditory sequences, as it happens during the imprinting phase. In contrast, the preference disappears in spontaneous choice paradigms, where chicks appear to base their decisions on extensive non-numerical variables, such as the total length of the sequence or the total amount of sound. It is important to note that spontaneous choice paradigms measure the animals' inherent predisposition towards a particular choice, rather than their intrinsic ability to make that choice (i.e., the ability to discriminate the two stimuli presented to them). It is possible that, at the group level, no significant preference emerges simply because the animals have no "reason" to choose one stimulus over the other. This, however, does not imply that the animals are incapable of discriminating. As a general rule, if in those tests animals show a significant preference, it indicates that they have discriminated between the two stimuli, but the opposite assumption cannot be made if no preference is shown. Indeed, the results of Experiment 2, clearly indicate chicks can distinguish between the two auditory sequences even when extensive variables are controlled. However, as for the current results, we cannot know with certainty what determined the preference for the larger numerosities. On one hand, it is possible that chicks imprinted on the single sound unit, and showed at the test a preference for the larger group of social companions, i.e. actually relying on numerical information. On the other hand, the mere exposure to sounds during the imprinting phase may have an activating effect that could resolve in a preference for the higher frequency rate, which is closer to the hen's call (Kent, 1993), or in a preference for the larger numerosity, irrespectively of the specific imprinting sounds. For example, it has been demonstrated that other kinds of early stimulation, such as motor stimulation, increased preference for social stimuli (Johnson et al., 1985). Various types of early stimulation, such as artificial sounds, maternal calls, motor activity in darkness, handling, or exposure to a prominent abstract visual stimulus, facilitate the emergence of predisposition in chicks (reviewed in Rosa Salva et al., 2015). Additional control experiments are needed to disentangle these two alternatives. For example, future studies could test the role of a different kind of early stimulation to sound preferences, or different sounds could be used between the imprinting and the test phase to test if it is a sound-specific preference.

As mentioned earlier, the relationship between continuous variables and numerosity remains a hotly debated and ongoing topic. In auditory processing, continuous variables like sound duration, frequency, and intensity can influence the perception of numerosity. For example, when presented with a sequence of sounds, both humans and animals may rely on the duration of the sequence, the cumulative amount of sound, or the intensity of the sounds rather than solely on the number of distinct auditory events. This suggests that in both visual and auditory domains, numerical processing might not be entirely independent of other sensory attributes. Indeed, an alternative theory to the number sense has also been developed: A Theory of Magnitude (ATOM) (Walsh, 2003). This theory posits that the brain uses a common neural system to process different types of magnitude information, including time, space, and number. According to ATOM, the overlap in processing spatial, temporal, and numerical features arises because these features are processed in overlapping neural circuits. Humans and other vertebrate species process different dimensions in a similar manner (e.g., space, time, length, number) at both the behavioral (De Corte et al., 2017; Lourenco & Longo, 2010; Miletto Petrazzini & Brennan, 2020) and neural levels (Walsh, 2003). Overall, the results of this study in young domestic chicks, seem to support this view.

Another interesting point concerns the evidence of how continuous variables and numerosity interact, depending on the absolute value of the numerosities presented (Feigenson, 2005; Rugani et al., 2013; Xu & Spelke, 2000). According to the hypothesis that divides numerical perception into two systems, OTS and ANS, continuous variables may be less relevant in the presence of small numerosities. With small numerosities, stimuli may be processed more holistically as individual objects, while larger numerosities, being within the domain of the ANS, are represented more approximately, where, according to theories like ATOM, the boundaries between numerosity and continuous variables are more blurred. Further studies conducted with auditory sequences representing smaller numerosities, such as 2 or 3 sounds, are therefore necessary to understand how the performance of domestic chicks fits within the current literature.

In conclusion, this study provides new insights into the numerical discrimination abilities of domestic chicks in response to auditory stimuli. Our findings indicate that while chicks demonstrate a capacity to distinguish between different numerosities, this ability is influenced by the context in which the sounds are presented and by prior exposure to the stimuli during imprinting. This suggests that, in the absence of explicit motivation, such as familiarity from imprinting, chicks may default to relying on extensive, non-numerical cues rather than numerical information alone. This reliance on continuous variables aligns with theories like ATOM, which propose that magnitude processing, encompassing time, space, and number—draws upon shared neural resources, especially for larger quantities. Further research with controlled exposure and a focus on smaller numerosities could help clarify the boundary conditions under which numerical and non-numerical cues interact and inform our understanding of the innate versus learned components of numerical processing across species.

Chapter 4

Study 3

Cross-modal numerical transfer in domestic chicks

4.1. General Introduction

The ability to integrate information from different senses is crucial for environmental perception in both humans and non-human animals. Cross-modal correspondences refer to ‹‹*associations that appear to exist between different basic physical stimulus attributes, or features, in different sensory modalities*›› (Spence, 2011). Ratcliffe and colleagues, in a review, reported several studies that investigate how cross-modal correspondences are shared across non-human mammals, as they can naturally combine signals from different sensory modalities, especially when these signals share similar features or frequently co-occur (Ratcliffe et al., 2016).

Domestic chicks have been demonstrated to show forms of cross-modal correspondence in a couple of studies. Lonconsole and colleagues found a cross-modal correspondence between luminance and space, as chicks spontaneously associate a low-luminance (black) panel with the right side of the apparatus and a high-luminance (white) panel with the left-side (Loconsole et al., 2021). In 2017, Versace and colleagues demonstrated that visually naïve chicks could be imprinted on the combination of visual-audio stimuli, to the point that they spontaneously generalized abstract multimodal patterns. The animals were imprinted on XX or XY patterns, in either a visual condition, an acoustic condition, or a cross-modal condition (between-subjects). At the test, they were then presented with novel pairs of stimuli XX vs XY, in the same modality they were imprinted with. The authors found that the multimodal presentation of the stimuli elicits a stronger discrimination between the two patterns, than what was found in unimodal presentation, indicating that the two sources of information had been integrated (Versace, Spierings, et al., 2017). As we already mentioned in paragraph 1.1.2, also numerical cognition benefits some form of cross-modal correspondences, for example, audio-visual match and SNA.

While the latter has been already investigated in chicks (Rugani et al., 2015; Rugani, Vallortigara, et al., 2011; Rugani & Regolin, 2020), to our knowledge, there is still no evidence of cross-modal matching of numerical information in domestic chicks.

In the first experiment of this Study, inspired by the study of Izard and colleagues that demonstrated a-modal numerical representation in human newborns (Izard et al., 2009), we aimed to investigate whether a similar phenomenon could occur in chicks. We did that by adapting their paradigms for studying spontaneous preference for matching versus mismatching cross-modal numerical information. In the second experiment instead, we tried a different approach and we investigated whether chicks could transfer numerical information acquired acoustically through imprinting to later solve a visual task.

4.2. Experiment 3

Spontaneous preference for contingency

In the following experiment, we investigated whether numerically naïve, untrained chicks could spontaneously discriminate conditions in which the acoustical and visual modalities provide congruent or incongruent numerical information. To do so, we decided to test whether we could observe a spontaneous preference for approaching a set of objects whose number matched the number of sounds composing an acoustical sequence audible in the environment.

4.2.1. Methods

4.2.1.1. Subjects

We tested overall 149 animals, 84 chicks (51 females) with pink stimuli and 65 chicks (35 females) with red stimuli.

4.2.1.2. Stimuli

Acoustic stimuli

The 12-sound sequence was the same as described in all the other experiments of the thesis (see paragraph 3.3.1.2), while the 4-sound sequence was composed of 4 tones lasting 1 s each, with 200 ms pauses in between. In this case, the total length of the two sequences is equal, as the interstimulus intervals between each sound within the sequence.

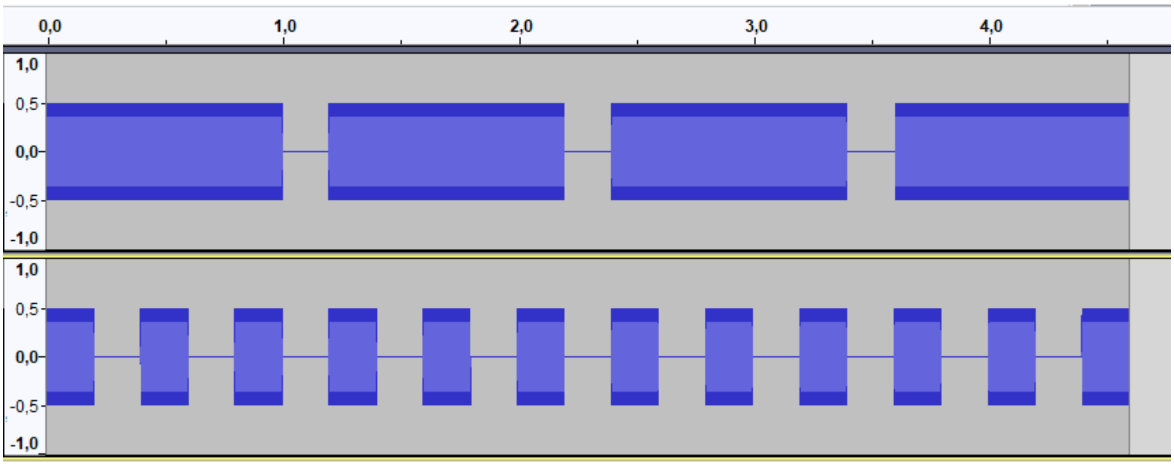


Figure 4.1: Spectrogram of stimuli used in Experiment 3
Screenshot from the Audacity program window. The top spectrogram is the 4-sounds sequence, the bottom spectrogram is the 12-sounds sequence

Visual stimuli

Visual stimulation consisted of two different sets of 3D objects, one composed of 4 elements and one of 12 elements. We used small rectangular plastic blocks (commercially available from Italveneta Didattica snc, via Risorgimento 32, 30010 Pegolotte di Cona, VE, Italy). These blocks were chosen as they have been already successfully used in numerical perception studies with naïve chicks (Rugani, Regolin, et al., 2010). Each block was a parallelepiped with a length of 1 cm, a width of 1 cm, and a height of 4 cm. They were arranged in random configurations of 4 and 12 elements and suspended in mid-air at each end of the apparatus, see Figure 4.2. In two separate conditions, either pink or red blocks were used, as initial observations seemed to suggest that stimuli color was relevant to determine the animals' approach to the stimuli.

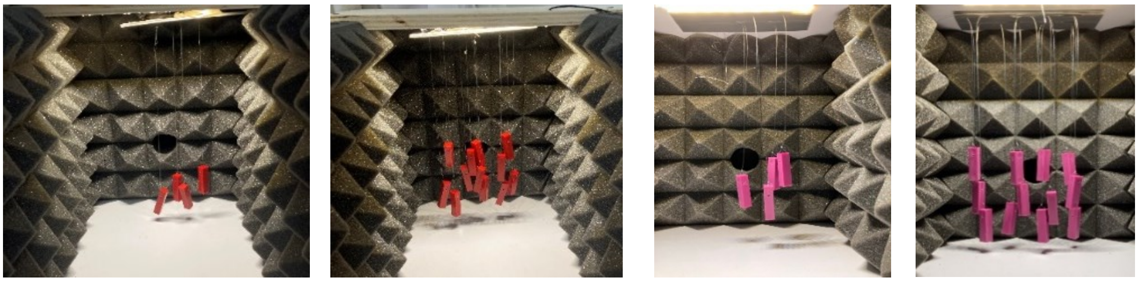


Figure 4.2: Colored blocks used as visual stimuli
The arrays of 4 and 12 blocks for each color used as visual stimuli in Experiment 3 and 4

4.2.1.3. Apparatus and procedure

The apparatus was the same as used in the previous study (see paragraph 3.2.1). In addition. The sets of 4 and 12 objects were suspended by a fine transparent thread at about the chick's head height at the end of each corridor, right in front of the speakers (Figure 4.3). The test procedure was the same as for the previous study, see paragraph 3.2.3, with the crucial difference that for each subject the same sequence, either 4 sounds or 12 sounds, was played for the whole duration of the test (including 1 minute of pre-exposure and 6 minutes of test). Thus, during the test, each subject was exposed only to one of the two sequences. The same sequence was played asynchronously between the left and right speakers with a 5 s pause between each other. The starting sequence and side of the apparatus were fully counterbalanced across subjects. In this way, one side of the apparatus had a *matching* number of visual elements and played sounds, while the other side presented *mismatching* visual and acoustical information.

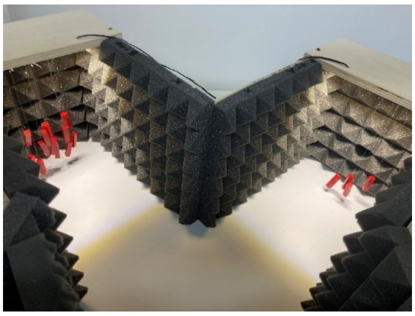


Figure 4.3: Test setting for Experiments 3 and 4
Picture of the apparatus during the test phase

4.2.1.4. *Statistical analysis*

For each group (pink and red stimuli) we measured the mean preference for the congruent location (i.e. the side with a number of visual objects congruent with the number of sounds played during the test) with the index described in 3.2.4. As for previous experiments, a non-parametric *two-tailed one-sample Wilcoxon signed rank test* was used to test preference against chance level *($\mu = 50$)*. A two-way analysis of variance (ANOVA) was performed to test Sex (Male vs Female) and Background audio (4-sound vs 12 sounds) effect, and homoscedasticity was verified with Mauchly's test: $P > 0.05$.

We also analyzed raw time spent close to one stimulus or the other. A three-way mixed ANOVA was performed with Sex and Background audio as between factors and Stimulus (familiar vs non-familiar numerosity) as within factor.

4.2.2. Results

Contrary to initial observations, the analysis on preference score revealed no significant main effect ($F_{(1, 141)} = 0.098$, $p_{value} = .775$) or interactions of Stimuli Color (Stimuli color*Sex: $F_{(1, 141)} = 0.190$, $p_{value} = .663$; Stimuli color*Background: $F_{(1, 141)} = 1.013$, $p_{value} = .316$; Stimuli color*Background*Sex: $F_{(1, 141)} = 1.170$, $p_{value} = .281$). Therefore, the subsequent analysis has been carried out collapsing on the two color groups.

Preference score

A two-way 2x2 ANOVA on the preference for congruent numerosity revealed a main effect of Background Audio ($F_{(1,145)} = 8.255$, $p_{value} = .005$, $\eta_p^2 = 0.054$). In particular, chicks listening to a 12-sound sequence during the test showed a significant preference per the congruent stimulus (Mean = 65.1%, SEM = 5.64%, *two-tailed one-sample Wilcoxon signed rank test:* $W_{(72)} = 1690$, $p_{value} = .0166$), while those that were listening to 4-sound sequence showed a tendency to approach incongruent stimulus (Mean = 39.8%, SEM = 5.53%, *two-tailed one-sample Wilcoxon signed rank test:* $W_{(31)} = 1193$, $p_{value} = .078$). No main effects of Sex ($F_{(1,145)} = 3.451$, $p_{value} = .065$) and non-significant trend for interaction between Sex and Background Audio was significant ($F_{(1,145)} = 0.087$, $p_{value} = .768$). The results are shown in Figure 4.4.

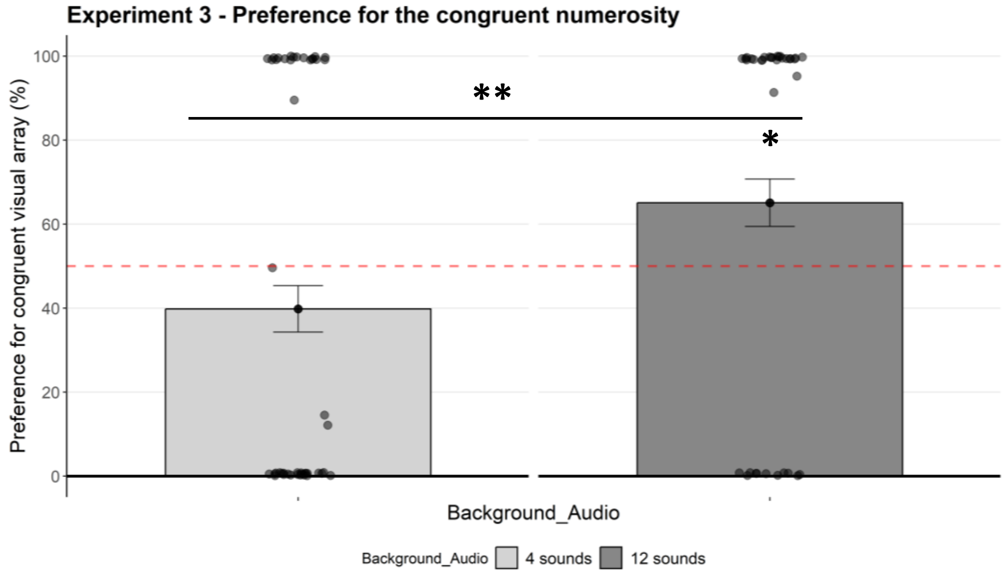


Figure 4.4: Results of Experiment 3 (Congruency preference score)
Overall preference for the congruent number of visual elements in the two Background audio conditions. The red dotted line indicates the chance level. The means and standard errors of the mean are reported (Anova analysis, ** = $p_{value} < .01$; two-tailed one-sample Wilcoxon signed rank test, * = $p_{value} < .05$).

We thus decided to also analyze the data in relation to larger numerosity preference. We performed a two-way 2x2 ANOVA that revealed no significant main effects (Sex: $F_{(1,145)} = 3.451$, $p_{value} = .065$, Background audio: $F_{(1,145)} = 0.248$, $p_{value} = .619$) or interaction ($F_{(1,145)} = 0.087$, $p_{value} = .768$). After collapsing all the data, we found a significant preference for the larger numerosity (Mean = 62.6%, SEM = 3.94%, *two-tailed one-sample Wilcoxon signed rank test:* $W_{(148)} = 6959$, $p_{value} = .0033$, Z = 0.241, Figure 4.5).

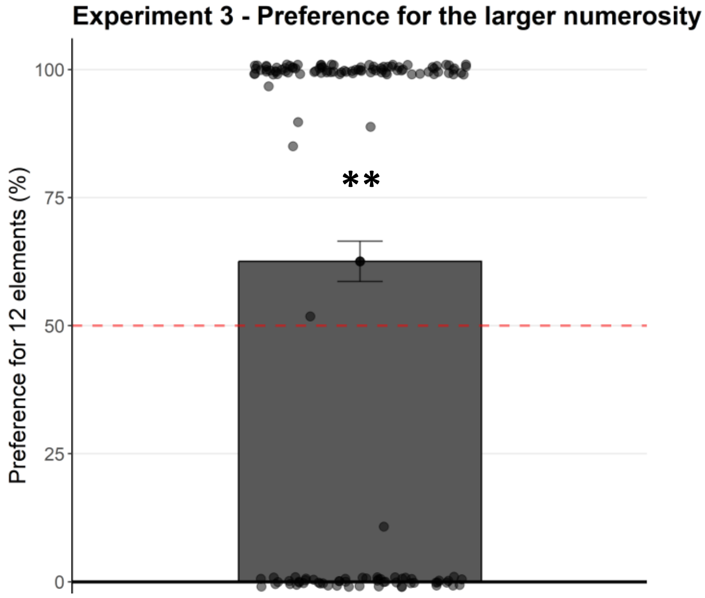


Figure 4.5: Results of Experiment 3 (Larger preference score)
Overall preference for the 12-sound sequence in the whole sample. The red dotted line indicates the chance level.
The mean and the standard error of the mean are reported (two-tailed one-sample Wilcoxon signed rank test, ** = p_{value} <.01).

Absolute time spent closer to stimuli

We first analyzed the time spent by chicks to congruent and incongruent stimuli. The three-way 2x2x2 ANOVA revealed no significant main effects (Sex: $F_{(1,145)} = 2.986$, $p_{value} = .086$; Background Audio: $F_{(1,145)} = 0.001$, $p_{value} = .974$ and Stimulus: $F_{(1,145)} = 0.584$, $p_{value} = .446$). Also, the Sex per Background Audio interaction ($F_{(1,145)} = 1.379$, $p_{value} = .242$) and the Sex per Stimulus interaction ($F_{(1,145)} = 0.022$, $p_{value} = .881$) were not significant. Instead, the Background Audio per Stimulus interaction ($F_{(1,145)} = 5.210$, $p_{value} = .024$, $\eta_p^2 = 0.035$), and the Sex per Background Audio per Stimulus interaction ($F_{(1,145)} = 4.725$, $p_{value} = .031$, $\eta_p^2 = 0.032$) were significant.

Pairwise comparisons revealed that females in the 4 sounds Background condition spent more time closer to the incongruent numerosity (Mean = 136 s, SEM = 19.2 s) than to the congruent numerosity (Mean = 71.9 s, SEM = 16.3 s; *two-tailed two-sample Wilcoxon signed rank test:* $W_{(42)} = 309$, $p_{value} = .048$, Z = 0.302), while females in the 12 sounds Background condition spent more time closer to the congruent numerosity (Mean = 161 s, SEM = 19.4 s) than to the incongruent numerosity (Mean = 64.6 s, SEM = 17.5 s; *two-tailed two-sample Wilcoxon signed rank test:* $W_{(42)} = 668$ $p_{value} = .018$, Z = 0.359). None of these effects were present in males, neither in the 4-sound background condition (*two-tailed two-sample Wilcoxon signed rank test:* $W_{(33)} = 322$ $p_{value} = .685$) nor in the 12-sound background condition (*two-tailed two-sample Wilcoxon signed rank test:* $W_{(42)} = 232$, $p_{value} = .766$). The results are shown in Figure 4.6.

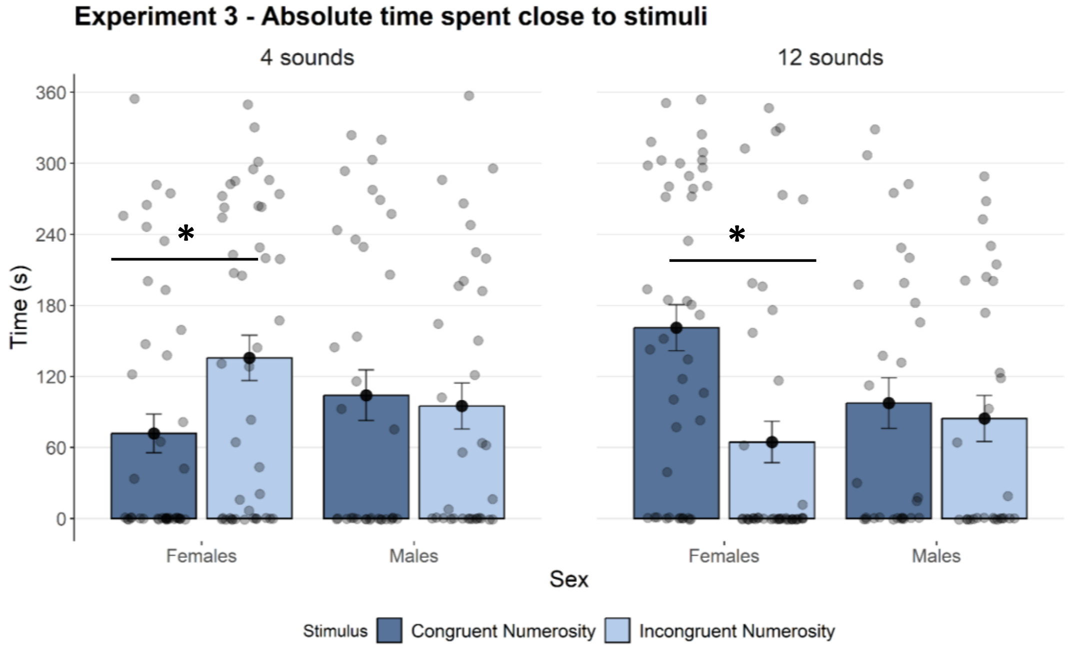


Figure 4.6: Results of Experiment 3 (Congruent vs Incongruent absolute time)
Absolute time spent close to the congruent and incongruent number of visual elements by Background audio and Sex. The means and standard errors of the mean are reported (two-tailed two-sample Wilcoxon signed rank test,* = $p_{value} < .05$).

As these results suggest that the numerosity of the stimulus itself and not familiarity is relevant in choice-making, we analyzed the time spent close to one stimulus or the other in relation to its numeric value (i.e. Larger vs Smaller). The three-way 2x2x2 ANOVA revealed no significant main effect of Sex ($F_{(1,145)} = 2.986$, $p_{value} = .086$) and Background Audio ($F_{(1,145)} = 0.001$, $p_{value} = .974$) while main effect of Stimulus was significant ($F_{(1,145)} = 5.210$, $p_{value} = .024$, $\eta_p^2 = 0.035$). The Sex per Stimulus interaction was also significant ($F_{(1,145)} = 4.725$, $p_{value} = .031$, $\eta_p^2 = 0.032$).

Pairwise comparisons revealed that females spend more time closer to the larger stimulus (Mean = 148 s, SEM = 13.6 s) than to smaller stimulus (Mean = 68.3, SEM = 11.9 s; *two-tailed one-sample Wilcoxon signed rank test:* $W_{(85)} = 2601$, $p_{value} = .002$, Z = 0.339), while males showed no difference (*two-tailed one-sample Wilcoxon signed rank test:* $W_{(62)} = 994$, $p_{value} = .926$). No other interactions (Sex*Background Audio: $F_{(1,145)} = 1.379$, $p_{value} = .242$, Background Audio*Stimulus: $F_{(1,145)} = 0.584$, $p_{value} = .446$ and Sex*Background Audio*Stimulus: $F_{(1,145)} = 0.022$, $p_{value} = .881$) were significant. Results are shown in Figure 4.7.

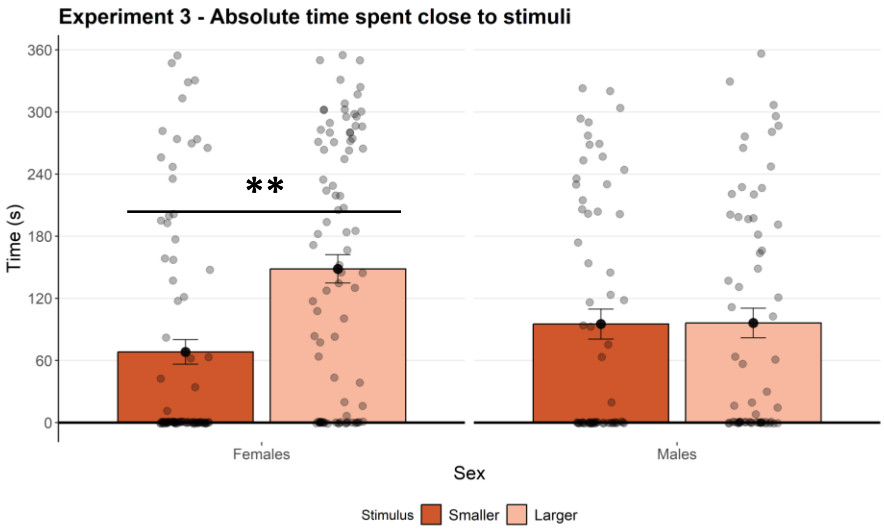


Figure 4.7: Results of Experiment 3 (Larger vs Smaller absolute time)
Absolute time spent close to smaller and larger numerosity by Background audio and Sex.
The means and standard errors of the mean are reported (two-tailed one sample Wilcoxon signed rank test, ** = p_{value} <.01).

4.2.3. Discussion

In this experiment, we investigated chicks' spontaneous preference for matching or mismatching audio-visual numerical information. As discussed in the introduction, there is convincing evidence for cross-modal numerical representation in both human and non-human animals. For instance, both rhesus monkeys and human infants have been shown to look longer at a video where the number of displayed elements matches the number of sounds played in the background (Izard et al., 2009; Jordan et al., 2005). However, a similar ability had never been tested in chicks, despite their well-known numerical capabilities.

Contrary to the results reported in human infants and monkeys, visually naïve chicks did not show any preference for approaching either the matching or mismatching side of the apparatus. Their choice was thus not influenced by the correspondence (or lack thereof) between the number of items available in the visual and in the acoustical modality. Instead, the animals showed a preference for approaching the larger set of objects they could detect visually, regardless of the number of sounds played in the background. As discussed in previous studies, a preference for a larger set of objects is a common phenomenon in domestic chicks (Lemaire et al., 2020; Pulliam, 1973; Roberts, 1996; Rugani, Regolin, et al., 2010), which is believed to reflect their social motivation to join the larger number of potential social companions. For the visual stimuli, we used blocks large enough to elicit a social response in the chicks rather than a food-seeking behavior. As mentioned, the color red (and possibly pink, due to chromatic proximity) is particularly influential in determining chicks' spontaneous preference. Therefore, it is possible that in this experiment, the number of objects was more salient to the chicks than the numerical cross-modal correspondence itself. The observed differences between sexes in the time spent near the stimuli, i.e. females spending more time closer to larger numerosities than males, seem to support this interpretation. Indeed, it has been argued that female chicks are generally more motivated to reach social companions (Vallortigara et al., 1990; see also Vallortigara and Zanforlin, 1988).

Based on the evidence we collected so far, it is impossible to determine whether the chicks did not recognize the difference between the congruent and incongruent sides of the apparatus, thus lacking cross-modal matching ability, or if they simply do not possess a spontaneous predisposition to prefer one condition over the other. It is possible that a spontaneous preference for congruency (or incongruency) existed, but it was not strong enough to override other predispositions, such as the above-mentioned preference for larger groups of companions. This possibility could be further investigated by using digital video stimuli instead of 3D objects. The advantage of 3D objects (and the original reason why we chose them) is that chicks are usually more responsive to them, increasing the chances that the animals would attend to the visual objects in our task. In future studies, instead, stimuli displayed on screen monitors could exploited as they can be more easily controlled for extensive variables that correlate with numerosities, clarifying which visual features may influence cross-modal matching. For instance, by dissociating numerosity from specific extensive continuous variables, we could investigate whether the presence of redundant numerical and non-numerical information in the visual stimuli can modulate the level of preference chicks display for the larger set of visual objects and their attention towards the audiovisual correspondence of numerical information. To partially overcome the limitations associated with spontaneous preference tests, in the next experiment, we thus tried to exploit an imprinting procedure as a way to probe the cross-modal transfer of numerical information in chicks.

Another point worth discussing is that we used relatively large numerosities, i.e., 4 and 12. As mentioned earlier, larger numerosities are generally chosen for studies investigating numerical cognition, particularly in the context of investigating *Number sense*'s properties, because they better support the argument that subjects are forming an abstract numerical representation of the stimuli, rather than tracking single objects. It has been suggested that the ability of infants to perform cross-modal matching with small numerosities is not necessarily due to an amodal representation of the numerosities perceived acoustically and visually, but rather to the representation of unitary objects with both visual and acoustic characteristics (Izard et al., 2009). According to this view, some evidence of numerical cross-modal transfer should instead be interpreted as a consequence of using an object-tracking system capable of representing objects with different sensory attributes. As larger numerosities could represent an additional level of complexity for young organisms with less developed cognitive capacities, future research should investigate the chicks' ability to perform numerical cross-modal transfers in the presence of small numerosities. Similar studies would help to shed light on the intricate relationship between ANS and cross-modal correspondences.

4.3. Experiment 4
Cross-modal imprinting

In the following experiment, we aimed to test whether chicks familiarised with a numerosity in the auditory modality would later recognize it visually. Specifically, we investigated if chicks imprinted on auditory sequences of different numerosities would show a preference for a sequence composed of the same number of visual events. Compared to our previous acoustical imprinting test (Experiment 2), we assumed that presenting the stimuli in a different sensory modality during the test could more effectively motivate the chicks to express a preference for the familiar stimulus rather than for the larger numerosity. In fact, here, neither of the two stimuli is perceptually identical to the imprinting stimulus, as it was the case in Experiment 2. In this situation, the animals might be more inclined to seek out familiar, non-perceptual characteristics, such as numerosity. Indeed, in past studies, it has been shown that the level of preference for the familiar features of the imprinting stimulus can be modulated by the level of novelty offered by the test stimuli (Vallrtigara & Andrew, 1991; Vallortigara, 1992; Versace, Fracasso, et al., 2017; see also references in Santolin et al., 2020). Moreover, compared to Experiment 3, here the presentation mode is consistent between visual and auditory stimulation. Indeed, while in Experiment 3 chicks had to match a visual array of either 4 or 12 blocks to an auditory sequence of 4 or 12 sounds (i.e. simultaneous presentation and sequential presentation respectively), in this experiment, both visual and acoustical stimuli are presented in a sequential manner. Even though it is known that chicks can process numerosities in both simultaneous and sequential formats (Rugani et al., 2009), a recent study conducted with crows argued for a neuronal two-stage process for these two formats of numerosities representations in the avian brain (Ditz &

Nieder, 2020). Therefore, given also the results of Experiment 3, to further facilitate the matching of acoustical and visual numerical information, here we employed a consistent presentation format across the two sensory modalities.

4.3.1. Methods

4.3.1.1. Subjects

We tested 100 animals on P1, 41 animals (21 female) were imprinted on a 4-sound sequence and 59 animals (40 female) were imprinted on a 12-sound sequence.

4.3.1.2. Stimuli

Acoustic stimuli

Acoustic stimuli were the same used in Experiment 1b and Experiment 2, see 3.3.2.2 for a detailed description.

Visual stimuli

Visual stimulation consisted of a single red disk flashing on a white background. The disk measured 4 cm in diameter and was centered on the screen at about the chick's head height (4 cm from the floor of the apparatus). The two visual sequences were created by flashing the red disk either 4 or 12 times. Flashes occurred at 5z Hz (0.2 s disk presentation followed by 0.2 s interval), with the 4-flash sequence lasting for a total of 1.4 s and the 12-flash sequence lasting a total of 4.6 s. During the test, the two sequences alternated asynchronously between the left and right screen with a 5 s pause between each other, mimicking the auditory sequence (see Figure 4.8). The side of the apparatus in which each sequence was presented, the starting numerosity, and the first screen of the apparatus to flash were fully counterbalanced across subjects.

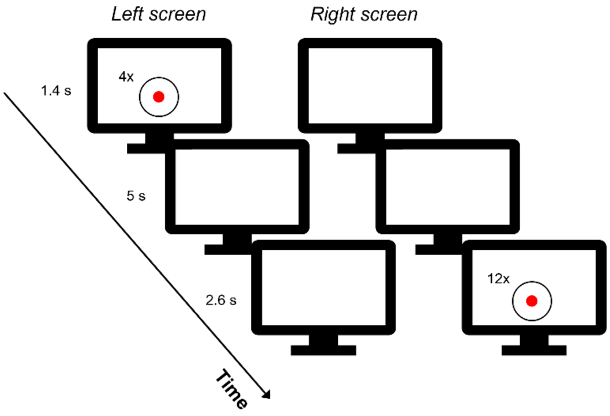


Figure 4.8: Schematic representation of visual stimulation in Experiment 4
A flashing dot was presented alternatively on the two monitors, flashing 4 times on one side and 12 times on the other side, with a 5-second pause.

4.3.1.3. Apparatus

The test apparatus was similar to the one of experiment one, with the exception that at the end of each arm, two high-frequency monitors (Asus MG248QR, 120 Hz) were placed. Choice areas were identified by fine pencil marks on the floor of the apparatus (Figure 4.9). Above the apparatus a webcam (Logitech, C922 Pro, HD stream webcam) recorded the chick's behavior, allowing for online data coding (see paragraph 3.2.4).

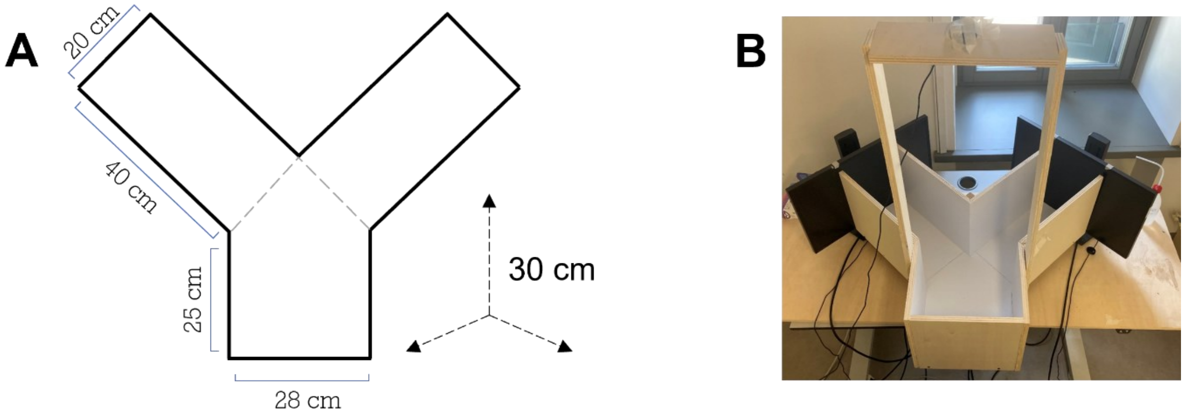


Figure 4.9: Test apparatus for Experiment 4
Panel A) Schematic representation of the apparatus with measures. The dotted lines represent the fine pencil lines delimiting the choice area.
Panel B) Photo of the apparatus, visual stimuli were presented centered on the monitors

4.3.1.4. Procedure

The imprinting procedure was the same as described in the previous study (see paragraph 3.4.1.2). As in the previous study, on test day each chick was carried from the dark incubator to the test apparatus inside a closed, opaque box ensuring no visual exposure until the beginning of the experiment. The test room was kept dark, except for light coming from the screens. Each chick was in turn placed in the central area of the apparatus, with the beak facing one of the long sides of the apparatus, and its behavior was recorded for 6 minutes. At the two ends of the corridors visual stimulation was displayed (see *Stimuli* paragraph), both the entering position of the chicks (i.e. pointing direction of the beak) and the position of the stimuli were counterbalanced across subjects. As described in paragraph 3.2.3 there was one minute of confined exposure to the stimuli to allow the chicks to observe both of them before making a choice. Data collection was the same as described in 3.2.4. After the test chicks were placed in a different incubator and then moved to the animal house.

4.3.1.5. *Statistical analysis*

For each group, we measured the mean preference for the familiar numerosity (i.e. the numerosity chicks were imprinted with) using the index described in 3.2.4. As for previous experiments, a non-parametric *two-tailed one-sample Wilcoxon signed rank test* was used to test preference against chance level ($\mu = 50$). A two-way analysis of variance (ANOVA) was performed to test the effects of Sex (Male vs Female) and Imprinting condition (4-sound vs 12 sounds), and homoscedasticity was verified with Mauchly's test: $P > 0.05$.

We also analyzed raw time spent close to one stimulus or the other. A three-way mixed ANOVA was performed with Sex and Imprinting condition as between factors and Stimulus (familiar vs non-familiar numerosity) as within factor.

4.3.2. Results

Preference score

A two-way 2x2 ANOVA on preference for familiar stimulus revealed no main effects (Sex: $F_{(1,96)} = 1.467$, $p_{value} = .229$; Imprinting condition ($F_{(1,96)} = 0.095$, $p_{value} = .229$), and no significant interaction ($F_{(1,96)} = 0.001$, $p_{value} = .984$). Therefore we collapsed the data over both Sex and Imprinting, and we found a non-significant trend in preference for familiar numerosity (Mean = 54.2 %, SEM = 4.91 %; *two-tailed one-sample Wilcoxon signed rank test:* $W_{(99)} = 2992$, $p_{value} = .073$, Figure 4.10).

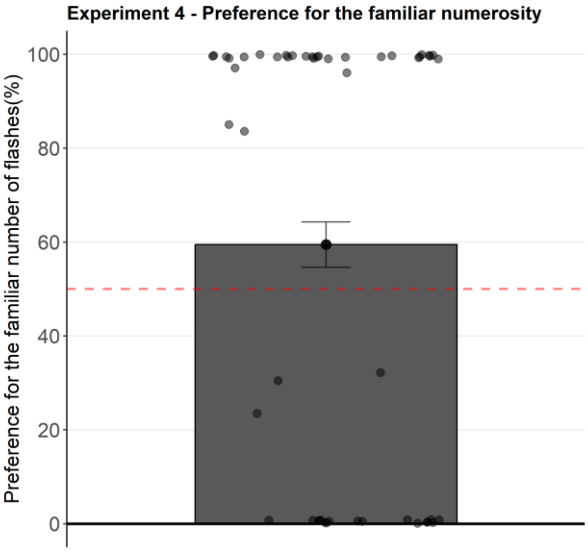


Figure 4.10: Results of Experiment 4 (Familiar preference score)
Overall preference for the familiar numerosity in the whole sample. The red dotted line indicates the chance level.
The mean and the standard error of the mean are reported.

Indeed, upon visual inspection (Figure 4.11), we can see that both imprinting groups tend towards a preference for the familiar numerosity, although this was not significant for either group. Also, a slightly higher preference was present in the group imprinted on the 12-sound sequence.

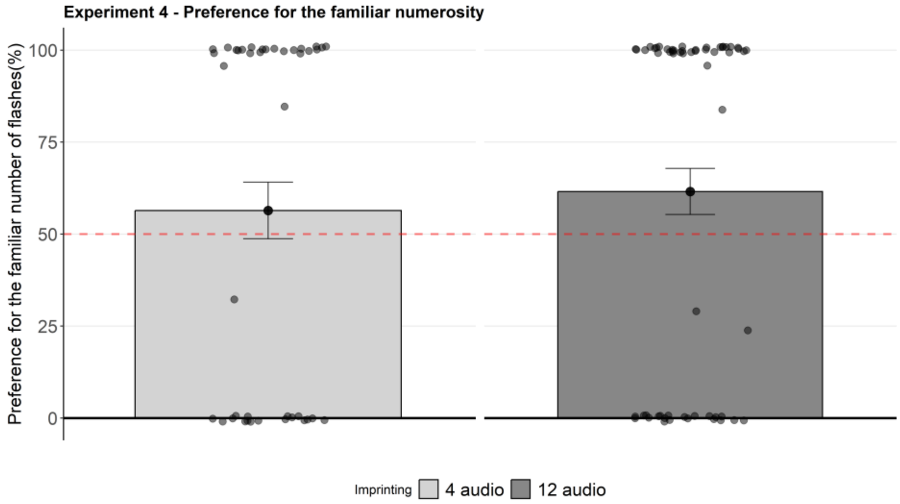


Figure 4.11: Results of Experiment 4 (Familiar preference score)
Preference for the familiar numerosity by Imprinting condition.
The means and standard errors of the mean are reported.

Absolute time spent closer to stimuli

The three-way 2x2x2 ANOVA on time spent close to stimuli no significant main effects of Sex ($F_{(1,96)} = 0.096$, $p_{value} = .757$), Imprinting Condition ($F_{(1,96)} = 0.652$, $p_{value} = .421$) and Stimulus ($F_{(1,96)} = 0.986$, $p_{value} = .323$, Figure 4.12). No significant interactions were found (Sex*Imprinting: $F_{(1,96)} = 0.011$, $p_{value} = .918$; Sex*Stimulus: $F_{(1,96)} = 1.977$, $p_{value} = .163$; Imprinting*Stimulus: $F_{(1,96)} = 1.867$, $p_{value} = .175$; Sex*Imprinting*Stimulus: $F_{(1,96)} = 0.701$, $p_{value} = .405$).

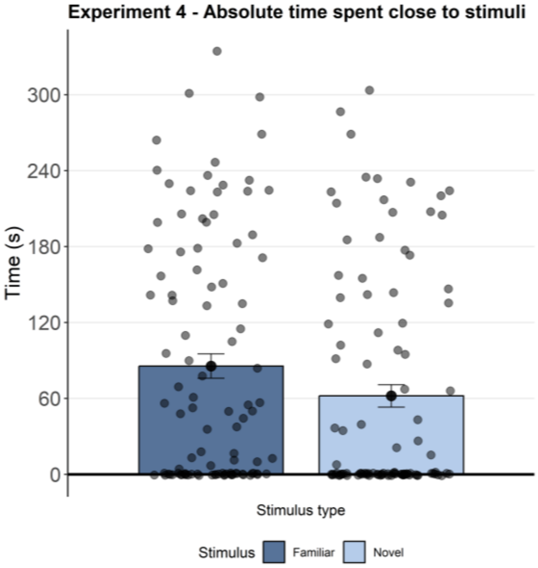


Figure 4.12: Results of Experiment 4 (Familiar vs Novel absolute time)
Absolute time spent close to familiar and novel stimuli by all the chicks.
The means and standard errors of the mean are reported.

4.3.3. Discussion

In this experiment, we investigated whether chicks could transfer numerical information acquired through acoustical imprinting to another sensory modality. Overall, we found no significant effects, neither in terms of a preference for the familiar stimulus nor for the larger numerosity. However, we would like to offer some speculative considerations, with the *caveat* that further experiments and/or a substantial increase in sample size would be necessary to substantiate them. It is worth noticing, that this is one of the two experiments among those presented in this thesis that did not yield any significant preference for the larger numerosity. In contrast, the data might be suggestive of an overall preference for the familiar stimulus. This is encouraging, given the strong bias for larger stimuli commonly observed when testing numerosity discrimination in domestic chicks (Lemaire et al., 2020; Rugani, Regolin, et al., 2010). As we have extensively discussed in previous experiments (Experiment 2 and Experiment 3), this robust predisposition to approach larger groups can mask other underlying effects. For instance, in the current experiment, a preference for the familiar numerosity (as a result of imprinting) and the predisposition to approach larger sets, could coexist originating the pattern observed in Figure 4.11 where preference for the larger numerosity looks slightly higher for chicks imprinted on 12 sounds.

The results of this experiment could be related to the modality of stimulus presentation. Instead of using visual arrays composed of a specific number of elements, we presented a single stimulus that flashed a specific number of times, mimicking the temporal structure of the auditory sequences. This approach may have had two main advantages compared to Experiment 3. First, the temporal structure of the stimuli to be compared was consistent. This could have favored the recognition of the similar temporal structure present in the sequence of flashes. However, this also means that we cannot exclude that chicks compared the auditory and the visual stimulation on such features rather than numerical information. The results could suggest a tendency to prefer the visual sequence that was familiar in terms of rhythm, rather than the number of sounds, as there is evidence that chicks can discriminate different acoustic rhythms (Chiandetti & Vallortigara, 2011; De Tommaso et al., 2019). Further research is necessary to disentangle these two components. For example, future studies should be carried out with sequences of stimuli that match in number but not in frequency rate. A second advantage, instead, is that the sequential presentation of the visual stimuli could have mitigated the spontaneous preference for larger groups. Unlike visual arrays, sequential flashes are less likely to be interpreted as indicating the number of social companions present in a group (e.g., this could represent the repeated appearance of the same individual objects), thereby reducing the motivational biases that could have led to a preference for larger numerosities in other experiments (Fontanari et al., 2011).

Furthermore, as discussed earlier, the difference between spontaneous choice paradigms and imprinting paradigms could account for the emergence of different effects. Indeed, beyond the advantages of the specific stimulus modality used in this experiment, it is also possible that in a task where learning mechanisms are emphasized, as in imprinting, cross-modal correspondences can be learned even in the absence of a spontaneous preference for congruent numerical information. For instance, when a naïve chick is presented with congruent versus incongruent options, it may lack a clear motivation to choose one over the other. However, when shown a stimulus resembling what it was familiarized with during imprinting, the chick's motivation to approach the perceived familiar stimulus likely increases. In conclusion, this experiment presents some methodological strengths that could be further explored to test numerical cross-modal correspondences in domestic chicks.

4.4. General Discussion

The broad competence of chicks in numerical cognition, and the similarity between their abilities and those of other species, provide key evidence supporting the evolutionary origins of the Approximate Number System (ANS). The conservation of certain cognitive abilities across evolution suggests that they confer significant behavioral advantages, and regarding numerical cognition, these advantages could manifest in more efficient foraging and mating strategies. Cross-modal correspondences are another example of evolutionarily conserved phenomena that seem to offer substantial advantages. For instance, as discussed in the introduction, it has been shown that multisensory information can enhance numerical matching abilities in young children. In a number-matching computer game, children's performance improved when multimodal information was presented, suggesting that combining sensory inputs, known as intersensory redundancy, enhances their ability to match quantities, at least in early childhood (Jordan & Baker, 2011). In this context, we anticipated that chicks might also demonstrate this ability, as studies on both humans and non-human animals suggest that cross-modal matching could provide significant ecological advantages.

The debate on cross-modal numerical representation in domestic chicks remains open. Our studies suggest that, while chicks may be capable of transferring numerical information across sensory modalities, such as from auditory to visual, this ability does not appear to be a spontaneous preference. A previous study by Fontanari and colleagues demonstrated that domestic chicks can use spatiotemporal information to form object representations. The study showed that when chicks were presented with two different objects simultaneously versus one object shown twice in sequence, they preferred to approach the larger group of objects (see Fontanari et al., 2011 for details on appropriate control experiments). This finding suggests that chicks can relate spatiotemporal information to numerical information, thus our results cannot indicate lacking this ability in chicks. The numerical cross-modal transfer may occur only when necessary, and when no other options are available as in the imprinting paradigm. While it does not seem to be a strong spontaneous behavior, particularly when compared to the spontaneous preference for approaching the stimulus with a larger numerosity.

However, it is important to note that evidence for cross-modal numerical matching appears generally inconsistent in the literature, both across human and non-human animals. For example, in 1987 Davis & Albert attempted to replicate the study previously conducted by Church & Meck in 1984 but were unsuccessful. Davis & Albert trained rats to discriminate between sequences of 3, 2, or 4 sounds and found no evidence that the rats could transfer their auditory numerical discrimination to the visual modality when presented with sequences of 2, 3, and 4 lights (Davis & Albert, 1987). The authors suggested that the rats in the original study made dichotomous, intensity-based judgments (e.g., associating a less intense sound with a less intense light) rather than performing a cross-modal match based on absolute numerical value. Similarly, studies investigating cross-modal representation across different sensory modalities have produced contradictory findings (see Jordan & Brannon, 2006). For instance, Starkey and colleagues initially found that infants preferentially looked towards the same number of objects as the sounds they were hearing (Starkey et al., 1983, 1990). However, subsequent studies, in which features such as the rate and tone of the sounds were manipulated, failed to replicate those findings as they did not observe any significant preference for either the equivalent or non-equivalent visual display (Mix et al., 1997; Moore et al., 1987).

The nature of the stimuli themselves may facilitate cross-modal mapping. For example, infants as young as 6 months old succeeded in a violation-of-expectation paradigm involving numerical cross-modal matching. In this paradigm, infants were first familiarized with a video showing two objects sequentially impacting a surface, each producing a sound. During the test phase, they heard either two or three sounds while the screen was covered. After a brief delay, the screen was raised to reveal two or three objects. Infants looked significantly longer when the number of objects on the screen matched the number of sounds they had heard, suggesting that they had formed an expectation based on the auditory information, which was then violated (Kobayashi et al., 2005). The ecological relevance of the stimuli may have facilitated this cross-modal correspondence, as suggested by Jordan and Brannon. In one of their study, they demonstrated that 7-month-old infants match the number of voices they heard with the number of faces presented on a screen (Jordan & Brannon, 2006). The authors discussed two important points when comparing their findings with the previous conflicting findings reported in the literature. First, the ecological relevance of the presented situations appears to be relevant for cross-modal transfer to emerge. Second, the role of continuous variables and the structure of the experimental design (e.g. within-subject versus between-subject) are crucial. They argued that a between-subject design could facilitate cross-modal matching based on non-numerical information, as participants have the opportunity to hear both sounds and see both visual stimuli. In such cases, the preference for matching or mismatching conditions may not reflect true cross-modal correspondence but rather a decision based on other dimensions, such as relative intensity. Jordan and colleagues' studies, in 2005 with monkeys and in 2006 with infants, took these points into account. First of all, they presented ecologically relevant situations, such as the faces and vocalizations of conspecifics. Then, both auditory and visual stimuli (2 vs. 3) were presented simultaneously, with the duration of the composite audio streams and videos matched. This ensures that participants cannot rely on cues related to the total duration to solve the task. Furthermore, by using a between-subjects design, they avoided the possibility of participants learning to match the more intense or complex auditory stimulus with the more intense or complex visual stimulus, as participants only experienced one of the two stimuli (either 3 or 2) in the auditory modality. In both studies the authors found that participants were looking longer at the video that showed the same number of conspecifics of the vocalization they were hearing, demonstrating reliable numerical cross-modal competencies (Jordan et al., 2005; Jordan & Brannon, 2006). In our studies, similarly, each chick was exposed to only one auditory sequence while facing both visual arrays in Experiment 3 or was imprinted to only one auditory sequence in Experiment 4. Therefore one of the criteria above mentioned for numerical cross-modal transfer was met. On the contrary, our stimuli were not ecologically relevant, as neither the visual nor the auditory stimulation were naturalistic stimuli. An additional key difference from Jordan and colleagues' studies, however, is that we used larger numerosities instead of a 2 vs 3 comparison. As previously discussed, this might lead to different representational phenomena, as larger numerosities may not be tracked as unitary elements with multimodal features, unlike smaller sets of 2 or 3 elements. However, it has been demonstrated that human newborns can perform cross-modal matching of large numerosities even when they cannot rely on other non-numerical information, even when presented with larger numerosities (Izard et al., 2009). In the study by Izard and colleagues, the habituation auditory sequence was a between-subject factor, meaning that one group listened to 12 sounds while facing 4 and 12 dots, and the other group listened to 4. Despite this, newborns successfully completed the task, as both groups looked longer at the visual numerosity that matched the auditory sequence, suggesting cross-modal pieces of evidence.

In conclusion, the lack of evidence for cross-modal numerical transfer in our study may not be due to limitations in the paradigm or the numerosities used, but rather other factors. In Experiment 3, the strong spontaneous preference for the larger group of elements may have overshadowed other behaviors. In Experiment 4, the situation presented to the chicks may not have been ecologically relevant enough. The observed trend toward a preference for the familiar numerosity, though not statistically significant, is nonetheless encouraging. Future studies should explore whether using stimuli more similar to those in Jordan's studies (e.g., imprinting with naturalistic calls and video presentations of chicks during testing) might yield significant results.

Chapter 5

Conclusion

Overall, we have observed that while there are significant innate components in numerical cognition, much remains to be discovered. As reported in the introduction, there is evidence suggesting that the ANS might originate from an evolutionary mechanism shared between humans and non-human species (Brannon & Roitman, 2003). And if the basic mechanism is shared (and potentially innate) across cultures and species, why shouldn't some of its specific characteristics also be shared (and potentially innate)? In this regard, we investigated the SNA phenomenon and the numerical auditory/cross-modal perception.

Study 1 fits within the framework of research aimed at understanding whether SNA are the result of cultural experiences or if they might refer to some innate predisposition. For about a decade, there has been evidence of SNA in domestic chicks and monkeys (Cantlon & Brannon, 2006; Drucker & Brannon, 2014; Rugani, Vallortigara, et al., 2011; Rugani et al., 2015; Rugani, Vallortigara, et al., 2016). More recently, evidence of SNA has been found in even more evolutionarily distant species such as bees, as well as in infants (de Hevia et al., 2017; Di Giorgio et al., 2019; Giurfa et al., 2022). However, given the strong impact that cultural contexts have on SNA in adult humans, the debate about its origin, cultural or biological, is still open (Göbel et al., 2011).

Researches that question an innate predisposition for SNA emphasize the role of cultural experiences in determining these associations. For example, Casasanto and colleagues propose the *CORrelations in Experience (CORE) principle*. The CORE principle posits that abstract domains are spatialized in people's minds according to how they are spatialized in the world. In particular, they claim that ‹‹*conceptions of time are selectively shaped by cultural practices that spatialize time, whereas conceptions of number are selectively shaped by practices that spatialize numbers*›› (Pitt & Casasanto, 2020). According to the authors, both space-time association and space-number association can be understood as mental metaphors, meaning mental constructs that metaphorically map point-to-point correspondences between continuous analog scales across different conceptual areas. In these metaphors, the source domain, such as space, provides a framework for understanding the target domain, like time or number (Casasanto & Hall, 2010; Lakoff & Johnson, 1980). According to this view, writing/reading habits unequivocally shape time over space, indeed during reading each new fixation occurs later in time and farther to the right in space but does not necessarily shape numbers into space, as reading plain text does not create a direct experienced correlation between numbers and space. The authors claim that studies correlating their SNARC results with reading/writing habits might have instead activated broader culture-specific associations. Possibly any cultural practice could have influenced how participants map numbers into space numbers, such as counting on fingers or encountering numbers in spatial arrangements like rulers, calendars, graphs, and computer keyboards. In summary, the spatial organization in everyday life of numeric-related contents would serve as a scaffolding for numerical space mapping (see Pitt & Casasanto, 2020 for an extensive review). Although this proposal accounts for SNA in literate adults and could explain how explicit culturally-determined forms of SNA might develop over a lifetime, it does not seem to fully explain the results of studies found in animals and human newborns. Indeed, what is peculiar about these latter studies is that not only are SNA found to be consistent within a species, but in all non-human species where SNA have been found, these have always been revealed to be oriented from left to right.

There is another theory that seeks to explain SNAs without necessarily considering it specifically number-dependent: the *Brain's Asymmetric Frequency Tuning* (BAFT) hypothesis. This hypothesis, as we have already mentioned, states that SNA might actually be the result of differential hemispheric involvement in response to different visual spatial frequency (SF) bands (Felisatti et al., 2020). This theory suggests that there is a parallel between the processing of SF bands in the brain and mapping numbers into space. Indeed, the processing of SF in vertebrate brains seems to be also lateralized: the right hemisphere is more attuned to low SFs, while the left hemisphere is more attuned to high SFs (Christman, 1989). Given the fact that the fibers connecting the eyes and brain cross at the optic chiasm, each brain hemisphere processes visual input from the opposite side of the visual field. In natural scenes, there is often a correlation between the numerosity of elements in a scene and SFs, where scenes with fewer elements low SF range, and those with many elements have a high SF range. Therefore the cognitive association of fewer items with left space and more items with right space may be rooted in the brain's hemispheric asymmetry in SF selection. (Felisatti et al., 2020, see Figure 5.1). However, as we have already discussed, this theory might not explain, for instance, the results obtained in the study by Di Giorgio and colleagues in 2019 (Di Giorgio et al., 2019). Moreover, Rugani and colleagues (2020), in a study exploiting the same paradigm used in 2015 (see Figure 1.5), demonstrated that chicks show left-to-right SNA irrespectively of the adaptation to high or low spatial frequency given by the different tests' order.

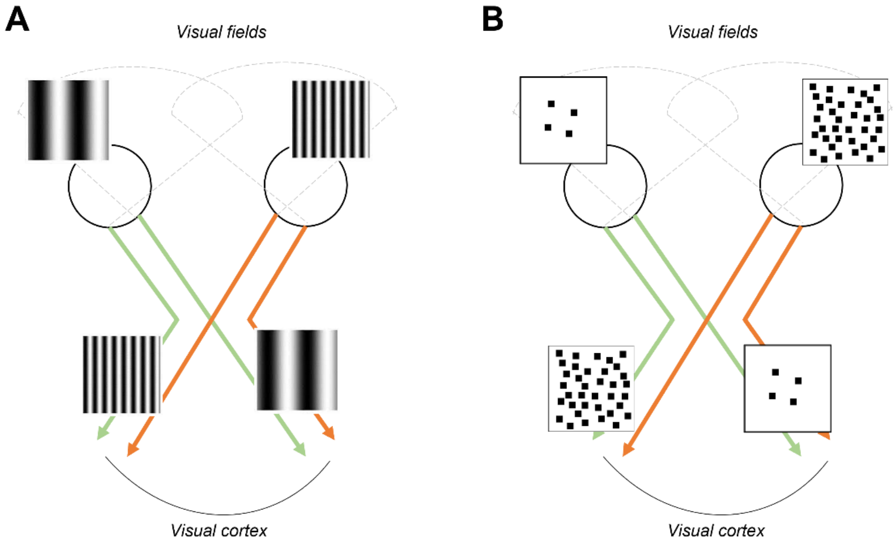


Figure 5.1: Schematic representation of the BAFT theory
Panel A) The left hemisphere preferentially filters for higher Spatial Frequency bands, while the right hemisphere preferentially filters for lower Spatial Frequency bands.
Panel B) According to the BAFT theory, given the spatial frequency components of small and large sets of dots, numerosities are accordingly processed in the two hemispheres.
Modified form Felisatti et al., 2020.

Study 1 of this thesis fits within this framework, aiming specifically to investigate how SNA might emerge in different situations. First, our Study 1a seems to rule out the possibility that SNA are solely the result of cultural influences. We found that in a numerosity comparison task, where the spatial component is implicitly processed, subjects who should theoretically not be influenced by cultural biases show a left-to-right oriented SNA. Indeed, both Himba adults, an indigenous population without a written system, and Italian preschoolers showed the same type of congruence effect as Italian adults in response to small numerosities presented on the left side of the screen. However, an equal congruence effect was not shown for large numerosities on the right suggesting that the nature of the stimuli, in this case non-symbolic, is particularly relevant. This asymmetry can be explained if we consider that the numerical value of the stimuli in relation to its spatial presentation is the relevant aspect, and also take into account how symbolic numerosities are represented differently in the brain. A series of studies have indeed shown that there is greater activation of the right hemisphere in response to non-symbolic numerosities, while the left hemisphere is more involved in response to symbolic stimuli (Dehaene & Cohen, 1997; Piazza et al., 2007; Verguts & Fias, 2004). This asymmetry is likely due to the role that language, localized in the left hemisphere, plays in processing symbolic numerosities (which are typically acquired through linguistic processes), to the point that the two hemispheres become differently involved with age (Cantlon et al., 2006; Kersey & Cantlon, 2017; Rivera et al., 2005).

Our Study 1b seems to provide further support for this interpretation of the asymmetric effect found in Study 1a. Indeed, when a similar experiment to the original one conducted with Himba and children was replicated with Italian adults, with the addition of symbolic stimuli, a response pattern was found that is equally justifiable by the above-mentioned hemispheric asymmetry. We found that in response to symbolic numerical stimuli, the congruency effect emerges for large numerosities presented on the right. Overall, these two studies seem to suggest that it is the numerical value of the stimuli that determines the emergence of SNA, at least in this kind of task. Indeed, hypotheses related to spatial frequencies do not seem suitable for explaining the asymmetry specifically in response to the symbolic or non-symbolic nature of the stimuli. If it were only a matter of visual frequency bands determining the response (BAFT hypothesis), we would have found a congruence effect on the right side as well in the non-symbolic experiments (where arrays of dots are used), and even more significantly we would have not found this asymmetry between the two hemispaces in function of symbolic versus non-symbolic format. It would be interesting for future perspectives to see what happens in the rare portion of the population that is left-handed and has inverted brain lateralization, to see if they would show opposite patterns for symbolic numerosity in Study 1b. It would also be interesting to replicate Study 1a with children from Arabic populations, potentially well-isolated from Western cultural influences, in a longitudinal perspective. What we could expect is that before the start of school, Arabic preschoolers would show the same type of pattern we found in Italian preschoolers, while growing up they would diverge from Italian adults. If, as hypothesized, the two tasks, different in their explicit and implicit SNA forms, refer to two different mechanisms, then in the preschool phase we should find results similar to the one we collected. Indeed, similarly to Italian and Himba children, when explicitly tested Arabic preschoolers should have no pre-established notions to rely on and therefore they should not show consistent explicit SNA at the population level. On the contrary, when tested implicitly they could rely on the shared biological mechanism that would lead them to show SNA oriented from left to right similar to the one of the Western population. In the early school years, instead, we should find that Arabic children show explicit SNA oriented from right to left, coherent with their reading/writing habits, while the response in the implicit task remains uncertain. One possibility is that they would exhibit the SNA observed in Italian adults, suggesting that such implicit tasks engage only biologically predisposed mechanisms. Alternatively, the cultural context may exert a strong enough influence to mask this biological predisposition, resulting in a right-to-left oriented SNA even in the implicit task. Indeed, while our data appear to support the hypothesis that explicit and implicit SNA involve distinct cognitive mechanisms, shaped by both cultural experiences and biological predispositions, it remains crucial to investigate how each mechanism develops over time and, more importantly, to what degree they may influence one another. For this reason, it would also be interesting to replicate Study 1b with non-Western adult populations. If, as we hypothesize, the response to symbolic stimuli is influenced by the hemisphere associated with language processing, particularly due to the strong cultural significance of these stimuli, it would be intriguing to investigate whether populations that do not read or write from left to right might exhibit an inverted congruence effect in similar tasks. All these future studies would help shed light on the intricate relationship that has been shown to exist between numbers, space, and presentation modes.

Even once it is demonstrated that at least some SNA phenomena have biological bases, why should they manifest specifically from left to right? While some hypotheses suggest that numerical and spatial cognition might share common neural circuits (Hubbard et al., 2005), these do not fully account for the specific left-to-right orientation of the mental number line. An alternative hypothesis, proposed by Vallortigara, links the *Valence hypothesis* with SNA. According to the *Valence hypothesis*, the anterior regions of the left and right hemispheres are specialized for processing approach and withdrawal behaviors, respectively (Davidson, 2004). Vallortigara in 2018 proposed that brain asymmetries related to motivation and emotion might account for the left-to-right orientation of the mental number line (Vallortigara, 2018). This hypothesis extends beyond human neuropsychology to suggest that fundamental motivational dimensions, such as approach and withdrawal, are present across species (see reviews in Vallortigara et al., 2018, Lichtenstein-Vidne et al., 2017; Quaranta et al., 2007; Rogers et al., 2013). For instance, animals may inherently associate larger magnitudes with positive emotions and approach behaviors, while smaller magnitudes are linked to negative emotions and withdrawal. In experiments with chicks, for example, increases or decreases in numerical quantities could evoke activation of the left hemisphere (associated with positive emotions and approach) or the right hemisphere (linked to negative emotions and withdrawal) respectively, leading to behavioral associations of left/small and right/large (Rugani et al., 2015). One of the first evidence of the correlation between SNA and valence was provided by Gerten and Topolinski in 2020. The authors demonstrated that when asked to judge the overall appearance of two digits on the screen, participants reported more positive feelings about the digit arrangements compatible with a left-to-right SNA (Gerten & Topolinski, 2021). In our first tentative attempt to validate this hypothesis, in a study conducted beyond the scope of this thesis, we tested whether Italian adults show SNA in response to positive/negative valence images (e.g. war scenario versus puppets). Preliminary results found that participants were actually faster to categorize negative valence images when they were presented on the right side of the screen and positive valence images when they were presented on the left side of the screen. Future analysis will focus on correlating the "*numerical SNA*" performance with the "*valence SNA"* performance within the same individual and expanding the sample to illiterate populations. According to the hypothesis, in implicit tasks, the valence of the stimuli should be a predictor of the SNA effect irrespectively to the cultural influences: left-to-right direction for positive stimuli and right-to-left direction for negative ones. Therefore participants who show a more lateralized behavior in response to positive valence stimuli should also manifest stronger left-to-right SNA. On the contrary, the use of negative valence stimuli should arise a right-to-left-oriented SNA. Indeed, if typical SNA behavior in response to positive/neutral stimuli emerges as a result of the "*more is better"* bias, in a negative situation "*less is better"* and we should observe a reversed SNA. However, some recent studies showed that emotional facial expressions seem to be spatial organized from left to right based on the intensity of the emotions displayed rather than on their valence value (i.e. happier faces are more rightward when judged on happiness, but more leftward when judged on angriness; Fantoni et al., 2019; Holmes et al., 2019; Holmes & Lourenco, 2011). Nevertheless, one strength of this hypothesis proposed by Vallortigara is that similar SNA could occur with continuous physical variables or other dimensions, and not be limited to just discrete numerosities. Right hemisphere activation and subsequent shifts in attention to the left side of the visual field (and vice versa) could be caused by general increases or decreases in perceptual changes (see Figure 5.2). Therefore future studies should be conducted to better isolate the role of numerical and non-numerical information, both regarding physical variables and other abstract dimensions.

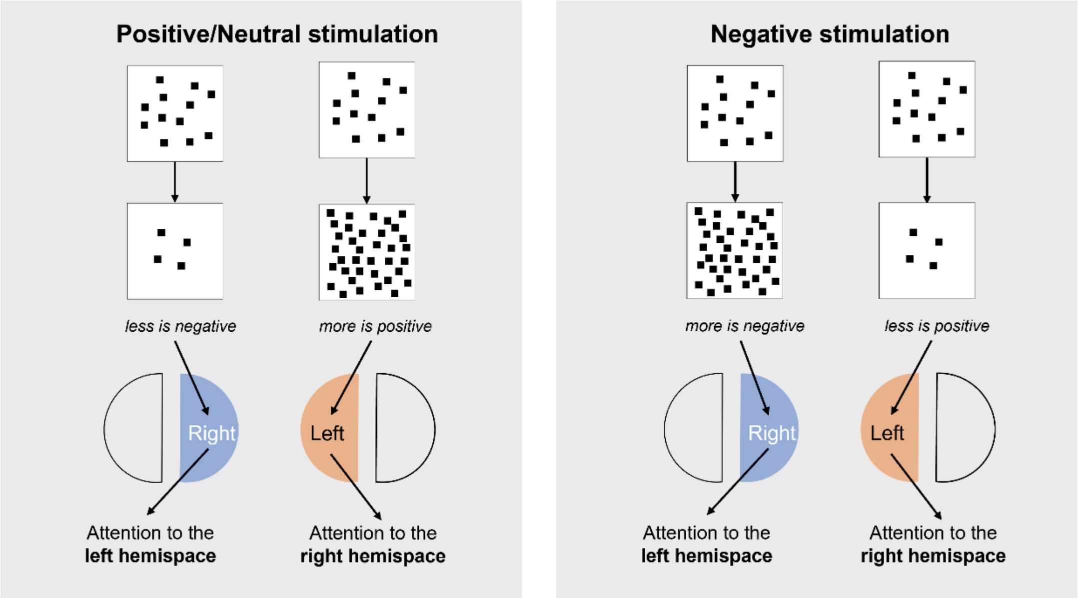


Figure 5.2: Schematic representation of the link between numerosity perception and the *valence hypothesis*
In case of positive/neutral stimuli, increasing changes in numerosity are associated with the prevalent activation of the left hemisphere (due to its positive valence value) and with the consequent attention to the contralateral hemispace. The opposite is true for decreasing changes in numerosity and the opposite pattern is expected.
Modified from Vallortigara, 2018

Our studies with chicks have not been particularly conclusive in relation to the original research questions we had, but they have nonetheless expanded our understanding of ANS in this species. Study 2 highlighted how predispositions previously observed in response to visual stimuli can also be found in response to acoustic stimuli. Additionally, Study 2 allowed us to investigate how even in the acoustic domain, confounding information from extensive variables is particularly relevant in influencing the behavior of these animals. Future studies should be conducted, particularly with a greater focus on the role of continuous variables.

As previously mentioned, the role of continuous variables is particularly significant when dealing with studies of numerical discrimination. An alternative theory to the *Number Sense* hypothesis suggests that a shared prelinguistic framework enables organisms to encode various dimensions that can be quantified as "more than" or "less than," including numbers, spatial dimensions, time, brightness, and length (Walsh, 2003; Figure 5.3). According to A Theory Of Magnitude (ATOM), processing different magnitudes simultaneously can lead to symmetrical interference across various dimensions. This is illustrated by research on human infants, who process numerical, temporal, and spatial dimensions similarly (Lourenco & Longo, 2010). Also, primates show similar behaviors, for example, Merritt and colleagues in 2010 investigated how space and time interact in both human adults and monkeys by having participants assess either the length or temporal duration of a line displayed on a computer screen. The lines varied in terms of length and duration. The findings revealed that humans experienced more interference from the spatial dimension compared to the temporal one. In contrast, monkeys exhibited mutual interference between spatial and temporal dimensions, with no significant dominance of spatial over temporal information (Merritt et al., 2010).

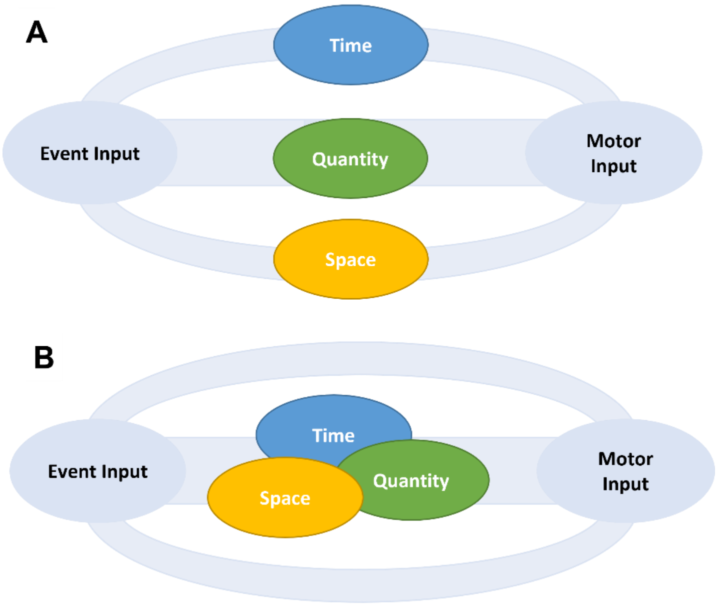


Figure 5.3: Schematic representation of possible links between different magnitudes
Panel A) Different magnitudes could be processed separately and compared in a second moment.
Panel B) According to Walsh 2003, a generalized magnitude system computes all magnitude according to a common metric. Modified from Walsh, 2003

In the specific context of our studies, controlling for extensive variables such as "total duration" and "total amount of sound" between Experiment 1a and Experiment 1b of Study 2 suggests that there was indeed an interaction between various dimensions, numerical and non-numerical, in determining the chicks' responses. In Experiment 1a, where extensive variables positively correlate with numerosity (i.e., larger numerosity corresponds to greater length and more sound), the chicks showed a preference for the larger numerosity (but also greater magnitudes), consistent with their social predispositions. In Experiment 1b, where extensive variables were equalized between the two stimuli, the chicks no longer showed this preference. These results imply that the animals' responses might have been based either on a direct comparison of extensive variables, leading them to choose the greater magnitude, or there may be a necessity for mutual support between numerical and non-numerical information to establish a preference. Further research is needed to clarify whether the chicks' responses were influenced by the magnitude or by the interaction of multiple dimensions. The need for mutual support between numerical and non-numerical information in determining preferences is something that should be investigated in future studies. For instance, one could conduct an experiment where chicks are presented with a single long sound versus a single short sound to test whether, with only extensive information, they still show a preference for the larger magnitude. Alternatively, as previously mentioned, different paradigms could be employed, such as training paradigms, to determine whether the issue is a lack of predisposition to make a particular choice at the group level or an actual inability to discriminate numerosities in the auditory domain.

In Study 3, the role of other non-numerical magnitudes is also relevant. In examining cross-modal numerical transfer, we presented chicks with acoustic sequences of either 4 or 12 sounds and visual stimuli consisting of arrays of 4 or 12 elements. We did not find any conclusive evidence of cross-modal transfer, but an encouraging trend to prefer the familiar stimulus in cross-modal imprinting (Experiment 4). According to ATOM theory, a choice for the familiar numerosity in chicks could be supported by the comparison of different types of magnitudes, such as sound duration and duration of visual stimulation or total area occupied by the visual stimuli, rather than making a direct numerical transfer. This would provide an alternative explanation for any numerical cross-modal abilities. For example, evidence supporting the ATOM theory has also been found in simpler organisms like honeybees. Bortot and colleagues demonstrated that honeybees transfer quantitative discrimination between numerosity and size, and vice versa (Bortot et al., 2020; Bortot & Vallortigara, 2023). Therefore, further research dissociating numerical and non-numerical variables is needed to explore which dimensions, this capacity might depend on in domestic chicks and how to facilitate it. Additionally, similar to what was noted in human studies, our chick studies also highlighted the importance of task demands on the emergence of specific behaviors. In Study 3, although only tendentially significant, we observed a potential difference in the emergence of cross-modal numerical abilities between spontaneous choice paradigms and imprinting. The focus of this thesis was primarily on early predispositions, so the research question was addressed from that perspective. However, it would also be interesting to determine if other paradigms reveal these cross-modal matching abilities in chicks. It could be that other types of paradigms or training might make these abilities more evident.

There are two general methodological limitations of the studies we conducted with chicks. First, a different incubation setting might be necessary because there was no acoustic isolation for these animals, which could have introduced a confound in studies where audio is relevant. Although this was not our initial approach due to technical difficulties, individual hatching and incubation should be considered for future research. Second, as observed particularly in contingency tasks, the stimuli used were not naturalistic. Future studies should consider using videos and calls of chicks and he hopes to enhance chick abilities in the function of a higher ecological value. Moreover, in the future, we could also take advantage of this paradigm to investigate if SNA is present in response to non-visual stimulation. As we already mentioned several times, evidence of SNA in chicks has been vastly reported (Rugani et al, 2020 the most recent), while still no studies on this topic have been conducted with other stimulus modalities or bi-modal stimulation.

In conclusion, numerical cognition is a fascinating area of research clearly deserving of further exploration. SNA remains a hot topic with no definitive answers to all issues it raises yet, and the studies of my thesis may contribute by offering a new perspective through the decomposition into simpler components. Of particular interest is cross-modal mapping, which appears to be a new frontier to explore, especially in chicks.

Appendix

Supplementary Figures and Table for Study 1a

I. Supplementary Figures

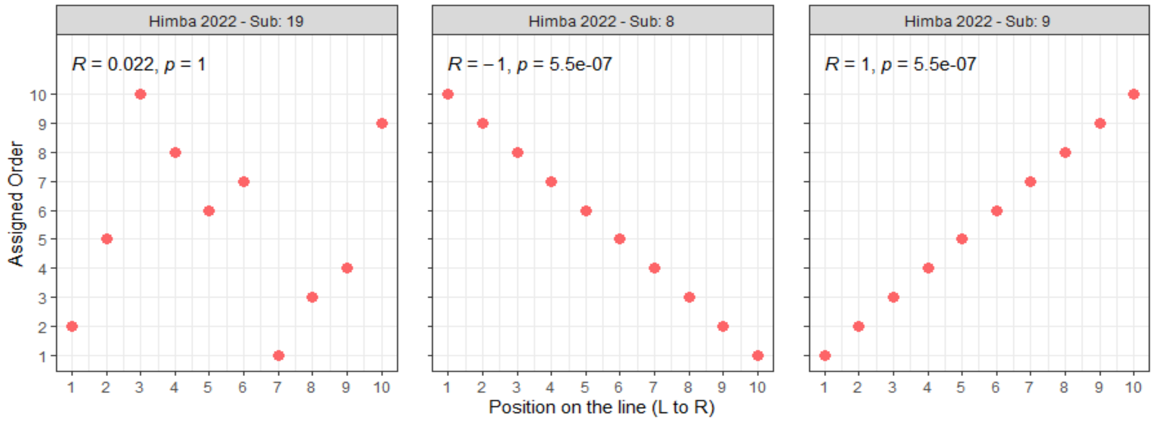


Figure S1: Examples of correlation plots between Cards order assigned by participants and 1 to 10 left-to-right order
For this task we calculated the Kendall Tau correlation between the cards' ordinal position (y-axis), and the supposed left-to-right order of the numerosities (x-axis). In the figure an example of random disposition (left graph), perfect right-to-left disposition (central graph), and perfect left-to-right disposition (right graph).

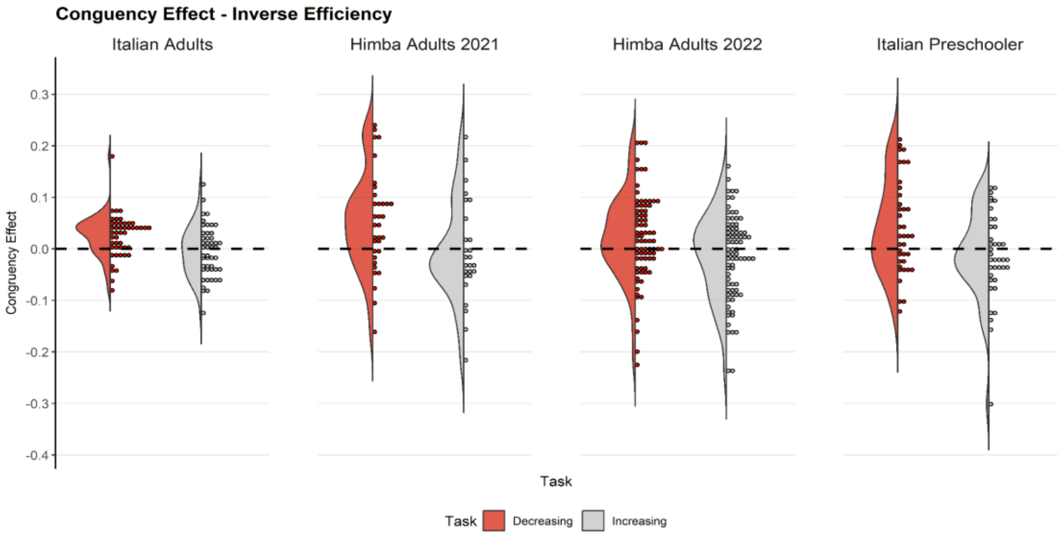


Figure S2: Distributions of congruency effect groups per Task instruction
The dashed line represents the 0-chance level (i.e. no effect of the congruent condition on the performance)

II. Supplementary Tables

One-way chi-square test for a difference from equal occurrence of shapes in each population

Group	*N*	*% for each shape*				*χ2 test statistic*		
		Lateral Line	Random	Sagittal/ Diagonal Line	Two dimensional	Df	*χ2*-value	p_{value}
Italian Adults	47	80.85	0	2.13	17.02	3	173.25	**< .0001**
Himba Adults 2021	81	74.08	18.52	3.7	3.7	3	134.29	**< .0001**
Himba Adults 2022	83	84.34	8.42	7.24	0	3	189.45	**< .0001**
Italian Preschoolers	43	65.12	18.60	2.33	13.95	3	91.455	**< .0001**

Table S1
In bold pvalue significant for alpha value = .05

Two-Way Analysis of Variance of Inverse Efficiency Scores in Himba 2021 by Task instruction and Congruency Condition

Factor	DF_n	DF_d	SS_n	SS_d	F-value	p_{value}	$\eta^2_{partial}$	η^2_p CI 90%
Task instruction	1	57	0.746	32.518	1.308	.258	0.022	[0 0.117]
Congruency Condition	1	57	0.185	2.274	3.520	0.66	0.058	[0 0.175]
Task instruction x Congruency Condition	1	57	0.196	2.274	4.901	**.031**	0.079	[0.004 0.204]

Table S2
In bold pvalue significant for alpha value = .05

Three-Way Analysis of Variance of Inverse Efficiency Scores in Groups by Task instruction and Congruency Condition

Factor	Df_n	Df_d	SS_n	SS_d	F-value	p_{value}	$\eta^2_{partial}$	η^2_p *CI 90%*
Group	2	160	100.674	59.699	134.9	**<.0001**	0.628	[0.551 0.680]
Task instruction	1	160	0.353	15.355	3.674	.057	0.022	[0 0.072]
Congruency Condition	1	160	0.024	2.089	1.379	.242	0.009	[0 0.047]
Group *x* Task instruction	2	160	0.233	15.355	1.216	.299	0.015	[0 0.051]
Group *x* Congruency Condition	2	160	0.005	2.809	0.129	.879	0.002	[0 0.011]
Task instruction *x* Congruency Condition	1	160	0.222	3.112	11.397	**.0009**	0.066	[0.017 0.136]
Group *x* Task instruction *x* Congruency Condition	2	160	0.120	3.112	3.074	**.048**	0.037	[0 0.088]

Table S3
In bold pvalue significant for alpha value = .05

Planned comparison for differences between Congruency Condition for Task instruction for each Group
One-tailed two-sample paired t-test (Condition 1 < Condition 2 alternative)

Group	*Task*	*Condition 1*	*Condition 2*	*Df*	*t-value*	*95% CI*	*p_{value}*	*Effect size*
Italian Adults	Decreasing	Congruent	Incongruent	46	-3.36	[-Inf -0.013]	<.001	-0.534
Italian Adults	Increasing	Congruent	Incongruent	46	0.676	[-Inf 0.015]	.749	0.099
Himba Adults 2021	Decreasing	Congruent	Incongruent	32	-2.762	[-Inf -0.039]	.005	-0.481
Himba Adults 2021	Increasing	Congruent	Incongruent	25	0.272	[-Inf 0.107]	.606	0.053
Himba Adults 2022	Decreasing	Congruent	Incongruent	77	-1.58	[-Inf 0.002]	.06	-0.178
Himba Adults 2022	Increasing	Congruent	Incongruent	77	0.772	[-Inf 0.039]	.779	0.088
Italian Preschooler	Decreasing	Congruent	Incongruent	37	-2.32	[-Inf -0.028]	.013	-0.376
Italian Preschooler	Increasing	Congruent	Incongruent	37	0.967	[-Inf 0.167]	.83	0.157

Table S4
In bold pvalue significant for alpha value = .05; underscored values show a tendency towards significance. Effect size measured as Cohen's D absolute value

Comparison for differences between Congruency Condition for Task instruction for each Group –
Two-tailed two-sample paired t-test

Group	*Task*	*Condition 1*	*Condition 2*	*Df*	*t-value*	*95% CI*	*p_{value}*	*Effect size*
Italian Adults	Decreasing	Congruent	Incongruent	46	-3.36	[-0.038 -0.011]	<.001	-0.534
Italian Adults	Increasing	Congruent	Incongruent	46	0.676	[-0.008 0.017]	.749	0.099
Himba Adults 2021	Decreasing	Congruent	Incongruent	32	-2.762	[-0.263 -0.039]	.009	-0.481
Himba Adults 2021	Increasing	Congruent	Incongruent	25	0.272	[-0.082 0.107]	.788	0.053
Himba Adults 2022	Decreasing	Congruent	Incongruent	77	-1.58	[-0.062 0.008]	.119	-0.178
Himba Adults 2022	Increasing	Congruent	Incongruent	77	0.772	[-0.019 0.044]	.442	0.088
Italian Preschooler	Decreasing	Congruent	Incongruent	37	-2.32	[-0.192 -0.013]	.026	-0.376
Italian Preschooler	Increasing	Congruent	Incongruent	37	0.967	[-0.067 0.189]	.83	0.157

Table S5
In bold pvalue significant for alpha value = .05; Effect size measured as Cohen's D absolute value.

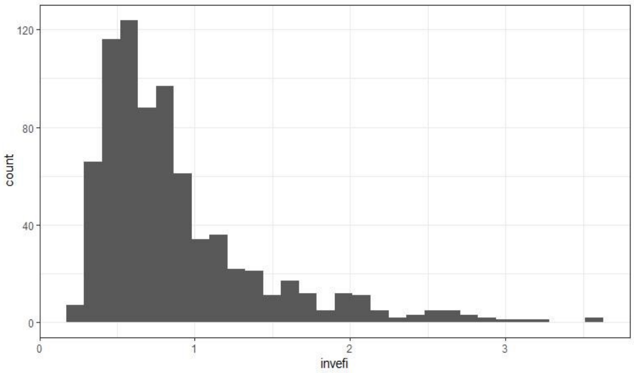
Shapiro-Will test for Normality on Congruency effect for each Group and Task instruction

Group	*Task*	*n*	*Statistic*	*p_{value}*
Italian Adults	Decreasing	47	0.923	.004
Italian Adults	Increasing	47	0.992	.985
Himba Adults 2021	Decreasing	33	0.974	.606
Himba Adults 2021	Increasing	26	0.968	.578
Himba Adults 2022	Decreasing	78	0.979	.222
Himba Adults 2022	Increasing	78	0.980	.245
Italian Preschooler	Decreasing	38	0.957	.148
Italian Preschooler	Increasing	38	0.938	.035

Table S6
In bold pvalue significant for alpha value = .05; Effect size measured as Cohen's D absolute value.

III. Supplementary Analysis

a. Inverse Efficiency Scores



For this kind of distribution (namely, typical time distribution), it is common to apply a **Gamma distribution** or an inverse Gaussian distribution. Due to better convergence, we use here Gamma distribution.

Model 1: Ideal model

We use an optimizer to allow the model to converge

```
Generalized linear mixed model fit by maximum likelihood (Laplace
 Approximation) [glmerMod]
 Family: Gamma ( inverse )
Formula: invefi ~ cong * Task * group + (1 + cong | participant_id)
  Data: df_invefi
Control: glmerControl(optimizer = "bobyqa")

   AIC   BIC  logLik deviance df.resid
 -630.7  -537.8  335.3  -670.7   750

Scaled residuals:
  Min   1Q Median   3Q   Max
-2.0028 -0.4614 0.0003 0.4671 3.7365

Random effects:
 Groups     Name      Variance Std.Dev. Corr
 participant_id (Intercept)   0.082652 0.28749
        congincongruent 0.005458 0.07388 -0.25
 Residual            0.042373 0.20585
Number of obs: 770, groups: participant_id, 222

Fixed effects:
                             Estimate Std. Error t value Pr(>|z|)
(Intercept)                       1.368371  0.091338 14.981 < 2e-16 ***
congincongruent                    -0.106894  0.044478 -2.403 0.01625 *
Taskincreasing                     -0.053947  0.132833 -0.406 0.68465
groupHimba Adults 2022                 0.129978  0.109628 1.186 0.23577
groupItalian Adults                   0.938937  0.118548 7.920 2.37e-15 ***
groupItalian Preschoolers               -0.419875  0.150194 -2.796 0.00518 **
congincongruent:Taskincreasing             0.110632  0.059919 1.846 0.06484 .
congincongruent:groupHimba Adults 2022          0.057150  0.051823 1.103 0.27011
congincongruent:groupItalian Adults           -0.006688  0.080634 -0.083 0.93390
congincongruent:groupItalian Preschoolers         0.051370  0.047394 1.084 0.27841
Taskincreasing:groupHimba Adults 2022           0.034927  0.136732 0.255 0.79838
Taskincreasing:groupItalian Adults           0.150734  0.151162 0.463 0.64353
Taskincreasing:groupItalian Preschoolers         0.062340  0.134710 0.997 0.31868
congincongruent:Taskincreasing:groupHimba Adults 2022  -0.046555  0.075311 -0.618 0.53647
congincongruent:Taskincreasing:groupItalian Adults    0.019416  0.117496 0.165 0.86875
congincongruent:Taskincreasing:groupItalian Preschoolers -0.049079  0.067760 -0.724 0.46888

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2: without the triple interaction

To reduce the model complexity, another possibility is to remove the triple interactions and focus solely on the interactions for which we have prior hypotheses, namely:

- Cong & Task (effect of congruence stronger for decreasing, as we saw this effect in the pre-testing with Himbas 2021)
- Cong & group (effect of congruence stronger for the groups not exposed to formal education)

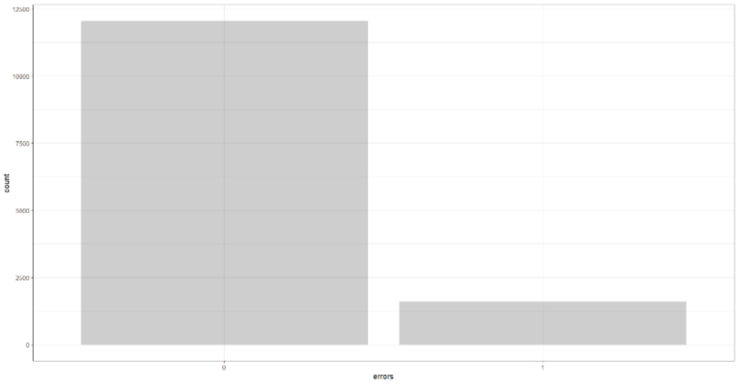
```
Generalized linear mixed model fit by maximum likelihood (Laplace
 Approximation) [glmerMod]
 Family: Gamma ( inverse )
Formula: invefi ~ cong * Task + cong * group + (1 + cong | participant_id)
  Data: df_invefi
Control: glmerControl(optimizer = "bobyqa")

   AIC   BIC  logLik deviance df.resid
 -634.5  -569.5  331.3  -662.5   756

Scaled residuals:
  Min   1Q Median   3Q   Max
-1.9616 -0.4584 -0.0017 0.4517 3.6474

Random effects:
 Groups     Name      Variance Std.Dev. Corr
 participant_id (Intercept)   0.082496 0.28722
        congincongruent 0.005382 0.07336 -0.24
 Residual            0.042773 0.20682
Number of obs: 770, groups: participant_id, 222

Fixed effects:
                Estimate Std. Error t value Pr(>|z|)
(Intercept)           1.342368  0.068711 19.536 < 2e-16 ***
congincongruent        -0.091299  0.037771 -2.417 0.01564 *
Taskincreasing         0.001902  0.017253  0.110 0.91223
groupVT_2022          0.144690  0.089814  1.611 0.10718
groupVT_adul          1.010536  0.095934 10.534 < 2e-16 ***
groupVT_pres          -0.391846  0.136728 -2.866 0.00416 **
congincongruent:Taskincreasing 0.073276  0.023388  3.133 0.00173 **
congincongruent:groupVT_2022  0.038471  0.040031  0.961 0.33654
congincongruent:groupVT_adul  0.002230  0.061377  0.036 0.97102
congincongruent:groupVT_pres  0.031344  0.038330  0.818 0.41351
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Then we apply model comparison:

```
Data: df_invefi
Models:
infevi_model2: invefi ~ cong * Task + cong * group + (1 + cong | participant_id)
infevi_model1: invefi ~ cong * Task * group + (1 + cong | participant_id)

        npar   AIC   BIC   logLik  deviance  Chisq  Df  Pr(>Chisq)
infevi_model2  14   -634.52 -569.47  331.26  -662.52
infevi_model1  20   -630.68 -537.76  335.34  -670.68   8.1677  6   0.2261
```

As the model are not significantly different, we chose to report the model with the triple interaction *Congruency Condition*Task instruction*Group* as it is the one more in line with our hypothesis and it is also coherent with the model used to analyze Reaction Times.

b. Errors



The *Error* variable clearly follows a **binomial** distribution. The Italian adults have made **no** errors. Thus, we exclude them from the following analysis.

Model 1: ideal model, with random slopes

The ideal model does not converge, despite the use of the optimizer

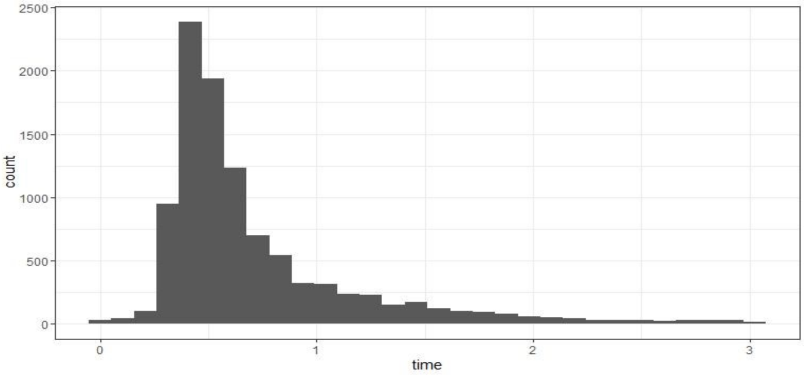
```
Generalized linear mixed model fit by maximum likelihood (Laplace
 Approximation) [glmerMod]
 Family: binomial (logit)
Formula: errors ~ cong * Task * group + (1 + cong | participant_id) + (1 | type_stim)
  Data: df_errors
```

Model 2: without random slopes

We remove the random slopes for congruency:

```
Generalized linear mixed model fit by maximum likelihood (Laplace
 Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: errors ~ cong * Task * group + (1 | participant_id) + (1 | type_stim)
  Data: df_errors
Control: glmerControl(optimizer = "bobyqa")

   AIC    BIC  logLik deviance df.resid
 4211.5  4310.5 -2091.7  4183.5    8678

Scaled residuals:
  Min    1Q Median    3Q   Max
-2.6127 -0.2265 -0.0977 -0.0320 29.8358

Random effects:
 Groups      Name     Variance Std.Dev.
 participant_id (Intercept) 3.181  1.784
 type_stim    (Intercept) 4.712  2.171
Number of obs: 8692, groups: participant_id, 175; type_stim, 6

Fixed effects:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)                -4.3428   1.2995 -3.342 0.000832 ***
congincongruent               0.5868   0.2786  2.107 0.035154 *
Taskincreasing               0.7152   1.8382  0.389 0.697212
groupVT_2022                 0.9003   0.4509  1.997 0.045837 *
groupVT_pres                 1.3194   0.5060  2.607 0.009123 **
congincongruent:Taskincreasing        -0.4151   0.3965 -1.047 0.295133
congincongruent:groupVT_2022         -0.3562   0.3174 -1.122 0.261751
congincongruent:groupVT_pres         -0.3868   0.3468 -1.115 0.264704
Taskincreasing:groupVT_2022         -0.9418   0.5832 -1.615 0.106337
Taskincreasing:groupVT_pres         -1.5564   0.6027 -2.583 0.009807 **
congincongruent:Taskincreasing:groupVT_2022  0.2666   0.4543  0.587 0.557235
congincongruent:Taskincreasing:groupVT_pres  0.3475   0.5010  0.694 0.487950

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c. Reaction Times



Given the data distribution (similar to IES), we apply a **Gamma distribution** to the model.

Model 1: Ideal model with triple interaction

We use an optimizer to allow the model to converge

```
Generalized linear mixed model fit by maximum likelihood (Laplace
 Approximation) [glmerMod]
 Family: Gamma ( inverse )
Formula: time ~ cong * Task * group + (1 + cong | participant_id) + (1 |
  type_stim)
  Data: df_time
Control: glmerControl(optimizer = "bobyqa")

   AIC   BIC  logLik deviance df.resid
 -1057.1  -905.5  549.6 -1099.1  10088

Scaled residuals:
  Min   1Q Median   3Q   Max
-2.4371 -0.5120 -0.1865 0.2247 11.2087

Random effects:
 Groups     Name      Variance Std.Dev. Corr
 participant_id (Intercept)   0.065914 0.25674
        congincongruent 0.010409 0.10202 -0.29
 type_stim   (Intercept)   0.001998 0.04469
 Residual            0.164870 0.40604
Number of obs: 10109, groups: participant_id, 222; type_stim, 6

Fixed effects:
                       Estimate Std. Error t value Pr(>|z|)
(Intercept)                 1.409320  0.104100 13.538 < 2e-16 ***
congincongruent               -0.061529  0.039078 -1.575 0.115367
Taskincreasing                0.133730  0.151809  0.881 0.378365
groupVT_2022                 0.346018  0.101984  3.393 0.000692 ***
groupVT_adul                 0.963322  0.110637  8.707 < 2e-16 ***
groupVT_pres                -0.461603  0.124675 -3.702 0.000214 ***
congincongruent:Taskincreasing        0.073965  0.058531  1.264 0.206335
congincongruent:groupVT_2022         0.031121  0.048027  0.648 0.517001
congincongruent:groupVT_adul        -0.045888  0.058248 -0.788 0.430810
congincongruent:groupVT_pres         0.009675  0.049430  0.196 0.844817
Taskincreasing:groupVT_2022         -0.160451  0.126923 -1.264 0.206174
Taskincreasing:groupVT_adul         -0.034176  0.131361 -0.260 0.794736
Taskincreasing:groupVT_pres         -0.120546  0.125913 -0.957 0.338378
congincongruent:Taskincreasing:groupVT_2022 0.014659  0.068400  0.214 0.830299
congincongruent:Taskincreasing:groupVT_adul 0.046848  0.082940  0.565 0.572183
congincongruent:Taskincreasing:groupVT_pres -0.013234  0.064202 -0.206 0.836693

---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model 2: without the triple interaction

To reduce the model complexity, another possibility is to remove the triple interactions and focus solely on the interactions for which we have prior hypotheses, namely:

- *Cong* & *Task* (effect of congruence stronger for decreasing, as we saw this effect in the pre-testing with Himbas 2021)
- *Cong* & *group* (effect of congruence stronger for the groups not exposed to formal education)

```
Generalized linear mixed model fit by maximum likelihood (Laplace
 Approximation) [glmerMod]
 Family: Gamma ( inverse )
Formula: time ~ cong * Task + cong * group + (1 + cong | participant_id) +
  (1 | type_stim)
  Data: df_time
Control: glmerControl(optimizer = "bobyqa")

   AIC    BIC  logLik deviance df.resid
 -1049.5  -941.1  539.7 -1079.5  10094

Scaled residuals:
  Min    1Q Median    3Q    Max
-2.4334 -0.5131 -0.1843 0.2232 11.3276

Random effects:
 Groups     Name       Variance Std.Dev. Corr
 participant_id (Intercept)   0.066430 0.25774
       congincongruent 0.010431 0.10213 -0.29
 type_stim   (Intercept)   0.001988 0.04459
 Residual            0.165397 0.40669
Number of obs: 10109, groups: participant_id, 222; type_stim, 6

Fixed effects:
                Estimate Std. Error t value Pr(>|z|)
(Intercept)           1.468846  0.085047 17.271 < 2e-16 ***
congincongruent        -0.063149  0.031941 -1.977  0.0480 *
Taskincreasing         0.002807  0.074600  0.038  0.9700
groupVT_2022          0.271698  0.084480  3.216  0.0013 **
groupVT_adul          0.950917  0.093439 10.177 < 2e-16 ***
groupVT_pres          -0.516312  0.111190 -4.643 3.43e-06 ***
congincongruent:Taskincreasing 0.077513  0.018859  4.110 3.95e-05 ***
congincongruent:groupVT_2022  0.038827  0.037940  1.023  0.3061
congincongruent:groupVT_adul  -0.026304  0.045760 -0.575  0.5654
congincongruent:groupVT_pres  0.003396  0.041380  0.082  0.9346
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Then we apply model comparison:

```
Data: df_time
Models:    npar  AIC  BIC  logLik  deviance  Chisq  Df  Pr(>Chisq)
time_model2   15   -1049.5 -941.15  539.73  -1079.5
time_model1   21   -1057.1 -905.46  549.55  -1099.1  19.636  6  0.003214 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

    R2m   R2c
[1,] 0.5258673 0.6606472
    R2m   R2c
[1,] 0.5211745 0.6576967
```

It seems that the best model is the one with the triple interaction, however, this difference is not clear cut: while the AIC is lower for the model with a triple interaction, the BIC is lower for a model without triple interaction. The percentage of variance explained is relatively similar too.

Reference List

Adachi, I. (2014). Spontaneous Spatial Mapping of Learned Sequence in Chimpanzees: Evidence for a SNARC-Like Effect. *PLOS ONE*, *9*(3), e90373. https://doi.org/10.1371/JOURNAL.PONE.0090373

Anderson, U. S., Stoinski, T. S., Bloomsmith, M. A., & Maple, T. L. (2007). Relative numerousness judgment and summation in young, middle-aged, and older adult orangutans (Pongo pygmaeus abelii and Pongo pygmaeus pygmaeus). *Journal of Comparative Psychology*, *121*(1), 1–11. https://doi.org/10.1037/0735-7036.121.1.1

Ansari, D., & Dhital, B. (2006). Age-related changes in the activation of the intraparietal sulcus during nonsymbolic magnitude processing: An event-related functional magnetic resonance imaging study. *Journal of Cognitive Neuroscience*, *18*(11), 1820–1828. https://doi.org/10.1162/jocn.2006.18.11.1820

Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus-response compatibility in representational space. *Neuropsychologia*, *36*(8), 731–735. https://doi.org/10.1016/S0028-3932(98)00002-5

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86*(3), 201–221. https://doi.org/10.1016/S0010-0277(02)00178-6

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences*, *102*(39), 14116–14121. https://doi.org/10.1073/PNAS.0505512102

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/JSS.V067.I01

Bateson, P. P. G., & Jaeckel, J. B. (1976). Chicks' preferences for familiar and novel conspicuous objects after different periods of exposure. *Animal Behaviour*, *24*(2), 386–390. https://doi.org/10.1016/S0003-3472(76)80048-6

Benson-Amram, S., Gilfillan, G., & McComb, K. (2018). Numerical assessment in the wild: insights from social carnivores. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740). https://doi.org/10.1098/RSTB.2016.0508

Beran, M. J., Evans, T. A., & Harris, E. H. (2008). Perception of food amounts by chimpanzees based on the number, size, contour length and visibility of items. *Animal Behaviour*, *75*(5), 1793–1802. https://doi.org/10.1016/J.ANBEHAV.2007.10.035

Beran, M. J., Parrish, A. E., & Evans, T. A. (2015). *Numerical Cognition and Quantitative Abilities in Nonhuman Primates*. *1*, 91–119. https://doi.org/10.1016/B978-0-12-420133-0.00004-1

Berch, D. B., Foley, E. J., Hill, R. J., & Ryan, P. M. D. (1999). Extracting Parity and Magnitude from Arabic Numerals: Developmental Changes in Number Processing and Mental Representation. *Journal of Experimental Child Psychology*, *74*(4), 286–308. https://doi.org/10.1006/JECP.1999.2518

Bergen, B. K., & Chan Lau, T. T. (2012). Writing direction affects how people map space onto time. *Frontiers in Psychology*, *3*(APR), 109. https://doi.org/10.3389/FPSYG.2012.00109/BIBTEX

Blanca, M. J., Alarcón, R., Arnau, J., Bono, R., & Bendayan, R. (2017). Datos no normales: ¿es el ANOVA una opción válida? *Psicothema*, *29*(4), 552–557. https://doi.org/10.7334/psicothema2016.383

Bloom, P. (1994). Generativity within language and other cognitive domains. *Cognition*, *51*(2), 177–189. https://doi.org/10.1016/0010-0277(94)90014-0

Bolhuis, J. J., Johnson, M. H., & Horn, G. (1985). Effects of early experience on the development of filial preferences in the domestic chick. *Developmental Psychobiology*, *18*(4), 299–308. https://doi.org/10.1002/DEV.420180403

Bolhuis, Johan J. (1991). Mechanisms of avian imprinting: a review. *Biological Reviews*, *66*(4), 303–345. https://doi.org/10.1111/j.1469-185x.1991.tb01145.x

Bortot, M., Agrillo, C., Avarguès-Weber, A., Bisazza, A., Petrazzini, M. E. M., & Giurfa, M. (2019). Honeybees use absolute rather than relative numerosity in number discrimination. *Biology Letters*, *15*(6). https://doi.org/10.1098/RSBL.2019.0138

Bortot, M., Regolin, L., & Vallortigara, G. (2021). A sense of number in invertebrates. *Biochemical and Biophysical Research Communications*, *564*, 37–42. https://doi.org/10.1016/J.BBRC.2020.11.039

Bortot, M., Stancher, G., & Vallortigara, G. (2020). Transfer from Number to Size Reveals Abstract Coding of Magnitude in Honeybees. *IScience*, *23*(5), 101122. https://doi.org/10.1016/j.isci.2020.101122

Bortot, M., & Vallortigara, G. (2023). Transfer from continuous to discrete quantities in honeybees. *IScience*, *26*(10), 108035. https://doi.org/10.1016/j.isci.2023.108035

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240. https://doi.org/10.1016/S0010-0277(02)00005-7

Brannon, E. M., & Roitman, J. D. (2003). Nonverbal representations of time and number in animals and human infants. In *Functional and Neural Mechanisms of Interval Timing* (pp. 143–182). CRC Press. https://doi.org/10.1201/9780203009574.ch6

Brannon, E. M., & Terrace, H. S. (1998). Ordering of the numerosities 1 to 9 by monkeys. *Science (New York, N.Y.)*, *282*(5389), 746–749. https://doi.org/10.1126/SCIENCE.282.5389.746

Buijsman, S., & Tirado, C. (2019). Spatial-numerical associations: Shared symbolic and non-symbolic numerical representations. *Quarterly Journal of Experimental Psychology (2006)*, *72*(10), 2423–2436. https://doi.org/10.1177/1747021819844503

Bulf, H., de Hevia, M. D., & Macchi Cassia, V. (2016). Small on the left, large on the right: numbers orient visual attention onto space in preverbal infants. *Developmental Science*, *19*(3), 394–401. http://doi.wiley.com/10.1111/desc.12315

Call, J. (2000). Estimating and operating on discrete quantities in orangutans (Pongo pygmaeus). *Journal of Comparative Psychology*, *114*(2), 136–147. https://doi.org/10.1037/0735-7036.114.2.136

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*(5), 401–406. https://doi.org/10.1111/J.1467-9280.2006.01719.X/ASSET/IMAGES/LARGE/10.1111_J.1467-9280.2006.01719.X-FIG2.JPEG

Cantlon, J. F., Brannon, E. M., Carter, E. J., & Pelphrey, K. A. (2006). Functional Imaging of Numerical Processing in Adults and 4-y-Old Children. *PLoS Biology*, *4*(5), e125. https://doi.org/10.1371/journal.pbio.0040125

Carazo, P., Font, E., Forteza-Behrendt, E., & Desfilis, E. (2009). Quantity discrimination in Tenebrio molitor: Evidence of numerosity discrimination in an invertebrate? *Animal Cognition*, *12*(3), 463–470. https://doi.org/10.1007/S10071-008-0207-7/FIGURES/2

Casasanto, D., & Hall, J. (2010). Space for Thinking. *Language, Cognition and Space: The State of the Art and New Directions*, 453–478.

Catani, M., & Thiebaut de Schotten, M. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, *44*(8), 1105–1132. https://doi.org/10.1016/j.cortex.2008.05.004

Chantrey, D. F. (1974). Stimulus preexposure and discrimination learning by domestic chicks: Effect of varying interstimulus time. *Journal of Comparative and Physiological Psychology*, *87*(3), 517–525. https://doi.org/10.1037/h0036982

Chantrey, D. F. (1976). The behaviour of domestic chicks during exposure to two stimuli. *Animal Behaviour*, *24*(4), 780–785. https://doi.org/10.1016/S0003-3472(76)80008-5

Chiandetti, C., & Vallortigara, G. (2011). Chicks like consonant music. *Psychological Science*, *22*(10), 1270–1273. https://doi.org/10.1177/0956797611418244

Chittka, L., & Geiger, K. (1995). Can honey bees count landmarks? *Animal Behaviour*, *49*(1), 159–164. https://doi.org/10.1016/0003-3472(95)80163-4

Chomsky, N. (2006). Language and Mind. *Language and Mind*. https://doi.org/10.1017/CBO9780511791222

Christman, S. (1989). Perceptual characteristics in visual laterality research. *Brain and Cognition*, *11*(2), 238–257. https://doi.org/10.1016/0278-2626(89)90020-1

Church, Russell M and Meck, W. H. (1983). Acquisition and cross-modal transfer of classification rules for temporal intervals. *Quantitative Analysis of Behavior: Discriminative Processes*, *4*, 75--97.

Cipora, K., Loenneker, H., Soltanlou, M., Lipowska, K., Domahs, F., Goebel, S. M., Haman, M., & Nuerk, H.-C. (2019). *Syntactic influences on numerical processing in adults: Limited but detectable*. https://doi.org/10.31234/OSF.IO/EWTD4

Cipora, K., Schroeder, P. A., Soltanlou, M., & Nuerk, H.-C. (2018). *More Space, Better Mathematics: Is Space a Powerful Tool or a Cornerstone for Understanding Arithmetic?* 77–116. https://doi.org/10.1007/978-3-319-98767-5_4

Cipora, K., Soltanlou, M., Reips, U. D., & Nuerk, H. C. (2019). The SNARC and MARC effects measured online: Large-scale assessment methods in flexible cognitive effects. *Behavior Research Methods*, *51*(4), 1676–1692. https://doi.org/10.3758/S13428-019-01213-5/FIGURES/5

Cipora, K., van Dijck, J.-P., Georges, C., Masson, N., Goebel, S. M., Willmes, K., Pesenti, M., Schiltz, C., & Nuerk, H.-C. (2019). *A Minority pulls the sample mean: on the individual prevalence of robust group-level cognitive phenomena – the instance of the SNARC effect*. https://doi.org/10.31234/OSF.IO/BWYR3

Clearfield, M. W., & Mix, K. S. (1999). Number Versus Contour Length in Infants' Discrimination of Small Visual Sets. *Https://Doi.Org/10.1111/1467-9280.00177*, *10*(5), 408–411. https://doi.org/10.1111/1467-9280.00177

Cleland, A. A., Corsico, K., White, K., & Bull, R. (2020). Non-symbolic numerosities do not automatically activate spatial–numerical associations: Evidence from the SNARC effect. *Quarterly Journal of Experimental Psychology*, *73*(2), 295–308. https://doi.org/10.1177/1747021819875021/ASSET/IMAGES/LARGE/10.1177_1747021819875021-FIG6.JPEG

Cohen Kadosh, R., Lammertyn, J., & Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. In *Progress in Neurobiology* (Vol. 84, Issue 2, pp. 132–147). https://doi.org/10.1016/j.pneurobio.2007.11.001

Cooperrider, K., Marghetis, T., & Núñez, R. (2017). Where Does the Ordered Line Come From? Evidence From a Culture of Papua New Guinea. *Psychological Science*, *28*(5), 599–608. https://doi.org/10.1177/0956797617691548/ASSET/IMAGES/LARGE/10.1177_0956797617691548-

FIG3.JPEG

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review*, *8*(4), 698–707. https://doi.org/10.3758/BF03196206

Dacke, M., & Srinivasan, M. V. (2008). Evidence for counting in insects. *Animal Cognition*, *11*(4), 683–689. https://doi.org/10.1007/S10071-008-0159-Y

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. *Brain and Cognition*, *20*(1), 125–151. https://doi.org/10.1016/0278-2626(92)90065-T

Davidson, R. J. (2004). Well–being and affective style: neural substrates and biobehavioural correlates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1449), 1395–1411. https://doi.org/10.1098/rstb.2004.1510

Davis, H., & Albert, M. (1986). Numerical discrimination by rats using sequential auditory stimuli. *Animal Learning & Behavior*, *14*(1), 57–59. https://doi.org/10.3758/BF03200037

Davis, H., & Albert, M. (1987). Failure to transfer or train a numerical discrimination using sequential visual stimuli in rats. *Bulletin of the Psychonomic Society*, *25*(6), 472–474. https://doi.org/10.3758/BF03334744/METRICS

De Corte, B. J., Navarro, V. M., & Wasserman, E. A. (2017). Non-cortical magnitude coding of space and time by pigeons. *Current Biology*, *27*(23), R1264–R1265. https://doi.org/10.1016/J.CUB.2017.10.029

De Hevia, M. D., Girelli, L., Addabbo, M., & Cassia, V. M. (2014). Human Infants' Preference for Left-to-Right Oriented Increasing Numerical Sequences. *PLOS ONE*, *9*(5), e96412. https://doi.org/10.1371/JOURNAL.PONE.0096412

de Hevia, M. D., & Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition*, *110*(2), 198–207. https://doi.org/10.1016/J.COGNITION.2008.11.003

de Hevia, M. D., Veggiotti, L., Streri, A., & Bonn, C. D. (2017). At Birth, Humans Associate "Few" with Left and "Many" with Right. *Current Biology*, *27*(24), 3879-3884.e2. https://doi.org/10.1016/j.cub.2017.11.024

De Tommaso, M., Kaplan, G., Chiandetti, C., & Vallortigara, G. (2019). Naïve 3-day-old domestic chicks (Gallus gallus) are attracted to discrete acoustic patterns characterizing natural vocalizations. *Journal of Comparative Psychology (Washington, D.C. : 1983)*, *133*(1), 118–131. https://doi.org/10.1037/COM0000132

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition*, *44*(1–2), 1–42. https://doi.org/10.1016/0010-0277(92)90049-N

Dehaene, S. (2011). The Number Sense: How the Mind Creates Mathematics, Revised and Updated Edition. In *Oxford University Press,*.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology: General*, *122*(3), 371–396. https://doi.org/10.1037/0096-3445.122.3.371

Dehaene, S., & Cohen, L. (1991). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia*, *29*(11), 1045–1074. https://doi.org/10.1016/0028-3932(91)90076-K

Dehaene, S., & Cohen, L. (1997). Cerebral Pathways for Calculation: Double Dissociation between Rote Verbal and Quantitative Knowledge of Arithmetic. *Cortex*, *33*(2), 219–250. https://doi.org/10.1016/S0010-

9452(08)70002-9

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. In *Cognitive Neuropsychology* (Vol. 20, Issues 3–6, pp. 487–506). https://doi.org/10.1080/02643290244000239

Di Giorgio, E., Lunghi, M., Rugani, R., Regolin, L., Dalla Barba, B., Vallortigara, G., & Simion, F. (2019). A mental number line in human newborns. *Developmental Science*, *22*(6), 1–10. https://doi.org/10.1111/desc.12801

Di Giorgio, E., Lunghi, M., Simion, F., & Vallortigara, G. (2017). Visual cues of motion that trigger animacy perception at birth: the case of self-propulsion. *Developmental Science*, *20*(4), e12394. https://doi.org/10.1111/desc.12394

Ditz, H. M., & Nieder, A. (2016). Numerosity representations in crows obey the Weber–Fechner law. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1827). https://doi.org/10.1098/RSPB.2016.0083

Ditz, H. M., & Nieder, A. (2020). Format-dependent and format-independent representation of sequential and simultaneous numerosity in the crow endbrain. *Nature Communications*, *11*(1), 1–10. https://doi.org/10.1038/s41467-020-14519-2

Dormal, V., Andres, M., Dormal, G., & Pesenti, M. (2010). Mode-dependent and mode-independent representations of numerosity in the right intraparietal sulcus. *NeuroImage*, *52*(4), 1677–1686. https://doi.org/10.1016/j.neuroimage.2010.04.254

Drucker, C. B., & Brannon, E. M. (2014). Rhesus monkeys (Macaca mulatta) map number onto space. *Cognition*, *132*(1), 57–67. https://doi.org/10.1016/J.COGNITION.2014.03.011

Fantoni, C., Baldassi, G., Rigutti, S., Prpic, V., Murgia, M., & Agostini, T. (2019). Emotional Semantic Congruency based on stimulus driven comparative judgements. *Cognition*, *190*, 20–41. https://doi.org/10.1016/J.COGNITION.2019.04.014

Feigenson, L. (2005). A double-dissociation in infants' representations of object arrays. *Cognition*, *95*(3). https://doi.org/10.1016/J.COGNITION.2004.07.006

Feigenson, L. (2011). Predicting sights from sounds: 6-month-olds' intermodal numerical abilities. *Journal of Experimental Child Psychology*, *110*(3), 347–361. https://doi.org/10.1016/j.jecp.2011.04.004

Feigenson, L., Carey, S., & Hauser, M. (2002). The Representations Underlying Infants' Choice of More: Object Files Versus Analog Magnitudes. *Https://Doi.Org/10.1111/1467-9280.00427*, *13*(2), 150–156. https://doi.org/10.1111/1467-9280.00427

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' Discrimination of Number vs. Continuous Extent. *Cognitive Psychology*, *44*(1), 33–66. https://doi.org/10.1006/COGP.2001.0760

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*(7), 307–314. https://doi.org/10.1016/j.tics.2004.05.002

Felisatti, A., Laubrock, J., Shaki, S., & Fischer, M. H. (2020). *A biological foundation for spatial-numerical associations: the brain's asymmetric frequency tuning*. *1477*(1), 44–53. https://pubmed.ncbi.nlm.nih.gov/32645221/

Fischer, M. H., Mills, R. A., & Shaki, S. (2010). How to cook a SNARC: Number placement in text rapidly changes spatial–numerical associations. *Brain and Cognition*, *72*(3), 333–336. https://doi.org/10.1016/J.BANDC.2009.10.010

Fischer, M. H., & Rottmann, J. (2005). Do negative numbers have a place on the mental number line ? *Psychology*

Science, *47*(1), 22–32.

Fischer, M. H., Shaki, S., & Cruise, A. (2009). It takes just one word to quash a SNARC. *Experimental Psychology*, *56*(5), 361–366. https://doi.org/10.1027/1618-3169.56.5.361

Fontanari, L., Rugani, R., Regolin, L., & Vallortigara, G. (2011). Object individuation in 3-day-old chicks: use of property and spatiotemporal information. *Developmental Science*, *14*(5), 1235–1244. https://doi.org/10.1111/J.1467-7687.2011.01074.X

Fuhrman, O., & Boroditsky, L. (2010). Cross-Cultural Differences in Mental Representations of Time: Evidence From an Implicit Nonlinguistic Task. *Cognitive Science*, *34*(8), 1430–1451. https://doi.org/10.1111/J.1551-6709.2010.01105.X

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*(1–2), 43–74. https://doi.org/10.1016/0010-0277(92)90050-R

Galton, F. (1880). Visualised Numerals. *Nature 1880 21:533*, *21*(533), 252–256. https://doi.org/10.1038/021252a0

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003

Gerten, J., & Topolinski, S. (2021). SNARC compatibility triggers positive affect. *Cognition and Emotion*, *35*(2), 356–366. https://doi.org/10.1080/02699931.2020.1846018

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 32–44. https://doi.org/10.1037/0096-1523.32.1.32

Gibson, L. C., & Maurer, D. (2016). Development of SNARC and distance effects and their relation to mathematical and visuospatial abilities. *Journal of Experimental Child Psychology*, *150*, 301–313. https://doi.org/10.1016/J.JECP.2016.05.009

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The Development of Automaticity in Accessing Number Magnitude. *Journal of Experimental Child Psychology*, *76*(2), 104–122. https://doi.org/10.1006/JECP.2000.2564

Giurfa, M., Marcout, C., Hilpert, P., Thevenot, C., & Rugani, R. (2022). An insect brain organizes numbers on a left-to-right mental number line. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(44), 2203584119. https://doi.org/10.1073/PNAS.2203584119/-/DCSUPPLEMENTAL

Göbel, S. M., Shaki, S., & Fischer, M. H. (2011). The cultural number line: A review of cultural and linguistic influences on the development of number processing. In *Journal of Cross-Cultural Psychology* (Vol. 42, Issue 4, pp. 543–565). https://doi.org/10.1177/0022022111406251

Grier, J. B., Counter, S. A., & Shearer, W. M. (1967). Prenatal auditory imprinting in chickens. *Science*, *155*(3770), 1692–1693. https://doi.org/10.1126/science.155.3770.1692

Grinnell, J., Packer, C., & Pusey, A. E. (1995). Cooperation in male lions: kinship, reciprocity or mutualism? *Animal Behaviour*, *49*(1), 95–105. https://doi.org/10.1016/0003-3472(95)80157-X

Guida, A., Megreya, A. M., Lavielle-Guida, M., Noël, Y., Mathy, F., van Dijck, J. P., & Abrahamse, E. (2018). Spatialization in working memory is related to literacy and reading direction: Culture “literarily” directs our thoughts. *Cognition*, *175*, 96–100. https://doi.org/10.1016/J.COGNITION.2018.02.013

Halberda, J., & Feigenson, L. (2008). *Developmental Change in the Acuity of the "Number Sense": The Approximate Number System in 3-, 4-, 5-, and 6-Year-Olds and Adults*. https://doi.org/10.1037/a0012682

Ham, A. D., & Osorio, D. (2007). Colour preferences and colour vision in poultry chicks. *Proceedings of the Royal Society B: Biological Sciences*, *274*(1621), 1941–1948. https://doi.org/10.1098/RSPB.2007.0538

Hanus, D., & Call, J. (2007). Discrete Quantity Judgments in the Great Apes (Pan paniscus, Pan troglodytes, Gorilla gorilla, Pongo pygmaeus): The Effect of Presenting Whole Sets Versus Item-by-Item. *Journal of Comparative Psychology*, *121*(3), 241–249. https://doi.org/10.1037/0735-7036.121.3.241

Haun, D. B. M., Jordan, F. M., Vallortigara, G., & Clayton, N. S. (2010). Origins of spatial, temporal and numerical cognition: Insights from comparative psychology. *Trends in Cognitive Sciences*, *14*(12), 552–560. https://doi.org/10.1016/J.TICS.2010.09.006

Hauser, M. D., Tsao, F., Garcia, P., & Spelke, E. S. (2003). Evolutionary foundations of number: Spontaneous representation of numerical magnitudes by cotton-top tamarins. *Proceedings of the Royal Society B: Biological Sciences*, *270*(1523), 1441–1446. https://doi.org/10.1098/RSPB.2003.2414

Heinsohn, R. (1997). Group territoriality in two populations of African lions. *Animal Behaviour*, *53*(6), 1143–1147. https://doi.org/10.1006/ANBE.1996.0316

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition 1982 10:4*, *10*(4), 389–395. https://doi.org/10.3758/BF03202431

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hess, E. H. (1959). Imprinting. *Science*, *130*(3368), 133–141. https://doi.org/10.1126/SCIENCE.130.3368.133/ASSET/842E0CD8-F700-4F0C-82C1-3331FCF7201D/ASSETS/SCIENCE.130.3368.133.FP.PNG

Heubner, L., Cipora, K., Soltanlou, M., Schlenker, M. L., Lipowska, K., Göbel, S. M., Domahs, F., Haman, M., & Nuerk, H. C. (2018). A mental odd-even continuum account: Some numbers may be "more odd" than others and some numbers may be "more even" than others. *Frontiers in Psychology*, *9*(JUN), 364587. https://doi.org/10.3389/FPSYG.2018.01081/BIBTEX

Hines, T. M. (1990). An odd effect: Lengthened reaction times for judgments about odd digits. *Memory & Cognition*, *18*(1), 40–46. https://doi.org/10.3758/BF03202644/METRICS

Hochman, S., Havedanloo, R., Heysieattalab, S., & Soltanlou, M. (2024). *How Does Language Modulate the Association Between Number and Space? A Registered Report of a Cross-cultural Study of the SNARC Effect*. https://doi.org/10.31234/OSF.IO/SME3Z

Hoffmann, D., Hornung, C., Martin, R., & Schiltz, C. (2013). Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds. *Journal of Experimental Child Psychology*, *116*(4), 775–791. https://doi.org/10.1016/J.JECP.2013.07.013

Hohol, M., Willmes, K., Nęcka, E., Brożek, B., Nuerk, H. C., & Cipora, K. (2020). Professional mathematicians do not differ from others in the symbolic numerical distance and size effects. *Scientific Reports 2020 10:1*, *10*(1), 1–12. https://doi.org/10.1038/s41598-020-68202-z

Holmes, K. J., Alcat, C., & Lourenco, S. F. (2019). Is Emotional Magnitude Spatialized? A Further Investigation. *Cognitive Science*, *43*(4), e12727. https://doi.org/10.1111/COGS.12727

Holmes, K. J., & Lourenco, S. F. (2011). Common spatial organization of number and emotional expression: A mental magnitude line. *Brain and Cognition*, *77*(2), 315–323.

https://doi.org/10.1016/J.BANDC.2011.07.002

Howard, S. R., Avarguès-Weber, A., Garcia, J. E., Greentree, A. D., & Dyer, A. G. (2018). Numerical ordering of zero in honey bees. *Science*, *360*(6393), 1124–1126. https://doi.org/10.1126/SCIENCE.AAR4975/SUPPL_FILE/AAR4975_HOWARD_SM.PDF

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews Neuroscience 2005 6:6*, *6*(6), 435–448. https://doi.org/10.1038/nrn1684

Hung, Y. hui, Hung, D. L., Tzeng, O. J. L., & Wu, D. H. (2008). Flexible spatial mapping of different notations of numbers in Chinese readers. *Cognition*, *106*(3), 1441–1450. https://doi.org/10.1016/J.COGNITION.2007.04.017

Hyde, D. C., Boas, D. A., Blair, C., & Carey, S. (2010). Near-infrared spectroscopy shows right parietal specialization for number in pre-verbal infants. *NeuroImage*, *53*(2), 647–652. https://doi.org/10.1016/j.neuroimage.2010.06.030

Inc., T. M. (2022). *MATLAB version: 9.13.0 (R2022b)*. The MathWorks Inc. https://www.mathworks.com

Izard, V., Dehaene-Lambertz, G., & Dehaene, S. (2008). Distinct Cerebral Pathways for Object Identity and Number in Human Infants. *PLoS Biology*, *6*(2), e11. https://doi.org/10.1371/journal.pbio.0060011

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106

JASP Team. (2024). *JASP (Version 0.18.3)[Computer software]*. https://jasp-stats.org/

Jeffreys, H. (1961). *Theory of Probability. 3rd Edition, Clarendon Press, Oxford.* 470. https://global.oup.com/academic/product/theory-of-probability-9780198503682

Johnson, M. H., Bolhuis, J. J., & Horn, G. (1985). Interaction between acquired preferences and developing predispositions during imprinting. *Animal Behaviour*, *33*(3), 1000–1006. https://doi.org/10.1016/S0003-3472(85)80034-8

Jordan, K. E., & Baker, J. (2011). Multisensory information boosts numerical matching abilities in young children. *Developmental Science*, *14*(2), 205–213. https://doi.org/10.1111/J.1467-7687.2010.00966.X

Jordan, K. E., & Brannon, E. M. (2006). The multisensory representation of number in infancy. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(9), 3486–3489. https://doi.org/10.1073/PNAS.0508107103/SUPPL_FILE/08107MOVIE1.MOV

Jordan, K. E., Brannon, E. M., Logothetis, N. K., & Ghazanfar, A. A. (2005). Monkeys match the number of voices they hear to the number of faces they see. *Current Biology*, *15*(11), 1034–1038. https://doi.org/10.1016/j.cub.2005.04.056

Jordan, K. E., MacLean, E. L., & Brannon, E. M. (2008). Monkeys match and tally quantities across senses. *Cognition*, *108*(3), 617–625. https://doi.org/10.1016/J.COGNITION.2008.05.006

Kent, J. P. (1993). The Chick's Preference for Certain Features of the Maternal Cluck Vocalization in the Domestic Fowl (Gallus Gallus). *Behaviour*, *125*(3–4), 177–187. https://doi.org/10.1163/156853993X00227

Kersey, A. J., & Cantlon, J. F. (2017). Neural tuning to numerosity relates to perceptual tuning in 3– 6-year-old children. *Journal of Neuroscience*, *37*(3), 512–522. https://doi.org/10.1523/JNEUROSCI.0065-16.2016

Kitchen, D. M. (2004). Alpha male black howler monkey responses to loud calls: effect of numeric odds, male companion behaviour and reproductive investment. *Animal Behaviour*, *67*(1), 125–139.

https://doi.org/10.1016/J.ANBEHAV.2003.03.007

Knecht, S., Deppe, M., Dräger, B., Bobe, L., Lohmann, H., Ringelstein, E. B., & Henningsen, H. (2000). Language lateralization in healthy right-handers. *Brain*, *123*(1), 74–81. https://doi.org/10.1093/BRAIN/123.1.74

Kobayashi, T., Hiraki, K., & Hasegawa, T. (2005). Auditory-visual intermodal matching of small numerosities in 6-month-old infants. *Developmental Science*, *8*(5), 409–419. https://doi.org/10.1111/J.1467-7687.2005.00429.X

Kobylkov, D., Mayer, U., Zanon, M., & Vallortigara, G. (2022). Number neurons in the nidopallium of young domestic chicks. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(32), e2201039119. https://doi.org/10.1073/PNAS.2201039119/SUPPL_FILE/PNAS.2201039119.SM01.MP4

Kovach, J. K. (1971). Effectiveness of Different Colors in the Elicitation and Development of Approach Behavior in Chicks. *Behaviour*, *38*(1–2), 154–168. https://doi.org/10.1163/156853971X00069

Krajcsi, A., Kojouharova, P., & Lengyel, G. (2022). Processing Symbolic Numbers: The Example of Distance and Size Effects. *Language, Cognition, and Mind*, *11*, 379–394. https://doi.org/10.1007/978-3-030-66175-5_27

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2016). The source of the symbolic numerical distance and size effects. *Frontiers in Psychology*, *7*(NOV), 220477. https://doi.org/10.3389/FPSYG.2016.01795/BIBTEX

Krajcsi, A., Lengyel, G., & Kojouharova, P. (2018). Symbolic number comparison is not processed by the analog number system: Different symbolic and non-symbolic numerical distance and size effects. *Frontiers in Psychology*, *9*(FEB), 332345. https://doi.org/10.3389/FPSYG.2018.00124/BIBTEX

Kutter, E. F., Bostroem, J., Elger, C. E., Mormann, F., & Nieder, A. (2018). Single Neurons in the Human Brain Encode Numbers. *Neuron*, *100*(3), 753-761.e4. https://doi.org/10.1016/J.NEURON.2018.08.036

Lakoff, G., & Johnson, M. (1980). The Metaphorical Structure of the Human Conceptual System. *COGNITIVE SCIENCE*, *4*, 195–208.

Lemaire, B. S., Rugani, R., Regolin, L., & Vallortigara, G. (2020). Response of male and female domestic chicks to change in the number (quantity) of imprinting objects. *Learning and Behavior*, 1–13. https://doi.org/10.3758/s13420-020-00446-1

Lemaire, B. S., & Vallortigara, G. (2021). *Filial imprinting and social predispositions in chicks (Gallus gallus domesticus)*. 1–97. https://doi.org/10.15168/11572_320462

Lichtenstein-Vidne, L., Gabay, S., Cohen, N., & Henik, A. (2017). Lateralisation of emotions: evidence from pupil size measurement. *Cognition & Emotion*, *31*(4), 699–711. https://doi.org/10.1080/02699931.2016.1164668

Lipton, J. S., & Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy*, *5*(3), 271–290. https://doi.org/10.1207/s15327078in0503_2

Loconsole, M., Pasculli, M. S., & Regolin, L. (2021). Space-luminance crossmodal correspondences in domestic chicks. *Vision Research*, *188*, 26–31. https://doi.org/10.1016/J.VISRES.2021.07.001

Lorenzi, E., Perrino, M., & Vallortigara, G. (2021). Numerosities and Other Magnitudes in the Brains: A Comparative View. *Frontiers in Psychology*, *12*, 641994. https://doi.org/10.3389/fpsyg.2021.641994

Lorenzi, E., & Vallortigara, G. (2021). Evolutionary and Neural Bases of the Sense of Animacy. *The Cambridge Handbook of Animal Cognition*, 295–321. https://doi.org/10.1017/9781108564113.017

Lourenco, S. F., & Longo, M. R. (2010). General magnitude representation in human infants. *Psychological*

Science : A Journal of the American Psychological Society / APS, *21*(6), 873–881. https://doi.org/10.1177/0956797610370158

Lyon, B. E. (2003). Ecological and social constraints on conspecific brood parasitism by nesting female American coots (Fulica americana). *Journal of Animal Ecology*, *72*(1), 47–60. https://doi.org/10.1046/J.1365-2656.2003.00674.X

Macchi Cassia, V., Picozzi, M., Girelli, L., & de Hevia, M. D. (2012). Increasing magnitude counts more: Asymmetrical processing of ordinality in 4-month-old infants. *Cognition*, *124*(2), 183–193. https://doi.org/10.1016/J.COGNITION.2012.05.004

Marino, L. (2017). Thinking chickens: a review of cognition, emotion, and behavior in the domestic chicken. *Animal Cognition*, *20*(2), 127. https://doi.org/10.1007/S10071-016-1064-4

Mccabe, B. J. (2013). Imprinting. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(4), 375–390. https://doi.org/10.1002/WCS.1231

McComb, K., Packer, C., & Pusey, A. (1994). Roaring and numerical assessment in contests between groups of female lions, Panthera leo. *Animal Behaviour*, *47*(2), 379–387. https://doi.org/10.1006/ANBE.1994.1052

McCrink, K., & de Hevia, M. D. (2018). From innate spatial biases to enculturated spatial cognition: The case of spatial associations in number and other sequences. *Frontiers in Psychology*, *9*(MAR), 332362. https://doi.org/10.3389/FPSYG.2018.00415/BIBTEX

McCrink, K., & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science*, *15*(11), 776–781. https://doi.org/10.1111/j.0956-7976.2004.00755.x

McCrink, K., & Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *Journal of Experimental Child Psychology*, *103*(4), 400–408. https://doi.org/10.1016/J.JECP.2009.01.013

Merritt, D. J., Casasanto, D., & Brannon, E. M. (2010). Do monkeys think in metaphors? Representations of space and time in monkeys and humans. *Cognition*, *117*(2), 191–202. https://doi.org/10.1016/J.COGNITION.2010.08.011

Miletto Petrazzini, M. E., & Brennan, C. H. (2020). Application of an abstract concept across magnitude dimensions by fish. *Scientific Reports 2020 10:1*, *10*(1), 1–7. https://doi.org/10.1038/s41598-020-74037-5

Mingolo, S., Prpic, V., Mariconda, A., Brugger, P., Drack, T., Bilotta, E., Agostini, T., & Murgia, M. (2024). It's SNARC o' clock: manipulating the salience of the context in a conceptual replication of Bächtold et al.'s (1998) clockface study. *Psychological Research*, *88*(3), 837–851. https://doi.org/10.1007/s00426-023-01893-x

Miura, M., & Matsushima, T. (2016). Biological motion facilitates filial imprinting. *Animal Behaviour*, *116*, 171–180. https://doi.org/10.1016/J.ANBEHAV.2016.03.025

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in infants: another look. *Developmental Psychology*, *33*(3), 423–428. https://doi.org/10.1037/0012-1649.33.3.423

Moore, D., Benenson, J., Reznick, J. S., Peterson, M., & Kagan, J. (1987). Effect of Auditory Numerical Information on Infants' Looking Behavior: Contradictory Evidence. *Developmental Psychology*, *23*(5), 665–670. https://doi.org/10.1037/0012-1649.23.5.665

Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education*, *1*(1), 21–31. https://doi.org/10.1016/J.TINE.2012.09.003

Nachshon, I. (1985). Directional preferences in perception of visual stimuli. *International Journal of Neuroscience*, *25*(3–4), 161–174. https://doi.org/10.3109/00207458508985369

Nachson, I., Argaman, E., & Luria, A. (1999). Effects of directional habits and handedness on aesthetic preference for left and right profiles. *Journal of Cross-Cultural Psychology*, *30*(1), 106–114. https://doi.org/10.1177/0022022199030001006

Nemeh, F., Humberstone, J., Yates, M. J., & Reeve, R. A. (2018). Non-symbolic magnitudes are represented spatially: Evidence from a non-symbolic SNARC task. *PLOS ONE*, *13*(8), e0203019. https://doi.org/10.1371/JOURNAL.PONE.0203019

Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience 2016 17:6*, *17*(6), 366–382. https://doi.org/10.1038/nrn.2016.40

Nieder, A. (2019). *A brain for numbers : the biology of the number instinct*. 376.

Nieder, A. (2020). The Adaptive Value of Numerical Competence. *Trends in Ecology & Evolution*, *35*(7), 605–617. https://doi.org/10.1016/J.TREE.2020.02.009

Nieder, A., & Dehaene, S. (2009). Representation of number in the brain. *Annual Review of Neuroscience*, *32*(Volume 32, 2009), 185–208. https://doi.org/10.1146/ANNUREV.NEURO.051508.135550/CITE/REFWORKS

Nuerk, H.-C., Wood, G., & Willmes, K. (2005). The Universal SNARC Effect. *Experimental Psychology*, *52*(3), 187–194. https://doi.org/10.1027/1618-3169.52.3.187

Nuerk, H. C., Iversen, W., & Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *57*(5), 835–863. https://doi.org/10.1080/02724980343000512

O'Shaughnessy, D. M., Cordero, T. C., Mollica, F., Boni, I., Jara-Ettinger, J., Gibson, E., & Piantadosi, S. T. (2023). Diverse mathematical knowledge among indigenous Amazonians. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(35), e2215999120. https://doi.org/10.1073/PNAS.2215999120/SUPPL_FILE/PNAS.2215999120.SAPP.PDF

Panteleeva, S., Reznikova, Z., & Vygonyailova, O. (2013). Quantity judgments in the context of risk/reward decision making in striped field mice: First "count," then hunt. *Frontiers in Psychology*, *4*(FEB), 1–8. https://doi.org/10.3389/fpsyg.2013.00053

Patro, K., Fischer, U., Nuerk, H.-C., & Cress, U. (2015). *How to rapidly construct a spatial-numerical representation in preliterate children (at least temporarily)*. https://doi.org/10.1111/desc.12296

Patro, K., & Haman, M. (2011). *The spatial-numerical congruity effect in preschoolers*. https://doi.org/10.1016/j.jecp.2011.09.006

Patro, K., & Nuerk, H. C. (2017). Limitations of Trans-Species Inferences: The Case of Spatial-Numerical Associations in Chicks and Humans. *Cognitive Science*, *41*(8), 2267–2274. https://doi.org/10.1111/COGS.12432

Patro, K., Nuerk, H. C., & Cress, U. (2016). Mental Number Line in the Preliterate Brain: The Role of Early Directional Experiences. *Child Development Perspectives*, *10*(3), 172–177. https://doi.org/10.1111/CDEP.12179

Patro, K., Nuerk, H. C., Cress, U., & Haman, M. (2014). How number-space relationships are assessed before formal schooling: A taxonomy proposal. *Frontiers in Psychology*, *5*(MAY), 55976. https://doi.org/10.3389/FPSYG.2014.00419/BIBTEX

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/S13428-018-01193-Y/FIGURES/3

Perdue, B. M., Talbot, C. F., Stone, A. M., & Beran, M. J. (2012). Putting the elephant back in the herd: Elephant relative quantity judgments match those of other species. *Animal Cognition*, *15*(5), 955–961. https://doi.org/10.1007/S10071-012-0521-Y/FIGURES/2

Piazza, M. (2011). Neurocognitive Start-Up Tools for Symbolic Number Representations. In *Space, Time and Number in the Brain* (Vol. 1, Issue c). Elsevier Inc. https://doi.org/10.1016/B978-0-12-385948-8.00017-7

Piazza, M., & Izard, V. (2009). How Humans Count: Numerosity and the Parietal Cortex. *The Neuroscientist*, *15*(3), 261–273. https://doi.org/10.1177/1073858409333073

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron*, *44*(3), 547–555. https://doi.org/10.1016/j.neuron.2004.10.014

Piazza, M., Mechelli, A., Price, C. J., & Butterworth, B. (2006). Exact and approximate judgements of visual and auditory numerosity: An fMRI study. *Brain Research*, *1106*(1), 177–188. https://doi.org/10.1016/J.BRAINRES.2006.05.104

Piazza, M., Pica, P., Izard, V., Spelke, E. S., & Dehaene, S. (2013). Education Enhances the Acuity of the Nonverbal Approximate Number System. *Psychological Science*, *24*(6), 1037–1043. https://doi.org/10.1177/0956797612464057

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A Magnitude Code Common to Numerosities and Number Symbols in Human Intraparietal Cortex. *Neuron*, *53*(2), 293–305. https://doi.org/10.1016/j.neuron.2006.11.022

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*(5695), 499–503. https://doi.org/10.1126/science.1102085

Pisa, P. E., & Agrillo, C. (2009). Quantity discrimination in felines: A preliminary investigation of the domestic cat (Felis silvestris catus). *Journal of Ethology*, *27*(2), 289–293. https://doi.org/10.1007/S10164-008-0121-0/TABLES/1

Pitt, B., & Casasanto, D. (2014). Experiential Origins of the Mental Number Line. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36).

Pitt, B., & Casasanto, D. (2020). The correlations in experience principle: How culture shapes concepts of time and number. *Journal of Experimental Psychology. General*, *149*(6). https://doi.org/10.1037/XGE0000696

Pitt, B., Ferrigno, S., Cantlon, J. F., Casasanto, D., Gibson, E., & Piantadosi, S. T. (2021). *Spatial concepts of number , size , and time in an indigenous culture*. *August*, 1–7.

Platt, M. L., & Spelke, E. S. (2009). What can developmental and comparative cognitive neuroscience tell us about the adult human brain? *Current Opinion in Neurobiology*, *19*(1), 1–5. https://doi.org/10.1016/J.CONB.2009.06.002

Prpic, V., Basamh, Y. A., Goodridge, C. M., Agostini, T., & Murgia, M. (2023). Contrasting symbolic and non-symbolic numerical representations in a joint classification task. *Psychonomic Bulletin and Review*, *30*(4), 1422–1430. https://doi.org/10.3758/S13423-023-02246-W/TABLES/2

Pulliam, H. R. (1973). On the advantages of flocking. *Journal of Theoretical Biology*, *38*(2), 419–422. https://doi.org/10.1016/0022-5193(73)90184-7

Quaranta, A., Siniscalchi, M., & Vallortigara, G. (2007). Asymmetric tail-wagging responses by dogs to different emotive stimuli. *Current Biology*, *17*(6), R199–R201. https://doi.org/10.1016/J.CUB.2007.02.008

R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Ratcliffe, V. F., Taylor, A. M., & Reby, D. (2016). Cross-Modal Correspondences in Non-human Mammal Communication. *Multisensory Research*, *29*(1–3), 49–91. https://doi.org/10.1163/22134808-00002509

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology*, *83*(2 PART 1), 274–278. https://doi.org/10.1037/h0028573

Rivera, S. M., Reiss, A. L., Eckert, M. A., & Menon, V. (2005). Developmental Changes in Mental Arithmetic: Evidence for Increased Functional Specialization in the Left Inferior Parietal Cortex. *Cerebral Cortex*, *15*(11), 1779–1790. https://doi.org/10.1093/CERCOR/BHI055

Roberts, G. (1996). Why individual vigilance declines as group size increases. *Animal Behaviour*, *51*(5), 1077–1086. https://doi.org/10.1006/ANBE.1996.0109

Rogers, L. J., Vallortigara, G., & Andrew, R. J. (2013). Divided brains: The biology and behaviour of brain asymmetries. In *Divided Brains: The Biology and Behaviour of Brain Asymmetries*. https://doi.org/10.1017/CBO9780511793899

Rosa-Salva, O., Hernik, M., Fabbroni, M., Lorenzi, E., & Vallortigara, G. (2023). Naïve chicks do not prefer objects with stable body orientation, though they may prefer behavioural variability. *Animal Cognition*, *26*(4), 1177–1189. https://doi.org/10.1007/S10071-023-01764-3/FIGURES/5

Rosa-Salva, O., Mayer, U., Versace, E., Hébert, M., Lemaire, B. S., & Vallortigara, G. (2021). Sensitive periods for social development: Interactions between predisposed and learned mechanisms. *Cognition*, *213*, 104552. https://doi.org/10.1016/J.COGNITION.2020.104552

Rosa-Salva, O., Regolin, L., & Vallortigara, G. (2010). Faces are special for newly hatched chicks: Evidence for inborn domain-specific mechanisms underlying spontaneous preferences for face-like stimuli. *Developmental Science*, *13*(4), 565–577. https://doi.org/10.1111/j.1467-7687.2009.00914.x

Rosa Salva, O., Mayer, U., & Vallortigara, G. (2015). Roots of a social brain: Developmental models of emerging animacy-detection mechanisms. *Neuroscience & Biobehavioral Reviews*, *50*, 150–168. https://doi.org/10.1016/J.NEUBIOREV.2014.12.015

Roth, L., Jordan, V., Schwarz, S., Willmes, K., Nuerk, H. C., van Dijck, J. P., & Cipora, K. (2024). Don’t SNARC me now! Intraindividual variability of cognitive phenomena – Insights from the Ironman paradigm. *Cognition*, *248*, 105781. https://doi.org/10.1016/J.COGNITION.2024.105781

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. http://www.rstudio.com/

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The Development of Internal Representations of Magnitude and Their Association with Arabic Numerals. *Journal of Experimental Child Psychology*, *81*(1), 74–92. https://doi.org/10.1006/JECP.2001.2645

Rugani, R. (2018). Towards numerical cognition’s origin: insights from day-old domestic chicks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740). https://doi.org/10.1098/RSTB.2016.0509

Rugani, R., Fontanari, L., Simoni, E., Regolin, L., & Vallortigara, G. (2009). Arithmetic in newborn chicks. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1666), 2451–2460. https://doi.org/10.1098/rspb.2009.0044

Rugani, R., Kelly, D. M., Szelest, I., Regolin, L., & Vallortigara, G. (2010). Is it only humans that count from left to right? *Biology Letters*, *6*(3), 290–292. https://doi.org/10.1098/RSBL.2009.0960

Rugani, R., McCrink, K., De Hevia, M. D., Vallortigara, G., & Regolin, L. (2016). Ratio abstraction over discrete magnitudes by newly hatched domestic chicks (Gallus gallus). *Scientific Reports 2016 6:1*, *6*(1), 1–8. https://doi.org/10.1038/srep30114

Rugani, R., Platt, M. L., Zhang, Y., & Brannon, E. M. (2024). Magnitude shifts spatial attention from left to right in rhesus monkeys as in the human mental number line. *IScience*, *27*(2), 108866. https://doi.org/10.1016/j.isci.2024.108866

Rugani, R., & Regolin, L. (2020). Hemispheric specialization in spatial versus ordinal processing in the day-old domestic chick (Gallus gallus). *Annals of the New York Academy of Sciences*, *1477*(1), 34–43. https://doi.org/10.1111/NYAS.14345

Rugani, R., Regolin, L., & Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *Journal of Experimental Psychology. Animal Behavior Processes*, *34*(3), 388–399. https://doi.org/10.1037/0097-7403.34.3.388

Rugani, R., Regolin, L., & Vallortigara, G. (2010). Imprinted numbers: newborn chicks' sensitivity to number vs. continuous extent of objects they have been reared with. *Developmental Science*, *13*(5), 790–797. https://doi.org/10.1111/j.1467-7687.2009.00936.x

Rugani, R., Regolin, L., & Vallortigara, G. (2011). Summation of large numerousness by newborn chicks. *Frontiers in Psychology*, *2*(SEP), 10302. https://doi.org/10.3389/FPSYG.2011.00179/BIBTEX

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science*, *347*(6221), 534–536. https://doi.org/10.1126/science.aaa1379

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2020). Numerical magnitude, rather than individual bias, explains spatial numerical association in newborn chicks. *ELife*, *9*, 1–12. https://doi.org/10.7554/ELIFE.54662

Rugani, R., Vallortigara, G., & Regolin, L. (2013). Numerical Abstraction in Young Domestic Chicks (Gallus gallus). *PLoS ONE*, *8*(6), e65262. https://doi.org/10.1371/journal.pone.0065262

Rugani, R., Vallortigara, G., & Regolin, L. (2014). From small to large: Numerical discrimination by young domestic chicks (Gallus gallus). *Journal of Comparative Psychology*, *128*(2), 163–171. https://doi.org/10.1037/a0034513

Rugani, R., Vallortigara, G., & Regolin, L. (2016). Mapping number to space in the two hemispheres of the avian brain. *Neurobiology of Learning and Memory*, *133*, 13–18. https://doi.org/10.1016/J.NLM.2016.05.010

Rugani, R., Vallortigara, G., Vallini, B., & Regolin, L. (2011). Asymmetrical number-space mapping in the avian brain. *Neurobiology of Learning and Memory*, *95*(3), 231–238. https://doi.org/10.1016/J.NLM.2010.11.012

Salzen, E. A., Lily, R. E., & McKeown, J. R. (1971). Colour preference and imprinting in domestic chicks. *Animal Behaviour*, *19*(3), 542–547. https://doi.org/10.1016/S0003-3472(71)80109-4

Santolin, C., Rosa-Salva, O., Lemaire, B. S., Regolin, L., & Vallortigara, G. (2020). Statistical learning in domestic chicks is modulated by strain and sex. *Scientific Reports*, *10*(1), 1–8. https://doi.org/10.1038/s41598-020-72090-8

Sasanguie, D., De Smedt, B., & Reynvoet, B. (2017). Evidence for distinct magnitude systems for symbolic and non-symbolic number. *Psychological Research*, *81*(1), 231–242. https://doi.org/10.1007/S00426-015-0734-1/TABLES/2

Schaefer, H. H., & Hess, E. H. (1959). Color Preferences in Imprinting Objects1. *Zeitschrift Für Tierpsychologie*, *16*(2), 161–172. https://doi.org/10.1111/J.1439-0310.1959.TB02051.X

Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Developmental Science*, *20*(3), e12372. https://doi.org/10.1111/DESC.12372

Shaki, S., & Fischer, M. H. (2008). Reading space into numbers: a cross-linguistic comparison of the SNARC effect. *Cognition*, *108*(2), 590–599. https://doi.org/10.1016/J.COGNITION.2008.04.001

Shaki, S., & Fischer, M. H. (2018). *Deconstructing spatial-numerical associations*. https://doi.org/10.1016/j.cognition.2018.02.022

Shaki, S., Fischer, M. H., & Göbel, S. M. (2012). Direction counts: A comparative study of spatially directional counting biases in cultures with different reading directions. *Journal of Experimental Child Psychology*, *112*(2), 275–281. https://doi.org/10.1016/J.JECP.2011.12.005

Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review 2009 16:2*, *16*(2), 328–331. https://doi.org/10.3758/PBR.16.2.328

Shaki, S., Petrusic, W. M., & Leth-Steensen, C. (2012). *SNARC Effects With Numerical and Non-Numerical Symbolic Comparative Judgments: Instructional And Cultural Dependencies*. https://doi.org/10.1037/a0026729

Shumaker, R. W., Beck, B. B., Palkovich, A. M., Guagnano, G. A., & Morowitz, H. (2001). Spontaneous use of magnitude discrimination and ordination by the orangutan (Pongo pygmaeus). *Journal of Comparative Psychology*, *115*(4), 385–391. https://doi.org/10.1037/0735-7036.115.4.385

Skorupski, P., MaBouDi, H. Di, Galpayage Dona, H. S., & Chittka, L. (2018). Counting insects. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1740), 20160513. https://doi.org/10.1098/RSTB.2016.0513

Smyth, R. E., & Ansari, D. (2020). Do infants have a sense of numerosity? A p-curve analysis of infant numerosity discrimination studies. *Developmental Science*, *23*(2), e12897. https://doi.org/10.1111/DESC.12897

Spelke, E. S. (2000). Core Knowledge of Objects. *American Psychologist*, *November*, 1233–1243.

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics 2011 73:4*, *73*(4), 971–995. https://doi.org/10.3758/S13414-010-0073-7

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science (New York, N.Y.)*, *222*(4620), 179–181. https://doi.org/10.1126/SCIENCE.6623069

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*(2), 97–127. https://doi.org/10.1016/0010-0277(90)90001-Z

Sugita, Y. (2008). Face perception in monkeys reared with no exposure to faces. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(1), 394–398. https://doi.org/10.1073/PNAS.0706079105

Thomas, R. K., & Chase, L. (1980). Relative numerousness judgments by squirrel monkeys. *Bulletin of the Psychonomic Society*, *16*(2), 79–82. https://doi.org/10.3758/BF03334444/METRICS

Togoli, I., & Arrighi, R. (2021). Evidence for an A-Modal Number Sense: Numerosity Adaptation Generalizes Across Visual, Auditory, and Tactile Stimuli. *Frontiers in Human Neuroscience*, *15*, 448. https://doi.org/10.3389/FNHUM.2021.713565/BIBTEX

Tokita, M., & Ishiguchi, A. (2016). Precision and Bias in Approximate Numerical Judgment in Auditory, Tactile, and Cross-modal Presentation. *Perception*, *45*(1–2), 56–70. https://doi.org/10.1177/0301006615596888

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *TRENDS in Sport Sciences*, *1*(21), 19–25.

Toomarian, E. Y., & Hubbard, E. M. (2018). On the genesis of spatial-numerical associations: Evolutionary and cultural factors co-construct the mental number line. *Neuroscience & Biobehavioral Reviews*, *90*, 184–199. https://doi.org/10.1016/J.NEUBIOREV.2018.04.010

Townsend, J. T., Ashby, F. G., & others. (1983). *Stochastic modeling of elementary psychological processes*. CUP Archive.

Tversky, B., Kugelmass, S., & Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cognitive Psychology*, *23*(4), 515–557. https://doi.org/10.1016/0010-0285(91)90005-9

Vallortigara, G., & Andrew, R. J. (1991). Lateralization of response by chicks to change in a model partner. *Animal Behaviour*, *41*(2), 187–194. https://doi.org/10.1016/S0003-3472(05)80470-1

Vallortigara, Giorgio. (1992). Right hemisphere advantage for social recognition in the chick. *Neuropsychologia*, *30*(9), 761–768. https://doi.org/10.1016/0028-3932(92)90080-6

Vallortigara, Giorgio. (2012). Aristotle and the Chicken: Animacy and the Origins of Beliefs. In *The Theory of Evolution and Its Impact* (pp. 189–199). Springer, Milano. https://doi.org/10.1007/978-88-470-1974-4_12

Vallortigara, Giorgio. (2018). Comparative cognition of number and space: the case of geometry and of the mental number line. In *Philosophical Transactions of the Royal Society B: Biological Sciences* (Vol. 373, Issue 1740, p. 20170120). Royal Society Publishing. https://doi.org/10.1098/rstb.2017.0120

Vallortigara, Giorgio, Chiandetti, C., Rugani, R., Sovrano, V. A., & Regolin, L. (2010). Animal cognition. In *Wiley Interdisciplinary Reviews: Cognitive Science* (Vol. 1, Issue 6, pp. 882–893). John Wiley & Sons, Ltd. https://doi.org/10.1002/wcs.75

Vallortigara, Giorgio, & Versace, E. (2018). Filial Imprinting. In *Encyclopedia of Animal Cognition and Behavior* (pp. 1–4). Springer, Cham. https://doi.org/10.1007/978-3-319-47829-6_1989-1

Vallortigara, Giorgio, & Zanforlin, M. (1988). Open-field behavior of young chicks (Gallus gallus): Antipredatory responses, social reinstatement motivation, and gender effects. *Animal Learning & Behavior*, *16*(3), 359–362. https://doi.org/10.3758/BF03209088

Van Dijck, J.-P., Abrahamse, E., & Fias, W. (2020). Do preliterate children spontaneously employ spatial coding for serial order in working memory? *Ann. N.Y. Acad. Sci*. https://doi.org/10.1111/nyas.14430

van Galen, M. S., & Reitsma, P. (2008). Developing access to number magnitude: A study of the SNARC effect in 7- to 9-year-olds. *Journal of Experimental Child Psychology*, *101*(2), 99–113. https://doi.org/10.1016/J.JECP.2008.05.001

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Versace, E., Fracasso, I., Baldan, G., Dalle Zotte, A., & Vallortigara, G. (2017). Newborn chicks show inherited variability in early social predispositions for hen-like stimuli. *Scientific Reports 2017 7:1*, *7*(1), 1–7.

https://doi.org/10.1038/srep40296

Versace, E., Spierings, M. J., Caffini, M., ten Cate, C., & Vallortigara, G. (2017). Spontaneous generalization of abstract multimodal patterns in young domestic chicks. *Animal Cognition*, *20*(3), 521–529. https://doi.org/10.1007/s10071-017-1079-5

Versace, E., & Vallortigara, G. (2015). Origins of knowledge: Insights from precocial species. *Frontiers in Behavioral Neuroscience*, *9*(DEC), 147219. https://doi.org/10.3389/FNBEH.2015.00338/BIBTEX

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/J.TICS.2003.09.002

Wang, J., & Feigenson, L. (2021). Dynamic changes in numerical acuity in 4-month-old infants. *Infancy : The Official Journal of the International Society on Infant Studies*, *26*(1), 47–62. https://doi.org/10.1111/INFA.12373

Ward, C., & Smuts, B. B. (2007). Quantity-based judgments in the domestic dog (Canis lupus familiaris). *Animal Cognition*, *10*(1), 71–80. https://doi.org/10.1007/S10071-006-0042-7/TABLES/2

Watts, D. P., & Mitani, J. C. (2002). Hunting behavior of chimpanzees at Ngogo, Kibale National Park, Uganda. *International Journal of Primatology*, *23*(1), 1–28. https://doi.org/10.1023/A:1013270606320/METRICS

Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., & Dehaene, S. (2006). An open trial assessment of “The Number Race”, an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions : BBF*, *2*. https://doi.org/10.1186/1744-9081-2-20

Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, *61*(6), 1203–1216. https://doi.org/10.1006/ANBE.2000.1706

Wittlinger, M., Wehner, R., & Wolf, H. (2006). The ant odometer: stepping on stilts and stumps. *Science (New York, N.Y.)*, *312*(5782), 1965–1967. https://doi.org/10.1126/SCIENCE.1126912

Wood, G., Willmes, K., Nuerk, H.-C., & Fischer, M. H. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychology Science Quarterly*, *50*(4), 489–525.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature 1992 358:6389*, *358*(6389), 749–750. https://doi.org/10.1038/358749a0

Xu, F., & Arriaga, R. I. (2007). Number discrimination in 10-month-old infants. *British Journal of Developmental Psychology*, *25*(1), 103–108. https://doi.org/10.1348/026151005X90704

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1). https://doi.org/10.1016/S0010-0277(99)00066-9

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. https://doi.org/10.1002/CNE.920180503

Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: The SNARC effect and the REVERSE SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *Journal of Cognition and Culture*, *5*(1–2), 165–190. https://doi.org/10.1163/1568537054068660

Zohar-Shai, B., Tzelgov, J., Karni, A., & Rubinsten, O. (2017). It does exist! a left-to-right spatial-numerical association of response codes (SNARC) effect among native hebrew speakers. *Journal of Experimental Psychology: Human Perception and Performance*, *43*(4), 719–728. https://doi.org/10.1037/xhp0000336