UNIVERSITY OF PADOVA

Department of General Psychology

Master Degree in Cognitive Neuroscience and Clinical Neuropsychology

Final dissertation

On the functional independence of the Approximate Number System and Visual Working Memory in infants

Supervisor

Professor Judit Gervain

Co-supervisor

PhD Caterina Marino

Candidate: Vittoria Volpi

Student ID number: 2069469

Academic Year 2023/2024

Abstract

Understanding the developmental origins of cognitive systems provides critical insight into the broader landscape of human cognitive abilities. The relationship between Visual Working Memory (VWM) and the Approximate Number System (ANS) in early infancy has remained largely underexplored, despite adult studies suggesting these systems operate as functionally independent stages of information processing. This study investigates VWM span and the acuity of ANS in infants aged 9-12 months, aiming to determine if individual VWM span correlates with ANS acuity in this age group. A cohort of 14 infants was assessed through a series of visual change detection tasks with varying set sizes (1, 2, 3, and 4 elements) and numerical ratios (1:2, 2:3 and 3:4). Preference scores were derived from looking times, and correlations between VWM and ANS performance were analyzed. In the VWM tasks, infants exhibited a significant preference for changes in set sizes of 1 and 2 elements, indicating a VWM span of up to 2 elements. No significant preference was observed for larger set sizes. No significant preference was measured for numerical change, at any given ratio The lack of correlation between VWM and ANS supports the hypothesis of their functional independence, suggesting distinct developmental origins for these systems. Nevertheless, Bayesian correlation analyses provided insufficient evidence for either the presence or the absence of a significant correlation, highlighting the need for further research to substantiate these findings. These findings advance our understanding of early cognitive development and clarify the nature of number sense and its integration into complex cognitive architecture. However, the small sample size limits the generalizability of these findings, underscoring the need for further investigation to confirm these results.

Table of contents

References

Supplementary materials

Chapter 1 – Introduction

Early mathematical abilities play a fundamental role in education and predict future academic performance (Vanbecelaere et al., 2021). Research suggests that early numerical skills, such as magnitude estimation, are crucial for the subsequent development of robust mathematical and arithmetic abilities. These early, and phylogenetically shared, numerical skills are supported by the Approximate Number System (ANS), that serve as a foundational module upon which more complex numerical concepts can be built through formal education (Piazza, 2010). However, other cognitive systems and mechanisms also contribute to shaping mathematical abilities. Experimental evidence shows that formal math test performance is predicted by both ANS acuity (Mazzocco et al., 2011b) and Visual Working Memory (VWM) measures (Alloway & Passolunghi, 2011). Surprisingly, although both VWM and ANS contribute to the acquisition of mathematical knowledge, they do not appear to be associated with each other in adults. It is therefore relevant to investigate whether these systems are functionally independent during early phases of life. By testing this hypothesis in infants, the present research will shed light on the developmental trajectories of VWM and ANS: are they two distinct mechanisms even in infancy? Or do they originate as one and then, through development, diverge?

1.1 No correlation between Visual Working Memory (VWM) and the Approximate Number System (ANS) in adults

VWM is the ability to retain a limited amount of visual information and keep it available for a brief interval of time, e.g. for cognitive operations (Rouder et al., 2011). VWM has a limited capacity by nature, and, although there is considerable inter-individual variability, its capacity gravitates around 3 elements in human adults (Vogel et al., 2005).

In adults, VWM span correlates with many aspects of cognitive efficiency and numerous higher-level cognitive functions, including executive functions, language skills, mathematical abilities, learning, fluid intelligence, and even practical skills such as driving (Dell'Acqua et al., 2024).

By employing the event-related potential (ERP) technique during VWM tasks, a sustained negative wave over the posterior electrode sites in the contralateral hemisphere has been described, and referred to as the contralateral delay activity. The individual contralateral delay activity acts as a correlate of VWM capacity and is a reliable predictor of individual differences of many different cognitive abilities (Fukuda et al., 2015).

Interestingly, two recent studies found one ability that VWM does not correlate with: the ANS. The ANS is a cognitive ability that allows a non-precise estimation of the magnitude/numerosity of elements in a set, without relying on language or mathematical symbols. We rely on the ANS, for instance, to decide which queue has fewer people in it when lining up at the supermarket check-out.

In a study by Piazza et al. (2011) a dual-task was administered to adult participants, that had to perform simultaneously both a numerical magnitude comparison task (which recruits ANS) and a VWM task. The results indicated that numerical acuity in the magnitude comparison task was not affected by the load imposed by the VWM task, suggesting that these two systems, in adults, are not associated.

Similar conclusions were reached by Dell'Acqua et al. (2004). In this study, adult participants performed two tasks: a numerical set comparison task, as to estimate the participant's numerical acuity, and a change detection task to estimate their VWM span. The correlation between individual indices of numerical acuity and VWM capacity was calculated and found to be null. The Bayes factor supported these results, suggesting strong evidence for the absence of a correlation.

Both studies seem to indicate that numerical abilities and VWM skills are independent and have evolved separately. This present study tests this hypothesis in infants. Specifically, if ANS and VWM constitute two systems that have evolved independently, then their developmental trajectories should also be independent from each other, even during the early stages of life.

1.2 Numerical and visual working memory abilities in development

Both working memory and numerical abilities undergo considerable change during human development already at an early age.

Numerical abilities represent the foundations of later-on acquired and complex mathematical notions (Dehaene, 2011). Evidence in support of this comes from different empirical findings, as suggested by Inglis and Gilmore (2013): (i) both non-symbolic arrays and Arabic numerals arrays activate ANS; (ii) children, before the acquisition of formal mathematical notions can perform approximate numerical operations; (iii) ANS acuity, that is the precision of the system, has been found to predict performance at standardized mathematical tests; and (iv) dyscalculic children seem to have a lower ANS acuity, implying that the precision the approximate number system is related to the difficulties found in the formal mathematic domain.

Only one study so far has examined whether VWM and ANS are correlated in early development. Like in adults, in 6-month-old infants, numerical discrimination does not predict the ability to detect a colour change, a measure of VWM, at 9 months (Libertus & Brannon, 2010). This result, consistent with findings in adults by Piazza et al. (2011) and Dell'Acqua et al. (2011), indicate the absence of a predictive relationship between the two systems. Our study, on the other hand, measures individual ability in both systems simultaneously, thus adding a dimension to these findings. An absence of correlation would suggest that not only is there no predictive relationship over time, but there is also no concurrent association between VWM and ANS, strengthening the conclusion that they are functionally and independently distinct systems, and providing a more comprehensive understanding of their relationship and their developmental trajectories.

Furthermore, the study by Libertus and Brannon (2010) did not estimate individual VWM capacities but rather presented infants with a single level of difficulty (Set Size = 2) and calculated the individual ability in this task. In our research, however, each infant was presented with four levels of difficulty, thereby pushing the individual limits of VWM and allowing for the estimation of individual subjects' abilities.

Exploring the relation between working memory and numerical abilities deepens our understanding of cognitive development and may help identify connections between different cognitive abilities and cognitive modules. Further, both working memory and numerical abilities are considered to be fundamental predictors (Figure 1) of mathematical ability during education (Alloway & Passolunghi, 2011; Libertus et al., 2011); and math skills are in turn essential for academic achievement and outcomes. Analyzing the relationship between VWM and ANS can, therefore, contribute to better defining the dynamics underlying learning processes from very early stages of life.

Through this research, we seek to provide new evidence and insights on the relationship between VWM and ANS during the first year of life, paving the way for future theoretical and applied advances in the field of early cognitive development.

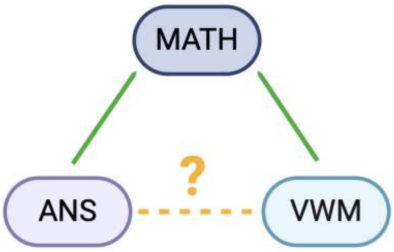

Figure 1: Schematic representation of ANS, VWM and mathematics associations. The green lines indicate a positive correlation, whereas the yellow dotted line highlights the research question of the present study.

1.3 The present study

The general objective of this research is to investigate the absence of a correlation between the span of VWM and ANS acuity in infants. Such a result would suggest that ANS and VWM depend on functionally distinct mechanisms within the cognitive architecture of the human mind from the early stages of life. The study thus aims to investigate the developmental origins of the distinction between the two systems, VWM and ANS, that both contribute to numerical cognition.

Indeed, while both ANS and VWM contribute to the performance in math ability (Alloway & Passolunghi, 2011; Libertus et al., 2011), evidence that these two precursor abilities are not correlated would strengthen the hypothesis that they are two independent precursor mechanisms or modules of the more general math ability. In this view, the separate contributions of ANS and VWM would be necessary to construct an efficient set of mathematical skills. If one of the two components is compromised, the final outcome would be impaired too, as can be observed in different learning disabilities cases. In fact, evidence shows that ANS acuity can distinguish between dyscalculic and non-dyscalculic children, further confirming that a deficit in ANS likely has an impact on children's formal mathematical skills (Mazzocco et al., 2011a). Evidence also shows that VWM skills can predict mathematical and arithmetical abilities, suggesting that working memory processes constitute a fundamental building block of efficient mathematical learning (Alloway & Passolunghi, 2011).

Since mathematical abilities, as well as VWM and ANS themselves, may also be related to general processing efficiency, attention and other cognitive skills, we also assess processing efficiency in a novelty detection task, and relate this to the VWM and ANS measures. For this third variable, we will use a classical a novelty detection (or habituation-dishabituation) paradigm (Benasich & Tallal, 1996; Marino & Gervain 2019). An increased response to a novel stimulus suggests that the memory traces of previous stimuli have been consolidated and that the new stimulus has been recognized as different. The novelty effect has thus been used as a measure of general attention and cognitive efficiency (Berg & Sternberg, 1985). Studies conducted in infants have demonstrated that novelty preference is related to overall cognitive development (Colombo et al., 1989) and later intelligence (Berg & Sternberg, 1985; Lewis et al., 1969).

1.4 Research questions and hypothesis

Following Dell'Acqua and colleagues' (2024) findings, the first research question of the present study asks whether VWM and ANS are two functionally separable modules of cognition. To achieve this, an individual measure of VWM ability and an individual measure of ANS ability will be obtained using well-established VWM and ANS tasks for infants. These two individual measures will then be correlated to examine the extent of covariance between the two variables and determine whether they are associated. Following Piazza et al.'s (2011) and Dell'Acqua and colleagues' (2024) findings in adults, we hypothesize that these two abilities constitute functionally distinct capacities already in infancy, and as a consequence we predict no correlation between the VWM and ANS measures. This ontogenetic continuity may result from our evolutionary inheritance: evolution may have endowed humans with two separate cognitive modules, one functionally suited for retaining visually acquired information, and another for the approximate computation of numerical quantities. If a (positive) correlation is nevertheless found, that could suggest that developmentally, VWM and ANS stem from the same cognitive module and grow to be functionally distinct over the course of life.

The second research question is whether either of these abilities are related to the novelty effect, an index of general cognitive efficiency. If the novelty effect correlates with both VWM and ANS, as expected based on the literature, then this suggests that both mnemonic and numerical capabilities draw upon general processing skills. If, by contrast, novelty detection only correlates with one of the two abilities, likely VWM but not with ANS, then this constitutes further evidence that ANS and VWM are a cognitively distinct modules, and may also be separable from broader attentional processes and cognitive efficiency.

Chapter 2 – Theoretical background of the study

2.1 Visual Working Memory (VWM)

VWM is a limited-capacity system that temporarily stores and manipulates visual information to support ongoing cognitive tasks. It involves the short-term retention of visual stimuli, allowing individuals to maintain and manipulate visual information in their mind for a brief period. Part of the intrinsic definition of VWM is its limited nature and decades of research converge on estimating this number to three or four elements that can be simultaneously held in memory (Luck & Vogel, 1997).

It is also known that VWM plays a crucial role in various cognitive processes, including problem-solving, decision-making, and comprehension, by enabling the temporary storage and manipulation of visual details (Unsworth et al., 2014).

Correlations have been found between VWM span and a broad range of different cognitive abilities, such as cognitive efficiency, executive functions, language skills, mathematical abilities, learning, fluid intelligence, and even practical skills such as driving (Dell'Acqua et al., 2024).

This evidence suggests that storing information about the visual identity of an object constitutes a fundamental ability both for infants and adults. Much research has focused on exploring VWM storage capacity in infants, its development during the first months and years of life as well as its mature state in adulthood.

2.1.1 Brief review of previous research methodologies

Starting from the work of Luck and Vogel (1997) the change-detection paradigm has been used to test the limits or span of the visual cache. In this task, participants are first presented with a sample array of a specific set size (from 1 to 12 elements in Luck and Vogel, 1997). Then after a short delay, a test array is presented. The test array is either identical to or different from the sample array and the subject is asked to detect the eventual presence of a change. For instance, the sample array may consist of three squares, each of a different color. The test array would be very similar, except that the color of one square has changed (Figure 1). Is the subject able to point out such a change?

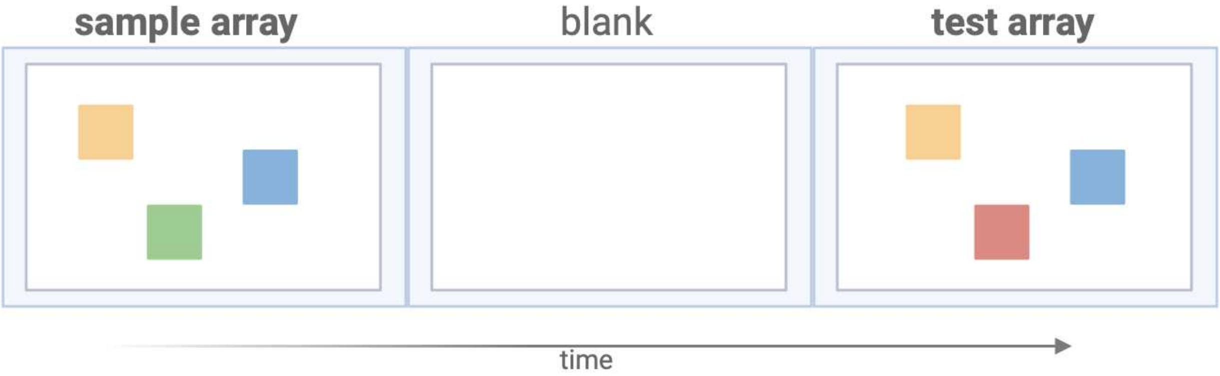

Figure 2: example of Luck and Vogel's (1997) change-detection paradigm

With this procedure, testing for different set sizes, we can estimate the number of elements that can be stored in VWM: if the subject notices changes in arrays of 4 elements, but not in arrays of 5 elements, then that person's VWM span is 4 elements.

The change detection paradigm has also been adapted to infants, who cannot yet verbally express if they noticed a change (Ross-Sheehy et al., 2003). The typical paradigm is implemented as a looking time task. Infants prefer looking at changing or novel displays rather than constant displays, and therefore they look for longer at displays showing changing information (Fantz, 1964).

Based on this change preference, Ross-Sheehy (2003) adapted the classical Luck and Vogel paradigm to infants, presenting simultaneously two displays side-by-side containing colored squares that blink. The two display are identical at the start. On one side, the display continues not changing upon successive blinks (non-changing stream), while in the display on the other side, the color of one square changes at every blink (changing stream). If the infant looks significantly longer at the changing display, then the infant has detected the change. Like for adults, the displays have different set sizes, i.e. different numbers of squares. The span of an infant's VWM is defined by the set size at which he/she still detects the change. In other words, infants' VWM span is assessed as the maximum number of items per display at which infants exhibit a significant preference for the changing display (Ross-Sheehy et al., 2003).

2.1.2 VWM development

Adults can maintain up to 3-4 elements in VWM (Luck & Vogel, 1997), but what is surprising is that the same capacity is observed as early as 10 months (Ross-Sheehy et al., 2003): adult-like VWM span, at least for single feature objects, is reached as early as 10 months.

Ross-Sheehy and colleagues (2003) conducted four experiments to assess VWM span in 4- to 13-month-old infants to map the development of VWM span.

In the first experiment, 6.5-, 10- and 13-month-old infants were tested. Given that VWM strongly depends on maturation of the prefrontal cortex, the functioning of which undergoes dramatical development in that age range, large changes in VWM span were to be expected (Ross-Sheehy et al., 2003). The results showed that 6.5-month-olds significantly looked longer at the changing display only at a set size of 1, while 10- and 13-month-olds showed a significant preference for the changing display at set sizes 1-3: infants' VWM span had increased, allowing them to detect a change in displays containing 2 and 3 elements.

In the second experiment, 4- and 6-month-old infants were tested, and both exhibited a preference for the changing display only at set size 1, thus replicating the results of the first experiment for the 6-month-old group.

In the third experiments, only 10-month-olds were tested, with the former experimental procedure, using set sizes of 2, 4 and 6. Significant preferences for the changing display were observed for set sizes of 2 and 4, but not for set size 6. The degree of preference remained relatively constant across set sizes 1-3 (tested in experiment one) and across set sizes 2 and 4, exhibiting decline only at set size 6. Hence, we can infer that VWM span is around 4 items in 10-month-old infants.

The fourth experiment served as a control to rule out the possibility that failure at larger set sizes observed in younger infants were caused by perceptual or attentional constraints rather than VWM. To do this, 6-month-olds were tested with the same procedure as before, using set sizes of 1-3, but with one exception: the blink was removed. In this version, one display showed a constant image, while in the other, one square changed color randomly every 500ms. This manipulation removed the memory component from the task and the results showed that, by doing so, a preference for the changing display was observed up to set size 3. Hence, the absence of a notable preference in infants at set sizes 2-3 in the prior experiments is attributable to memory constraints rather than limitations on perception or attention.

Interestingly, studies on VWM conducted in older children with the change detection paradigm are in contrast with the idea that VWM span reaches adult-like levels within the first year, by reporting a gradual increase in VWM span from 5 years to adulthood. Nevertheless, a likely reason behind the higher estimation of VWM span found in infants may be due to the method used to estimate VWM span. Infants are tested in an implicit looking time paradigm, while older children are typically tested using tasks that require explicit, often verbal, answers. Consequently, other processes, rather than just VWM span, could cause the developmental changes in performance seen at different ages (Simmering, 2012; Simmering & Perone, 2013).

It is important to underline that the span is only one feature of the VWM system, there are others, e.g. the time for which items are retained in memory. For instance, experimental evidence suggests that, through development, the duration of VWM improves. Thus younger infants may be able to represent multiple elements in VWM, but they may not be able to maintain this information as long as older infants (Diamond, 1990).

2.2 Approximate Number System (ANS) and its development

It might be easy to assume that mathematics is a skill acquired solely through experience and hard work within the school environment, leading to the belief that it is a completely learned competence. Nevertheless, research conducted on unschooled populations, infants, and animals has demonstrated that this is not entirely true: forms of numerical cognition are also present in individuals and organisms entirely unfamiliar with formal mathematics (Roitman & Brannon, 2003).

The simplest aspect of numerical cognition is the ability to distinguish which of two sets is bigger. The adaptive value of such capacity is evident: not only does it underlie an effective foraging strategy, but it can also determine the outcome of group conflict, a common occurrence in the animal kingdom. For instance, lionesses, living in social groups known as prides, base their decision to attack or not attack a rival group on the likelihood of prevailing by comparing the size of their own pride with that of the opposing one, which can be estimated by listening to the number of roars emanating from the rival group (McComb et al., 1994).

Numerical estimation abilities have also been demonstrated in human infants, from the very beginning of their life (reviewed Roitman & Brannon, 2003). Already a few hours after birth, newborns seem to associate abstract numericities across different sensory modalities (Izard et al., 2009). Newborns who were familiarized with an auditory stream, consisting of sequences of syllables each repeated a fixed number of times, demonstrated a significant preference for images presenting a congruent numerosity to that heard during the test phase, as opposed to an incongruent one, when the ratio between the two numerosities was at least 1:3. No significant preference for the congruent numerosity was observed if the ratio was 1:2. This shows that the ANS is operational at birth at least for ratios as large as 1:3 and it operates in a cross-modal manner.

The minimum ratio required for a participant to discriminate between two sets, and thus perceive them as different, serves as a measure of precision of the ANS (Inglis & Gilmore, 2014). The smaller the ratio, the more precisely an individual can estimate large numerosities, as they can discriminate between sets are more similar in quantity.

The ratio indicating the precision of the ANS is related to the Weber Fraction (*w*), i.e. a measure of the spread (standard deviation) of the Gaussian distribution of the internal representation of numerosity (Piazza, 2010). According to this perspective, upon viewing an array of n dots, individuals construct an internal representation that conforms to a normal distribution with a mean of *n* and a standard deviation of *w.* A larger *w* indicates poorer discriminability between similar numerosities, as it leads to greater overlap between the representations of neighboring numerosities, whereas individuals with *w*s closer to zero are more inclined to form representations that closely align with the true value of the numerosity n.

Different studies have tracked the developmental trajectory of the ANS by measuring either the minimum ratio or by calculating the *w* fraction at different ages.

Xu, Spelke and Goddard (2004) have investigated 6-month-old infants' capacity to represent numerosity in visual-spatial displays and discovered that the ANS acuity has already improved with respect to newborns: at 6 months of age, infants can discriminate between sets that differ by a ratio of 1:2.

By 10 months of age, acuity further improves, as infants are able to discriminate sets that differ by a ratio of 2:3 (Xu & Arriaga, 2007).

Other studies suggest that the ANS system spontaneously gets refined during development, until reaching the adult ratio of 7:8, or even 9:10 (Pica et al., 2004).

2.2.1 Brief review of previous research methodologies

Classically, in order to estimate a subject's ANS acuity, a nonsymbolic comparison task is used. The participant is presented with two arrays of different numerosities (n_1 and n_2), one next to the other or one right after the other, and asked which numerosity is the larger. After many trials, the individual's ANS acuity can be measured, in terms of accuracy, *w* fraction, or numerical ratio (Inglis & Gilmore, 2014).

However, when testing subjects who are not (yet) capable of providing a verbal response, such as infants, it is necessary to find an alternative method to estimate their numerical acuity.

A first methodology used to estimate numerical acuity in preverbal infants is the habituation paradigm. The infant is repetitively presented with an array of n_1 elements until their looking time declines to a set criterion of habituation (often, 50 or 60% decrement in looking times compared to an initial baseline). Then the infant is presented with a test array that alternates between the habituated numerosity n_1 and a new one, n_2 (the ratio between n_1 and n_2 can vary across different trials). If the subject detects the novel stimulus, looking times significantly increase, i.e. dishabituation occurs, when the test array is presented (Xu et al., 2005). Dishabituation to the test arrays suggests that infants have perceived the numerical change between the sample and the test array, and are thus sensitive to the tested ratio.

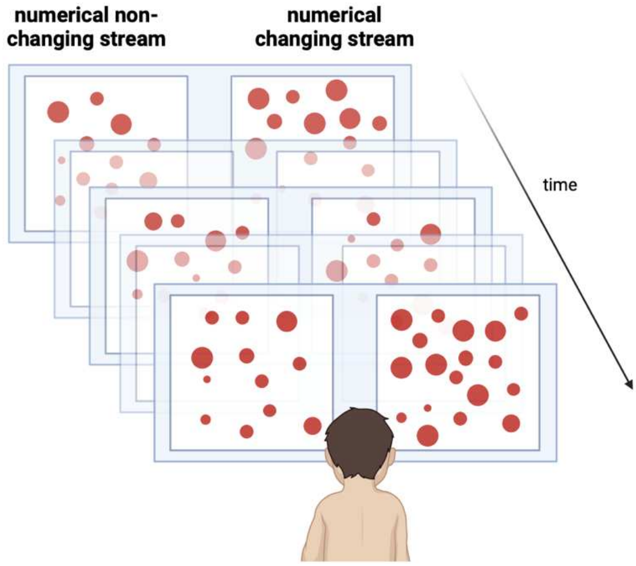

Figure 3: Numerical change detection paradigm. On one side a non-changing numerical stream is shown; on the other side a stream alternating between two numerosities is presented.

A second approach to estimating infants' numerical acuity involves using the numerical change detection paradigm. In this paradigm, two different arrays are presented to the infant on opposite sides of the screen. One array displays a constant numerosity, n_1, while on the other side of the screen, the second array alternates between two numerosities: n_1 and n_2 (and the ratio between n_1 and n_2 varies across different trials). An individual change preference score is calculated. Looking times are expected to be significantly bigger for the changing screen, if the infant can detect the difference between n_1 and n_2 (Libertus & Brannon, 2010). This is the methodology that has been adopted in the present study (see Chapter 3.3.3 ANS task), as it has been demonstrated that the degree of preference reflects inter-individual differences in the accuracy of ANS (Decarli et al., 2023). Furthermore, we also use a change detection task for measuring VMW span. Thus, the two measures are as comparable as possible in terms of task demands, reducing the probability of not find a correlation between the two measures for irrelevant methodological, task-related reasons.

When investigating the ANS it is crucial to control for variables that naturally would covary with numerosity, i.e. two cookies differ from one cookie not only in terms of numerosity (2 rather than 1) but also in volume, density, and surface. The non-numerical, continuous properties that naturally covary with numerosity are: (i) average diameter: the mean dimension of the elements' diameter; (ii) total contour: the sum of the perimeter of all the elements (Figure 2a); (iii) total surface: the sum of the surface of all the elements (Figure 2b); (iv) convex hull: the surface of the hypothetical encircling of all the elements (Figure 2c); (v) density: a measure of the distance between the elements (De Marco & Cutini, 2020). By controlling for these visual features, we can ensure that the participant is not relying on them to provide a response (verbal in the case of adults, or behavioral in the case of infants), and that what we obtain is a measure of the ANS itself. Stimuli used in the present study were controlled for these non-numerical continuous properties, as discussed in Chapter 3.3.3 ANS task.

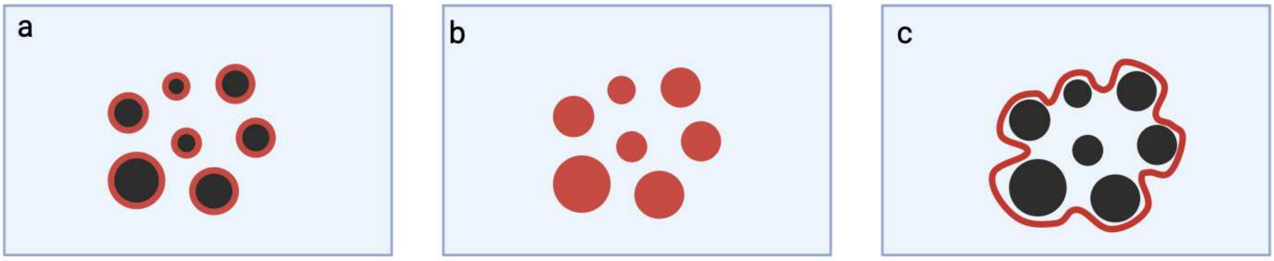

Figure 4: Demonstration of visual features, in red: total contour (a), total surface (b), and convex hull (c).

2.3 Previous studies on the association between Visual Working Memory and Numerical Abilities

Some studies have investigated the relationship between the ANS and VWM, but research has mainly focused on adult subjects, only a few studies have investigated infants and children.

In Piazza et al. (2011), conducted with adults, three tasks were presented to each participant: (i) dots counting task: 1-8 dots were presented and participants had to name the number as quickly and as accurately as possible - this task was used to estimate s, the subitizing capacity of the Object Tracking system; (ii) a VWM task using the classical paradigm (Luck and Vogel, 1997): participants had to judge a test array as being the same as or different from the sample array - this task was used to estimate K, the VWM span; and (iii) a dots comparison task: two dot arrays were presented and the participant had to indicate which is larger by pressing a button – this task was used to estimate w, ANS precision. The abilities were then correlated. Results showed no correlation between subitizing capacity (s) and numerosity estimation precision (w), suggesting that ANS and the Object Tracking System are two different systems, although they both contribute to the broader numerical cognitive module. A strong linear correlation was found between VWM span (K) and subitizing capacity (s) ($R = 0.73,\ p = 0.001$), supporting the hypothesis that the subitizing system reflects a boarder parallel individuation system common to any task that requires the individuation of multiple objects (such as the VWM task). No correlation was found between VWM span (K) and ANS precision (w) ($R = 0.07,\ p = 0.8$).

Dell'Acqua et al. (2024) identified some potential issues in Piazza et al.'s (2011) study, both in the design of the change detection task and in the calculation of K, which could have let a correlation between K and w go undetected. These authors thus set out to replicate Piazza et al. (2011) with improved methodologies, again testing adults. But even when correcting for these methodological flaws, Dell'Acqua and colleagues (2024) still found no correlation between ANS acuity and VWM span; even using Bayesian statistics, which indicated that the evidence supporting the lack of correlation was "substantial".

Libertus and Brennon (2010) tested sixteen infants longitudinally, both at 6 and at 9 months. At 6 months, infants were presented with a number discrimination task, whereas at 9 months, both with the same number discrimination task and a VWM task. While numerical change detection score at 6 months accurately predicts that obtained at 9 months, no correlation was found between the numerical change detection score at 6 months and the color change detection score at 9 months (Libertus & Brannon, 2010).

Despite evidence suggesting the independence of VMW and ANS, both these systems constitute fundamental components for the learning of formal mathematics (Alloway & Passolunghi, 2011).

The role of ANS in math learning has been well documented, with many studies suggesting that the ANS is a fundamental precursor of mathematical knowledge (Halberda et al., 2008; Libertus et al., 2011; Mazzocco et al., 2011a, 2011b). In a study by Libertus and Brannon (2011), 200 children aged 4 years were tested with a number comparison task to measure their ANS acuity and with the Form A of the Test of Early Mathematics Ability (TEMA-3; Ginsburg & Baroody, 2003), a measure of children's math abilities. Results showed that faster reaction times and greater accuracy on the ANS acuity task were associated with higher math ability. It needs to be noted, however, that Libertus and Brannon's (2011) study was correlational, therefore does not allow temporal nor causal inferences to be drawn. The relationship between ANS and formal maths ability could work either way: those kids who perform better in the standardized TEMA-3 test, since they are better in math, may be faster and more accurate in the ANS task, or the other way round, those children who have a sharper ANS can better perform in standardized math tests, or further still, the relationship between the two may be mediated by a third variable, e.g. attentional abilities etc. A longitudinal study with random assignment, e.g. a training study, will be needed to establish temporality and causation.

A longitudinal study by Mazzocco, Feigenson, and Halberda (2011) first obtained an ANS measure when children were preschoolers and then tested the same children 2 years later, once they were in school with the TEMA-3 test. The authors found an association between ANS precision prior to formal mathematical education and math performance after schooling, as the ANS precision accounted for 28% of the variance in TEMA-3 ($r^2 = .278,\ \mathrm{t}\,(16) = 2.405, \mathrm{p} = .030$).

Working memory also seems to play a role in math acquisition. Alloway and Passolunghi (2011) tested 206 children, whose age ranged between 7 and 8 years. Children were tested for their working memory through the administration of all the 12 tests from the Automated Working Memory Assessment (AWMA, Alloway, 2007), which assesses visuo-spatial and verbal working memory. To assess the formal mathematical and arithmetical abilities, children were presented with the Italian

AC-MT test (as the participants were Italian), which consists of four tasks with different levels of difficulty depending on the age of the student (Cornoldi, Lucangeli, & Bellina, 2002) and with the Numerical Operations subtest from the Wechsler Objective Numerical Dimensions (WOND, Wechsler, 1996). The study concluded that working memory can uniquely predict mathematical abilities, and, in particular, that VWM accounted for significant variance in predicting scores in many of the AC-MT subtests (Quantity Discrimination (14%) and Number Production (21%), Number Ranking (23%); Number Production (10%); Number Operations (19%)); and the Arithmetic test from the WOND (35%).

2.4 The present study

In light of the above findings, the present study asks whether VWM and ANS are two functionally separable modules of cognition and whether either of these abilities are related to the novelty effect, as an index of general cognitive efficiency, in 9–12-month-old infants. This age range was chosen based on existing literature. As the aim was to correlate the VWM and ANS abilities, it was important to establish an individual index of acuity for both VWM and ANS systems and thus select an age at which participants could be tested with varying levels of difficulty in which we could measure certain degree of variability in their performance.

To address these research questions, the present study utilizes a correlational approach. This is motivated by the existence of a prior study with infants (Libertus & Brannon, 2010) but with several key distinctions. First, the previous study did not estimate individual VWM span. Second, the previous study measured VWM and ANS abilities 3 months apart, a significant period in early development during which substantial changes could occur. In contrast, our study measures individual VWM and ANS abilities for each infant in a single testing session, and using multiple levels of difficulty, thus allowing for individual estimation of VWM and ANS abilities.

For both the VWM and ANS tasks, we employed a change detection paradigm, a well-validated approach for the estimation of both VWM and ANS abilities (e.g., Libertus & Brannon, 2010; Ross-Sheehy et al., 2003).

To assess individual VWM span, we presented infants with a change detection paradigm involving two displays: a non-changing display and a changing display. In the non-changing one, 1, 2, 3, or 4 colored squares were presented repeatedly, with no changes occurring. In the changing display, the color of 1 square (within a set of 1, 2, 3, or 4) changed color with each new presentation. This manipulation allowed us to estimate a change preference index for each set size and infants' VWM span is computed as the maximum number of items per display at which infants exhibit a significant preference for the changing display.

To assess individual ANS abilities, we employed a change detection paradigm similar to that used for VWM assessment. Infants were presented with two displays: a non-changing display and a changing display. In the static display, a constant number of dots was presented repeatedly. In contrast, in the dynamic display, two different numerosities, presented in a ratio of 1:2, 2:3, or 3:4, alternated with each new presentation.

For the novelty effect task, we employed an habituation-dishabituation paradigm, presenting images of adult faces until habituation was reached, and then introducing a novel infant face.

From these three tasks, we calculated three individual indices: a VWM span index, an ANS ability index, and a novelty effect index. We hypothesize that ANS and VWM constitute functionally distinct capacities already in infancy, and as a consequence we predict no correlation between the VWM and ANS measures. We further predict that if the novelty effect correlates with both VWM and ANS, as expected based on the literature, then this suggests that both mnemonic and numerical capabilities draw upon general processing skills.

Chapter 3 – Methodology

3.1 Participants

Fourteen infants (6 females) participated to the study. One participant was later excluded from data analysis due to substantial parental intervention during the experiment, so that only 13 participants entered the final data sample. All infants were born full-term and had no history of hearing or visual impairments. Participants' mean age was 10 months and 26 days ($SD = 22{,}57$ days; range 9-12 months).

The study was approved by the ethics committee of the Università degli Studi di Padova. All parents of all participating infants gave written informed consent prior to participation in accordance with the Declaration of Helsinki (the informed consent document is included in the Supplementary Material section). Participants were recruited from the BabyLab, Università degli Studi di Padova database and contacted via e-mail.

3.2 Stimuli

3.2.1 VMW task

The stimuli were created in Keynote, then transformed into images and assembled in videos. The initial square colors were selected at random from a set of nine colors, which were green (R:114, G:246, B:74), brown (R:163, G:123, B:75), orange (R:241, G:153, B:55), violet (R:135, G:43, B:141), cyan (R:116, G:249, B:163), blue (R:20, G:49, B:244), yellow (R:255, G:252, B:83), red (R:234, G:64, B:37), and white (R:255, G:255, B:255), whereas the background was gray (R:135, G:135, B:135). Each square measured 2,7 x 2,7 cm on the experimental monitor. The colors within a display were always different from each other, but the same colors could appear on both monitors.

The changing stream was created from the non-changing one by randomly selecting one square and varying its color to another color (different from those already on the array). In this way, at every blink of the changing stream, all the squares, except one, were identical to their previous presentation. The set size (the number of squares on each array) was identical for the two arrays and remained constant throughout a trial (for both the changing and the non-changing streams).

Additionally, an attention getter stimulus was created to be used between trials. It was a blue flower that rotated and emitted a sound.

3.2.2 Novelty detection task

The habituation stimuli consisted of two identical copies of an adult female face with a neutral expression side by side (size on the screen: 26 × 20 cm). In the test stimuli, one of these was replaced by the face of a male child (size on the screen: 23 × 20 cm).

3.2.3 ANS task

The stimuli used in this task were a series of images containing dots of varying numerosity: 12, 16, 18, or 24. The “changing stream” consisted of videos with 12 dots alternating with 16 dots (ratio 3:4), or 12 with 18 (ratio 2:3), or 12 with 24 (ratio 1:2). In the "non-changing stream", the video always showed the same numerosity (12, 16, 18 or 24). Images within a video stream lasted 500 ms, interspersed with 250 ms of blank screen. The dots were white on a gray background. The stimuli were the same as those used in Dell’Acqua et al. (2024). These images were generated using the CUSTOM package in Matlab, controlling for various non-numerical parameters, which would otherwise naturally covary with numerosity. We thus ensured that for half of the images, the convex hull and total contour remained constant, while for the other half, the density and total area did. Dot size varied across all images.

3.3 Experimental procedures

Upon arrival at the BabyLab, the infants were acquainted with the experimental environment, while information about the child was gathered through a brief interview with their parents. When the child appeared comfortable, the infant and one parent entered in the testing booth. The caregiver was seated inside the testing booth and instructed to hold the child on their lap and advised not to interfere with the child's behavior, but to soothe them if necessary. Additionally, the parents were provided with glasses with opaque lenses to prevent them from seeing the stimuli presented on the screen and thus influencing the child's visual behavior. During testing, the lights in the booth were dimmed, and the stimuli were presented to the child on a central monitor, above which a camera was positioned. Three tasks were used: VWM span, novelty effect and ANS acuity. The three tasks were presented in the same order for every infant, not to introduce variability related to task order: the VWM task was first, followed by the novelty effect task and lastly by the ANS task.

The VWM task was first, as it was considered the engaging, thus the most likely to capture the child's curiosity without initially causing too much fatigue. The ANS task was last as it was deemed slightly more demanding and less engaging. If infants fussed out during this task, they at least provided data in the two previous ones.

Each task duration was infant controlled, meaning that, if the infant looked away from the screen for longer than 2 seconds, the software automatically proceeded to the following trial. Nevertheless, a maximum trial duration was set, so that the VWM task could last up 3 minutes, the novelty task up to 2 minutes and the ANS task up to 8 minutes. Thus the whole experimental session could last up to 18 minutes.

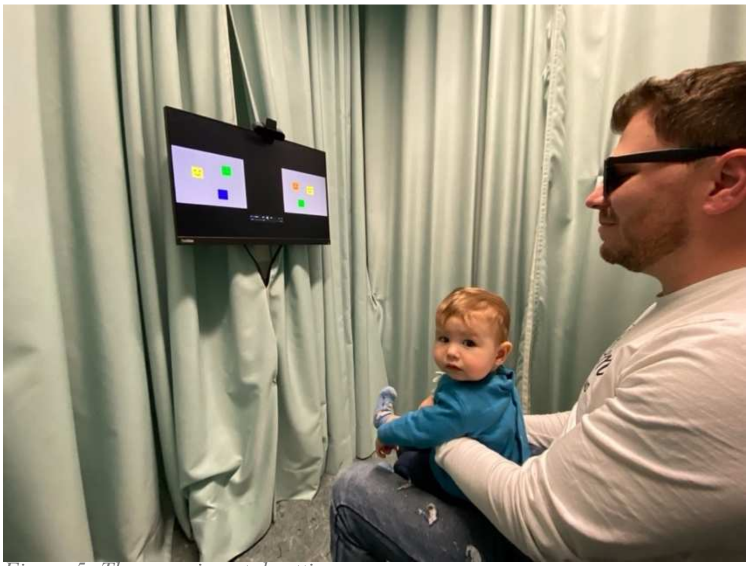

Figure 5: The experimental setting

The experiment was run with Habit 2.1 on a Mac with OS Mojave, version 10.14.6. The session was recorded through the camera placed above the screen. The camera transmitted the image to the experimenter, who was standing outside the booth and thus blind to the experiment.

Behind the curtain, the experimenter controlled stimulus presentation as a function of infants' online looking behavior, while the session was also video recorded, and subsequently coded offline. Data analysis was based on the offline coding. At the end of the testing session, families were provided with a souvenir photo and given a BabyLab diploma and a small gift (such as bibs, soap bubbles, small toys, or bags) as a token of gratitude.

3.3.1 VWM task

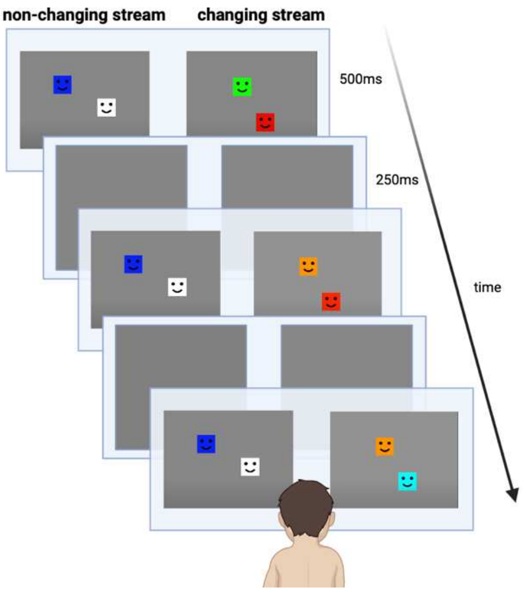

Figure 5: VWM span task. A changing and a non-changing stream are displayed, each on one side of the screen. Images are shown for 500ms, interleaved by 250ms of blank screens.

Based on Ross-Sheehy et al. (2003), VWM was assessed with the change detection paradigm. Two videos were simultaneously presented, one on the right and one to the left of the screen (Figure 5). Each video was composed of images representing one, two, three, or four smiling and colored squares, displayed for 500ms, alternating with 250ms of blank screen ("blink"). If the infant did not look at the screen for longer than 2 seconds, the trial was terminated and the next trial was automatically presented, i.e. trial length was infant controlled. Maximum trial duration was 20 seconds. There was a total of eight trials: two for each set size, with the changing stream appearing once on each side of the screen. The side of the changing stream was randomized across infants and counterbalanced within trials, so that each infant saw the changing stream on the left for half the trials and on the right for the other half. The changing side could not be the same more than twice in a row for each infant. The order of the trials was randomized across infants. Between each trial, an attention getter was displayed.

3.3.2 Novelty detection task

Based on Marino and Gervain (2019) infants were shown on a screen two identical copies of an adult female face until habituation was reached (habituation phase). Trials ended if the infant looked away for longer than 2 s. Maximum trial duration was 10 seconds. The mean looking time in the first two trials defined the baseline looking time (100%). The identical stimuli (adult faces) were presented repetedly until the habituation criterion was attained. This criterion was defined as a looking time equal to or less than 50% of the baseline.

Immediately after the habituation phase, the test phase followed, where one of the adult faces was replaced by the face a male child. Each infant underwent two test trials, one with the novel stimulus on each side of the screen. The side of the novel face was counterbalanced between the participants (for half of the participants the novel face appeared on the right in the first trial, for the other half on the left). Again, a look away longer than 2 seconds ended the trial. The session was recorded and coded offline.

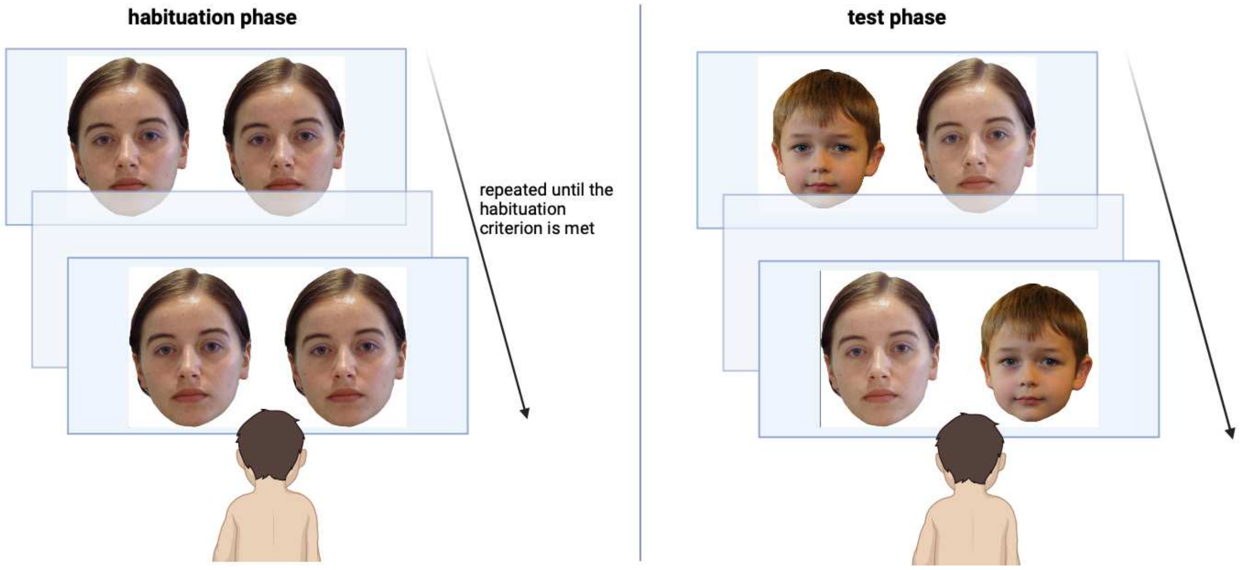

Figure 6: Novelty task: in the habituation phase two adult faces are shown until the habituation criterion is met. Then a test phase follows: a baby face replaces one of the two previous adult ones, first on one side and then on the other.

3.3.3 ANS task

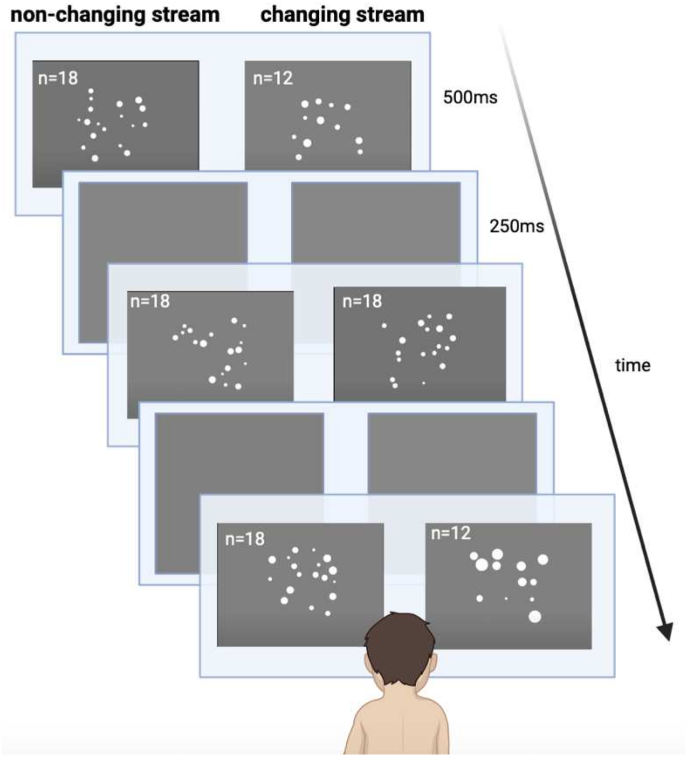

Figure 7: ANS task. A numerical changing and a non-changing stream are displayed, each on one side of the screen. Images are shown for 500ms, intervealed by 250ms of blank.

The paradigm used for the ANS task, like that of VWM, was also based on change detection. The “changing stream” was presented on one side of the screen, the "non-changing stream" on the other. Maximum trial duration was 40 seconds, but if the participant looked away from the screen for more than two seconds, the software automatically proceeded to the next trial (infant controlled). Each infant was presented with 12 trials, testing all possible combinations of the four numerosities on each side of the screen and each possible ratio (3:4, 2:3 and 1:2). The order of the trials was randomized between the participants and counterbalanced within the same subject so that, for each infant, the “changing stream” could not appear on the same side more than twice in a row. As in the previous tasks, the attention getter was displayed between trials.

3.4 Data Analysis

Looking times were offline coded using ELAN_6.7 software. On the basis of looking times, change preference scores were computer for both the VWM and the ANS task to determine whether the infant looked longer at the changing stream over the non-changing one. For the Novelty detection task, looking times to the novel face as compared to the familiar face were computed. Specifically, we computed the following measures for each task.

3.4.1 Individual VWM measures

- Total looking time: the total looking time calculated as the cumulative time that each infant spent looking at one of the two screens ($changing + non\_changing$ screen).
- Change Preference Score: for each subject, a change preference score was calculated for each span (1,2,3 and 4) to assess the difference in time spent looking at the changing and unchanging arrays. These preference scores were determined by dividing the time spent looking at the changing display by the total looking time (i.e., $changing \div (changing + non\_changing)$. A score of 0.5 indicates that the infant looked equally at both the changing and unchanging displays, while a score of 1.0 indicates exclusive attention to the changing display.
- Preference Score for the left side: a preference score for the left side was also computed to test an a-priori preference for one side of the screen rather than the other.

3.4.2 Individual Novelty Effect

Due to a technical error in the settings of the habituation procedure, measures recorded in the Novelty Effect task could not be analyzed.

3.4.3. Individual ANS measures

- Total looking time: the total looking time calculated as the cumulative time that each infant spent looking at one of the two screens ($changing + non\_changing$ screen).
- Change Preference Score: an individual change preference score was calculated for each ratio (1:2, 2:3 and 3:4) to assess the difference in time spent looking at the changing and unchanging arrays. As for VWM, preference scores were determined by dividing the time spent looking at the changing display by the total looking time (i.e., $changing \div (changing + non\_changing)$, so that a score of 0.5 indicates that the infant looked equally at both the

changing and unchanging displays, while a score of 1.0 indicates exclusive attention to the changing display.

- Preference Score for the left side: a preference score for the left side was also computed to test an a-priori preference for one side of the screen rather than the other.

Chapter 4 – Results

4.1 Preference for side

As a first analysis, I checked whether infants showed a bias for looking longer at one side of the screen. A preference score for the left side was computed. The results of a two-tailed one sample t-test revealed no a-priori preference for one side of the screen over the other ($M = 0.56,\ SD = 0{,}137,\ t(12) = 1.59,\ p = 0.137$). Since infants showed no side preference, side was not included as a factor in any further analysis.

4.2 VWM results

The average total looking time and change preference scores at each set size are reported in Table 1 and displayed in Figure 8.

Table 1: Descriptive statistics of change preference scores (PS) and overall looking times (LT) across different set sizes (1, 2, 3 and 4) in the VWM task.

Descriptive Statistics								
	PS				**LT**			
	1	**2**	**3**	**4**	**1**	**2**	**3**	**4**
Mean	0.592	0.598	0.520	0.570	14.403	12.941	13.799	13.387
Std. Deviation	0.122	0.090	0.127	0.154	5.572	5.861	5.208	6.205
Minimum	0.297	0.388	0.369	0.290	3.422	3.393	4.545	1.650
Maximum	0.763	0.727	0.746	0.876	19.822	19.895	19.895	19.718

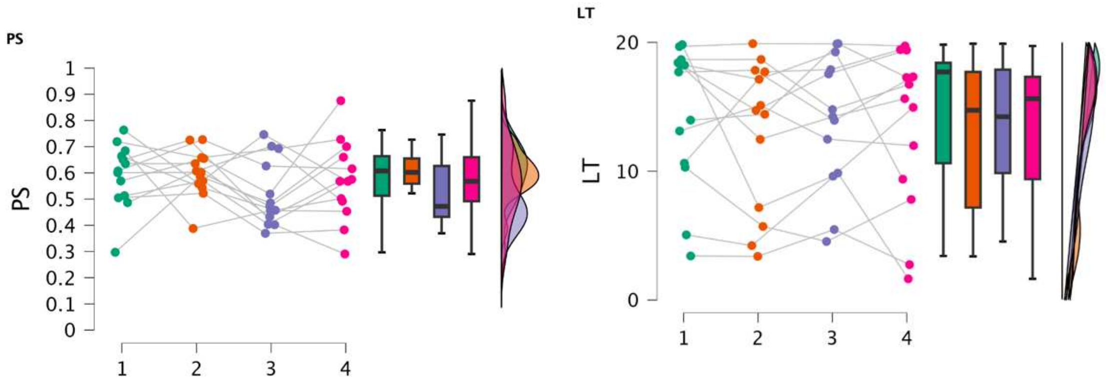

Figure 8: Boxplot of change preference scores (PS) and looking times (LT) across different set sizes in the VWM task.

Infants' mean preference scores across all set sizes significantly exceeded chance (equal to 0.5) ($M = 0.57,\ SD = 0.12,\ t(51) = 4.038,\ p =< .001,\ d = 0.56$) suggesting that infants significantly preferred to look at the changing display over the non-changing one. Nevertheless, two-tailed t-tests for each set size revealed that this was true only up to set size 2, whereas preference scores obtained for set sizes 3 and 4 were not significantly different from chance (results are reported in Table 2).

Table 2: Results of One Sample t-test for each Set Size

One Sample T-Test					
	t	**df**	**p**	**Cohen's d**	**SE Cohen's d**
PS Set Size 1	2.726	12	0.018	0.756	0.314
PS Set Size 2	3.945	12	0.002	1.094	0.351
PS Set Size 3	0.579	12	0.574	0.160	0.279
PS Set Size 4	1.644	12	0.126	0.456	0.291

Bayesian analyses were also conducted in order to check for evidence in favor of a genuine null preference in set sizes 3 and 4, as opposed to a simple lack of statistical power to detect an effect, and the results are presented in Table 3. These results indicate strong evidence that the mean preference score is greater than 0.5 for Set Size 1, with a Bayes Factor (BF) suggesting moderate support ($BF_{10} = 3.538$). For Set Size 2 ($BF_{10} = 22.5589$), there is strong evidence that the mean PS is greater than 0.5, marking this set size as having the most substantial evidence among the four. In contrast, Set Size 3 ($BF_{10} = 0.322, BF_{01} = 3.108$) provides evidence against the hypothesis that the mean PS is greater than 0.5, indicating a genuine lack of preference, i.e. no discrimination for the changing stream. Finally, for Set Size 4 ($BF_{10} = 0.812, BF_{01} = 1.232$), results provide no evidence for either preference or lack thereof due to lack of statistical power.

Table 3: Results of Bayesian One Sample T-Test

	BF_{10}
PS Set Size 1	3.538
PS Set Size 2	22.560
PS Set Size 3	0.322
PS Set Size 4	0.812

A repeated measure ANOVA with Set Size as a within-subject variable (1 / 2/ 3/ 4) revealed no effect of Set Size on change preference scores ($F = 1.216,\ n_p^2 = 0.092,\ p = 0.31$). A similar ANOVA over the looking times ($F = 0.543,\ n_p^2 = 0.045,\ p = 0.656$) did not show significant differences either.

The ANOVA analysis for preference scores was also performed using Bayesian statistics to test for evidence in favor of a lack of difference between the different set sizes, and the results suggest that the posterior probability is significantly higher for the null model (0.737) than for the model including Set Size as a factor (0.263) (Table 4). This indicates stronger support for the null model from the data. However, the Bayes Factor for the null model is not particularly strong ($BF_M = 2.798$), not going beyond the conventional threshold of 3. This analysis thus suggests that there is lack of statistical power to support either hypotheses in the data.

Table 4: Results of the Bayesian ANOVA.

Model Comparison					
Models	**P(M)**	**P(M|data)**	**BF_M**	**BF_{10}**	**error %**
Null model	0.500	0.737	2.798	1.000	
Set Size	0.500	0.263	0.357	0.357	0.399

4.3 ANS results

The average total looking time and change preference scores at each ratio are reported in Table 5 and displayed in Figure 9.

Table 5: Descriptive statistics of change preference scores (PS) and overall looking times (LT) across different ratios (1:2, 2:3 and 3:4) in the ANS task.

Descriptive Statistics						
	PS			**LT**		
	1:2	**2:3**	**3:4**	**1:2**	**2:3**	**3:4**
Mean	0.572	0.478	0.525	13.239	13.810	10.802
Std. Deviation	0.208	0.134	0.118	9.378	9.618	8.592
Minimum	0.289	0.257	0.229	1.299	2.000	1.275
Maximum	0.950	0.698	0.698	31.704	31.178	27.397

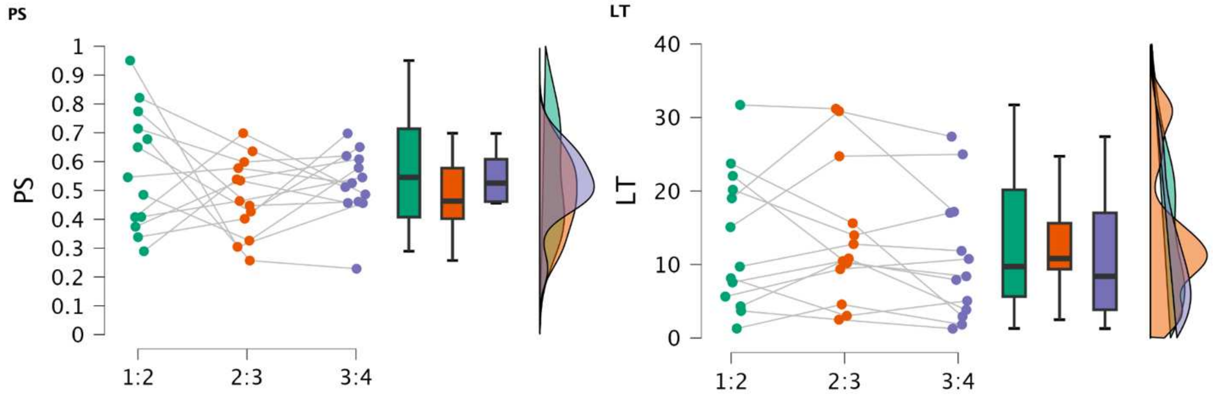

Figure 9: Boxplot of change preference scores (PS) and looking times (LT) across different ratios in the ANS task.

Overall, infants' mean preference scores over all ratios did not significantly differ from chance ($M = 0.57,\ SD = 0.12,\ t(51) = 4.038,\ p =< .001,\ d = 0.56$) suggesting that infants had no preference for the changing display over the non-changing one. Results of two-tailed t-tests for each Set Size are reported in Table 6.

Table 6: Results of One Sample t-test for each ratio

One Sample T-Test					
	t	**df**	**p**	**Cohen's d**	**SE Cohen's d**
PS 1:2 ratio	1.247	12	0.236	0.346	0.286
PS 2:3 ratio	-0.599	12	0.560	-0.166	0.279
PS 3:4 ratio	0.763	12	0.460	0.212	0.280

Bayesian analyses were also conducted, and the results are presented in Table 7. These comparisons do not show evidence for the discrimination of numerosities. The results for rations 1:2 and 3:4 provide no evidence in favor of discrimination or lack thereof due to low statistical power, while for the 2:3 ratio, there is moderate evidence favoring lack of discrimination ($1{:}2\, ratio{:}\, BF_{10} = 0.529, BF_{01} = 1.89; 2{:}3\, ratio{:}\, BF_{10} = 0.325, BF_{01} = 3.076; 3{:}4\, ratio{:}\, BF_{10} = 0.357, BF_{01} = 2{,}799$).

Table 7: Results of the Bayesian One Sample T-Test

	BF_{10}
PS 1:2 ratio	0.529
PS 2:3 ratio	0.325
PS 3:4 ratio	0.357

A repeated measure ANOVA with Ratio as a within-subject variable (1:2 / 2:3 / 3:4) revealed no effect of Ratio on the change preference score ($F = 1.011,\ n_p^2 = 0.078,\ p = 0.379$). A similar ANOVA over looking times ($F = 1.282,\ n_p^2 = 0.097,\ p = 0.296$) revealed no significant effects either.

The ANOVA analysis for preference scores was also performed using Bayesian statistics, and the results suggest that the posterior probability is significantly higher for the null model (0.701) compared to the model including Ratio as a factor (0.299) (Table 8). This indicates stronger support for the null model from the data. Yet, like before, the BF for the null model (BF_M=2.349) doesn't provide evidence favoring this model, likely due to lack of statistical power.

Table 8: Results of the Bayesian ANOVA.

Model Comparison					
Models	**P(M)**	**P(M|data)**	**BF_M**	**BF_{10}**	**error %**
Null model	0.500	0.701	2.349	1.000	
Ratio	0.500	0.299	0.426	0.426	0.585

4.4 Correlation between VWM and ANS

To measure the strength of the association between VWM and the ANS, a correlation was computed. The VWM index chosen was that obtained by subjects at Set Size 2, as this was the highest Span at which significant change detection was achieved. The ANS task did not yield significant results. Given this, there were two possible options to derive an ANS index. One was to use the mean preference score obtained across the three ratios as the ANS index. With this option, no statistically significant correlation was found between the VWM index and the ANS index ($r = -0.317,\ p = 0.292$).

The same analysis was also computed using a Bayesian analysis, comparing the likelihood of two competing hypotheses (an absence or a presence of a significant correlation). The Bayes Factor (BF_{10} = 0.566, BF_{01} = 1.768) suggests that there is insufficient evidence for either hypotheses. This suggests that the data do not provide sufficient support for the existence of a correlation, nor against it. More data or a larger effect size might be necessary to draw a definitive conclusion.

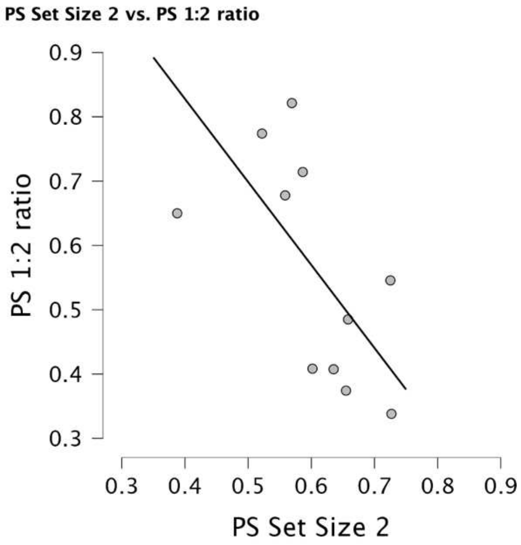

Figure 10: Pearson's correlation between preference scores (PS) obtained in the VWM task at Set Size 2 and those obtained in ANS task at Ratio 1:2

The second option was to use the preference score obtained at the 1:2 ratio, which, based on the literature, infants at the age tested in our study should succeed with. Using this option, a statistically significant negative correlation (r=-0.557,p=0.048) was found between the two indices. The distribution of the indices is shown in Figure 10.

The same analysis was also computed using a Bayesian analysis, comparing the likelihood of two competing hypotheses (an absence or a presence of a significant correlation). The Bayes Factor (BF_{10} = 2.003) indicates insufficient evidence for the presence of a correlation, but also insufficient evidence for the alternative hypothesis (BF_{01} = 0.499). So the evidence is not strong enough to make a definitive conclusion, and further investigation would be warranted to solidify these findings.

Chapter 5 – Discussion and Conclusions

5.1 Implications of the Results

The present study investigated the relationship between VWM and ANS in 10-month-old infants. While the lack of association between these two systems in adults suggests that they are functionally independent, it remained to be determined whether they shared a common developmental origin or if VWM and ANS were independent from the early stages of life.

The results obtained in infants aged 9-12 months suggest that the VWM span at this age is limited to 2 items. Infants showed a significant preference for the screen containing a change in sets of 1 and 2 items, but not when the set had a larger number of items. This measure of VWM capacity is lower than that reported by Ross-Sheehy et al. (2003) for infants of the same age, indicating a VWM span of 4 elements in 10-month-old infants, equivalent to that of an adult subject. The underestimation of the span in our study, compared to Ross-Sheehy's findings, could be attributed to differences in sample size. Our results were derived from a sample of just 13 subjects, while Ross-Sheehy's study included 48 participants. This smaller sample may lack the statistical power to detect preferences at larger set sizes, where effects are likely to be smaller due to increased task difficulty. In contrast, Ross's larger sample of 48 ten-month-old infants had sufficient power to identify these subtle preferences, underscoring the importance of a larger sample size for detecting small effects. In both the present study and that conducted by Ross-Sheehy et al. (2003), no significant effect of Set Size on preference scores was observed. Given the results of the Bayesian analysis, which confirmed that the model containing Set Size does not explain the data better than a Null model, this outcome does not appear to be simply due to a lack of statistical power stemming from the small sample size. The Bayesian evidence suggests that the inclusion of Set Size does not add significant explanatory value, indicating that the absence of a significant effect is not merely a consequence of insufficient sample size, but likely reflects the actual relationship in the data. The absence of a main effect suggests that there is no significant difference in the preference scores across the Set Sizes. It is interesting to note how this result aligns with findings regarding the subitizing system, which in adults exhibits strong correlations with the ANS (Piazza et al., 2010). Indeed, the subitizing system is capable of extracting numerical information from 3-4 elements in parallel (Trick & Pylyshyn, 1994). Set Size was found to exert a significant effect on accuracy in the VWM task in adults (Dell'Acqua et al., 2024). However, it is important to highlight a significant difference in tasks between adults and infants: while in adults, the task is explicit, in infants it is implicit.

From a methodological perspective, we can conclude that the VWM task worked well, showing significant preferences for change (at least up to Set Size 2) and thus contributing to the validation of the change detection paradigm for measuring memory abilities in infants. The looking time patterns obtained (e.g. sufficiently long looking times, frequent changes of eye gaze between the two streams etc.) further suggest that the stimuli were engaging for the participants.

Results obtained from the numerical change detection task, by contrast, did not allow us to establish a ratio at which infants of this age reliably perform the task. Despite literature suggesting that around 10 months of age, infants should be able to distinguish between two sets when they differ by at least a 2:3 ratio (Xu & Arriaga, 2007), participants in this study did not show a preference for change even at a 1:2 ratio, which is expected to be reached by around 6 months (Xu et al., 2005). This failure to replicate earlier findings might be due to the lack of statistical power resulting from the small sample size analyzed. Specifically, for the preference ratio of 1:2, the Bayesian analysis provided no evidence in favor nor of lack of discrimination, suggesting that the absence of significant effects for this ratio may be due to the limited sample size rather than the absence of the underlying effects observed in previous research. The lack of discrimination of numerosities is in contradiction with existing findings (Decarli et al., 2023). Unfortunately, since infants did not discriminate numerosities, we could not draw clear conclusions about the ANS system. We thus opted for two ways to derive the

ANS index: using the average preference scores across all ratios and using the preference score at the 1:2 ratio, which infants should theoretically be able to discriminate.

Finally, no significant correlation was found between VWM and ANS, when using the average preference score as the ANS index, whereas a strong negative correlation was found, when using the preference score at the 1:2 ratio. Considering the small sample size, this negative correlation needs to be interpreted with caution, and needs to be confirmed at larger sample sizes.

An absence of correlation between the measures obtained for the two systems suggests that they are functionally independent. This result is in line with the longitudinal findings of Libertus and Brannon (2010), which demonstrated that numerical preference scores at 6 months did not predict VWM preference scores at 9 months. Furthermore, this result is consistent with findings in adults, whose measures of numerical acuity do not correlate with those of ANS (Dell'Acqua et al., 2024), but show a strong correlation with those of subitizing (Piazza et al., 2011).

Research conducted in adults led to the theorical conclusion that while subitizing relies on a multi-purpose and attention-demanding system that represents elements as discrete entities, which are encoded in VWM, numerical estimation of large sets relies on an ancestral and pre-attentive system that does not tap onto VWM (Dell'Acqua et al., 2024).

The lack of a correlation between VWM capacity and ANS acuity, if confirmed in a larger group of infants, suggests that these cognitive processes depend on distinct mechanisms from early stages of life, confirming the modularity of our cognitive architecture, where specialized subsystems perform different functions.

If ANS and VWM are distinct systems, they are likely supported by separate neuronal substrates in the brain. This notion might appear contradictory to findings on the anatomical and functional organization of brain processes involved in these abilities, suggesting overlap between the cortical regions engaged in VWM and numerosity processing, specifically in the inter-parietal sulcus (IPS) region (Brigadoi et al., 2017; Harvey et al., 2015).

One of the possible mechanisms distinguishing the ANS and VWM systems is attention. The role of attention in VWM is well-documented, with evidence suggesting that VWM capacity is strictly intertwined with selective attention: attention is the mechanism through which we select relevant information, allowing us to correctly maintain it in VWM and subsequently retrieve it when necessary (Theeuwes et al., 2011). By contrast, evidence suggests that numerical estimation of large sets is a pre-attentive ability. Performance on subitizing is negatively affected by attentional load, while the estimation of large numerosities remains robust under such conditions (Piazza et al., 2011; Revkin et al., 2008).

The results of the present study, if confirmed with a larger sample of infants, corroborate these hypotheses and provide information about the developmental origin of such distinctiveness. Indeed, having demonstrated a lack of relationship between the two systems in infants, we can hypothesize that such differentiation does not stem from ontogeny but is instead phylogenetically endowed, allowing organisms to quickly estimate large sets in a ballistic, immediate, and pre-attentive way. Being able to to estimate the numerosity of food resources, or rivals, not depending on attentional resources could be crucial for an organism's survival.

5.2 Limitations of the Study

To draw meaningful and valid conclusions, it is essential to carefully examine the potential limitations of the study.

One challenge of the present research lies in its hypotheses: it was conceived with the intention of investigating whether, as in adults, the ANS and VWM systems are functionally independent in infants, as evidenced by their lack of correlation. Hypothesizing a null result is inherently challenging, as it is difficult to determine whether the effect is genuinely absent or if there is insufficient statistical power or other methodological difficulties to detect it. To address this issue, we used Bayesian statistics, which assesses not only the strength of the evidence in favor of the alternative hypothesis, but also that in favor of the null hypothesis. It thus allows us to distinguish between a null result supported by evidence and a null result deriving from methodological issues.

Indeed, the Bayesian statistics we conducted suggested that the data do not provide sufficient support for the existence of a correlation, nor against it as there is insufficient evidence for either hypotheses (an absence or a presence of a significant correlation). It is therefore essential to consider the limitations of the study, such as sample size and potential methodological constraints, which could affect the ability to detect an existing relationship.

Undoubtedly, a significant limitation of the present research is the sample size. A sample size this small (N=13) significantly impairs the ability to draw valid statistical conclusions, as it may lack the necessary statistical power to detect meaningful differences or associations.

Another potential limitation of this study lies in the decision to present the tasks in the same order for all subjects: first the VWM task, then the Novelty task, and finally the ANS task. As previously explained, maintaining a consistent task order is a useful strategy in research focused on examining individual performance across different tasks, as it helps minimize order effects. However, this choice might have disadvantaged the ANS task. By the time infants reached the final phase of testing, they might have been tired, less cooperative, and more distracted. This could contribute to the lack of effect in the ANS task.

The lack of significance in the ANS task also introduced a limitation we need to acknowledge. The ANS index used is not based on the maximal ratio infants are sensitive to. Interpretations drawn from such results are somewhat arbitrary. Without the ability to establish a ratio at which infants significantly preferred the screen showing numerical change, we cannot determine the acuity of the ANS ability. We used two ANS indices instead: the average performance at different ratios in the ANS task, as well as performance at the 1:2 ratio. But it is important to be aware that these choices were to some extent arbitrary.

5.3 Future Directions for Research

Analyzing the limitations of the present research highlights what needs to be improved in future studies. Most importantly, it is essential to reach a larger sample size to achieve greater statistical power and detect an effect if one exists.

Future research might also consider modifying the stimuli used to investigate ANS, as the current stimuli may not have been effective in detecting the numerical competencies of infants. Indeed, these stimuli were originally developed for adults. The numerical stimuli consisted of images of simple white dots of varying quantities that alternated on a grey screen. These stimuli may not have been particularly engaging for 9-12-month-old infants, potentially leading to reduced attention to or engagement with the stimuli.

Another significant contribution of future research could be the inclusion of a third task to measure subjects' abilities in addition to VWM and ANS. The present study attempted this, but due to an error in the experimental setup, the data obtained from the novelty effect task could not be analyzed.

By addressing these suggestions, future research can build on the current study's findings, overcome its limitations, and contribute to a deeper understanding of cognitive development in early childhood.

5.4 Conclusions

The present study tested the prediction, deriving from previous adult studies (Piazza et al. 2011, Dell'Acqua et al. 2024), that visual working memory capacity and approximate number estimation accuracy do not correlate in 9-12-month-old children, as they constitute distinct cognitive modules. The results make several contributions to a better understanding of this question.

Firstly, it has contributed to our knowledge of VWM capacities in infants aged 9-12 months, suggesting that, VWM span may be limited to approximately two elements at this range. However, the absence of significant results at larger set sizes suggests null findings, potentially due to the limited statistical power of our small sample size to detect subtle effects in more complex set sizes.

Secondly, given the observed lack of correlation between the VWM and ANS measures, this study has contributed to supporting the hypothesis that these two systems are functionally independent from the early stages of life. This finding suggests that although both systems are phylogenetically shared and ontogenetically precocious, and thus crucially important for organisms, they appear to be distinct and independent cognitive modules that do not share the same developmental origin. Nevertheless, Bayesian statistics suggests that there is insufficient evidence for this hypothesis and to draw a definite conclusion.

References

Alloway, T. P., & Passolunghi, M. C. (2011). The relationship between working memory, IQ, and mathematical skills in children. *Learning and Individual Differences*, *21*(1), 133–137. https://doi.org/10.1016/j.lindif.2010.09.013

Benasich, A. A., & Tallal, P. (1996). Auditory temporal processing thresholds, habituation, and recognition memory over the 1st year. *Infant Behavior and Development*, *19*(3), 339–357. https://doi.org/10.1016/S0163-6383(96)90033-8

Berg, C. A., & Sternberg, R. J. (1985). Response to Novelty: Continuity Versus Discontinuity in the Developmental Course of Intelligence. In *Advances in Child Development and Behavior* (Vol. 19, pp. 1–47). Elsevier. https://doi.org/10.1016/S0065-2407(08)60387-0

Brigadoi, S., Cutini, S., Meconi, F., Castellaro, M., Sessa, P., Marangon, M., Bertoldo, A., Jolicœur, P., & Dell'Acqua, R. (2017). On the Role of the Inferior Intraparietal Sulcus in Visual Working Memory for Lateralized Single-feature Objects. *Journal of Cognitive Neuroscience*, *29*(2), 337–351. https://doi.org/10.1162/jocn_a_01042

Colombo, J., Wayne Mitchell, D., Dodd, J., Coldren, J. T., & Horowitz, F. D. (1989). Longitudinal correlates of infant attention in the paired-comparison paradigm. *Intelligence*, *13*(1), 33–42. https://doi.org/10.1016/0160-2896(89)90004-4

De Marco, D., & Cutini, S. (2020). Introducing CUSTOM: A customized, ultraprecise, standardization-oriented, multipurpose algorithm for generating nonsymbolic number stimuli. *Behavior Research Methods*, *52*(4), 1528–1537. https://doi.org/10.3758/s13428-019-01332-z

Decarli, G., Piazza, M., & Izard, V. (2023). Are infants' preferences in the number change detection paradigm driven by sequence patterns? *Infancy*, *28*(2), 206–217. https://doi.org/10.1111/infa.12505

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

Dell'Acqua, R., Sessa, P., Brigadoi, S., Gervain, J., Luria, R., & Doro, M. (2024). On the functional independence of numerical acuity and visual working memory. *Frontiers in Psychology*, *15*, 1335857. https://doi.org/10.3389/fpsyg.2024.1335857

Diamond, A. (1990). The Development and Neural Bases of Memory Functions as Indexed by the AB and Delayed Response Tasks in Human Infants and Infant Monkeysa. *Annals of the New York Academy of Sciences*, *608*(1), 267–317. https://doi.org/10.1111/j.1749-6632.1990.tb48900.x

Fantz, R. L. (1964). Visual Experience in Infants: Decreased Attention to Familiar Patterns Relative to Novel Ones. *Science*, *146*(3644), 668–670. https://doi.org/10.1126/science.146.3644.668

Fukuda, K., Mance, I., & Vogel, E. K. (2015). α Power Modulation and Event-Related Slow Wave Provide Dissociable Correlates of Visual Working Memory. *The Journal of Neuroscience*, *35*(41), 14009–14016. https://doi.org/10.1523/JNEUROSCI.5003-14.2015

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. https://doi.org/10.1038/nature07246

Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences*, *112*(44), 13525–13530. https://doi.org/10.1073/pnas.1515414112

Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychologica*, *145*, 147–155. https://doi.org/10.1016/j.actpsy.2013.11.009

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

Lewis, M., Goldberg, S., & Campbell, H. (1969). A Developmental Study of Information Processing within the First Three Years of Life: Response Decrement to a Redundant Signal.

Monographs of the Society for Research in Child Development, *34*(9), iii–41. https://doi.org/10.2307/1165696

Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy: Number discrimination in infants. *Developmental Science*, *13*(6), 900–906. https://doi.org/10.1111/j.1467-7687.2009.00948.x

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability: Approximate number system and math abilities. *Developmental Science*, *14*(6), 1292–1300. https://doi.org/10.1111/j.1467-7687.2011.01080.x

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), Article 6657. https://doi.org/10.1038/36846

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011a). Impaired Acuity of the Approximate Number System Underlies Mathematical Learning Disability (Dyscalculia): Impaired Numerical Acuity Contributes to MLD. *Child Development*, *82*(4), 1224–1237. https://doi.org/10.1111/j.1467-8624.2011.01608.x

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011b). Preschoolers' Precision of the Approximate Number System Predicts Later School Mathematics Performance. *PLoS ONE*, *6*(9), e23749. https://doi.org/10.1371/journal.pone.0023749

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

Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences*, *14*(12), 542–551. https://doi.org/10.1016/j.tics.2010.09.008

Piazza, M., Fumarola, A., Chinello, A., & Melcher, D. (2011). Subitizing reflects visuo-spatial object individuation capacity. *Cognition*, *121*(1), 147–153. https://doi.org/10.1016/j.cognition.2011.05.007

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

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does Subitizing Reflect Numerical Estimation? *Psychological Science*, *19*(6), 607–614. https://doi.org/10.1111/j.1467-9280.2008.02130.x

Roitman, J., & Brannon, E. (2003). Nonverbal Representations of Time and Number in Animals and Human Infants. In W. Meck (Ed.), *Functional and Neural Mechanisms of Interval Timing* (Vol. 19). CRC Press. https://doi.org/10.1201/9780203009574.ch6

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The Development of Visual Short-Term Memory Capacity in Infants. *Child Development*, *74*(6), 1807–1822. https://doi.org/10.1046/j.1467-8624.2003.00639.x

Rouder, J. N., Morey, R. D., Morey, C. C., & Cowan, N. (2011). How to measure working memory capacity in the change detection paradigm. *Psychonomic Bulletin & Review*, *18*(2), 324–330. https://doi.org/10.3758/s13423-011-0055-3

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007

Simmering, V. R., & Perone, S. (2013). Working Memory Capacity as a Dynamic Process. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00567

Theeuwes, J., Kramer, A. F., & Irwin, D. E. (2011). Attention on our mind: The role of spatial attention in visual working memory. *Acta Psychologica*, *137*(2), 248–251. https://doi.org/10.1016/j.actpsy.2010.06.011

Trick, L., & Pylyshyn, Z. (1994). Trick, L. M. & Pylyshyn, Z. W. Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. Psychol. Rev. 101, 80-102. *Psychological Review*, *101*, 80–102. https://doi.org/10.1037/0033-295X.101.1.80

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. https://doi.org/10.1016/j.cogpsych.2014.01.003

Vanbecelaere, S., Matsuyama, K., Reynvoet, B., & Depaepe, F. (2021). The Role of the Home Learning Environment on Early Cognitive and Non-Cognitive Outcomes in Math and Reading. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.746296

Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, *438*(7067), 500–503. https://doi.org/10.1038/nature04171

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., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*(1), 88–101. https://doi.org/10.1111/j.1467-7687.2005.00395.x

Supplementary Materials

1. Informed consent modules

Modulo

MODULO CONSENSO INFORMATO PER RICERCHE CHE PREVEDONO LA SOMMINISTRAZIONE SU PARTECIPANTI MINORENNI

MODULO INFORMATIVO E DI CONSENSO ALLA PARTECIPAZIONE E AL TRATTAMENTO DEI DATI

DESCRIZIONE E SCOPI DELLA RICERCA

Gentili genitrici o genitori,

con il presente documento Vi chiediamo di fornire il Vostro consenso informato alla partecipazione di Vostro figlio/a alla ricerca “Abilità numeriche e di memoria” coordinata da Prof.ssa Judit Gervain. L’obiettivo della ricerca è quello di studiare le radici e i limiti di queste due abilità nei bambini.

METODOLOGIA DI RICERCA

Durante la ricerca Vostro figlio/a sarà seduto/a in braccio a voi in una cabina dove è posizionato uno schermo dal quale compariranno alcune immagini. Misureremo quindi se il vostro bambino / la vostra bambina passerà più tempo a guardare alcune piuttosto che altre, per determinare se riconosce o meno la presenza di cambiamenti tra queste, e quindi se distingue numeri ed oggetti diversi.

In dettaglio, verranno utilizzati i seguenti strumenti: HPP (Head-turn Preference procedure) è una tecnica non invasiva per misurare il tempo speso dal bambino rivolto verso uno schermo mentre è riprodotto un particolare stimolo visivo (in questo caso dei pallini, dei quadratini o delle facce). Nella cabina sperimentale saranno presenti: uno schermo e una telecamera per registrare i movimenti del bambino. Lo sperimentare, fuori la cabina, potrà seguire lo sguardo del bambino. I compiti e le immagini proposti sono semplici, piacevoli e non inducono stress.

LUOGO E DURATA DELLA RICERCA

La ricerca sarà svolta presso il Dipartimento di Psicologia dello Sviluppo e della Socializzazione e avrà una durata complessiva di circa 30 minuti considerando 5-10 minuti dentro la cabina sperimentale e 15-20 minuti per le spiegazioni e la preparazione.

La ricerca sarà svolta previa autorizzazione del Comitato Etico Della Ricerca Psicologica.

RECAPITI

- Responsabile della ricerca: Judit Gervain, professore ordinario, Dipartimento di Psicologia dello Sviluppo e della Socializzazione, Università degli studi di Padova

- Responsabile della raccolta dati:
Vittoria Volpi, studentessa, Università degli studi di Padova, vittoria.volpi@studenti.unipd.it

CONSENSO ALLA PARTECIPAZIONE E AL TRATTAMENTO DEI DATI

Le/I sottoscritte/i acconsentono liberamente alla partecipazione del proprio figlio/a allo studio "Abilità numeriche e di memoria".

Le/I sottoscritte/i dichiarano:

1. Di essere a conoscenza che lo studio è in linea con le vigenti leggi D. Lgs 196/2003 e UE GDPR 679/2016 sulla protezione dei dati e di acconsentire al trattamento e alla comunicazione dei dati personali, nei limiti, per le finalità e per la durata precisati dalle vigenti leggi (D. Lgs 196/2003 e UE GDPR 679/2016). Il/la responsabile della ricerca si impegna ad adempiere agli obblighi previsti dalla normativa vigente in termini di raccolta, trattamento e conservazione di dati sensibili.
2. Di sapere di poter ritirare la partecipazione della propria figlia/o dal presente studio in qualunque momento, senza fornire spiegazioni, senza alcuna penalizzazione e ottenendo il non utilizzo dei Suoi dati. Qualora ila genitrice o il genitore non fosse presente in sede di raccolta dati, potrà delegare altre figure (es., tutor, insegnanti, ricercatori, ecc.) ad interrompere la ricerca qualora queste lo ritenessero opportuno.
3. Di essere a conoscenza che i dati saranno raccolti in forma confidenziale (nome/codice).
4. Di essere a conoscenza che i dati della propria figlia/o saranno utilizzati esclusivamente per scopi scientifici e statistici e con il mantenimento delle regole relative alla riservatezza.
5. Di sapere che, qualora lo desiderassero, possono ottenere la restituzione dei dati grezzi. Poiché il presente studio non ha finalità cliniche, siamo stati informati che dovremo rivolgerci ad una/o specialista per l'eventuale interpretazione dei dati.
6. Di sapere che una copia del presente modulo ci sarà consegnata dalla ricercatrice o dal ricercatore.
7. Di acconsentire alla videoregistrazione e/o audioregistrazione.

Le/I sottoscritte/i
COGNOME E NOME (IN STAMPATELLO) ________________

COGNOME E NOME (IN STAMPATELLO) ________________

Genitori/genitrici di

COGNOME E NOME (IN STAMPATELLO) ________________

presa visione del presente modulo esprimono il proprio consenso al trattamento dei dati personali e alla partecipazione alla ricerca “Abilità numeriche e di memoria” della propria figlia/o.

Data ________________________

Firma leggibile ____

Firma leggibile ____

Si informano inoltre le figlie o i figli che i loro genitori/genitrici potrebbero avere accesso ai loro dati.

Firma leggibile (per conoscenza) _____________________________________ (minore > 12 anni)