UNIVERSITA' DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Psicologia dello Sviluppo e della Socializzazione

SCUOLA DI DOTTORATO DI RICERCA IN: SCIENZE PSICOLOGICHE

INDIRIZZO: SCIENZE COGNITIVE

CICLO: XX

Varieties of Numerical Representations

Direttore della Scuola: Ch.mo Prof. Luciano Stegagno

Supervisore: Ch.mo Prof. Marco Zorzi

Dottorando: Ilaria A. Berteletti

DATA CONSEGNA TESI

31 gennaio 2008

“Do not worry about your difficulties in Mathematics.
I can assure you mine are still greater.”

(A. Einstein, 1879-1955)

I would like to thank all those that have supported me in these three years and in particular …

My supervisor Marco Zorzi, for welcoming me in his lab, for his constant support and his scientific advices throughout my diverse projects.

Ed Hubbard, who has been a source of inspiration in the colourful world of synaesthesia. Of great value have been his constructive and very generous support, advice, encouragements and supervision.

Daniela Lucangeli, for making the developmental projects a reality and for her expert advices in Educational Sciences.

Stanislas Dehaene and Manuela Piazza from INSERM Paris, that have, directly and indirectly, brought great inspiration into my work.

Amongst those that have been encouraging and advising me throughout this period I shall specially thank Konstantinos Priftis and Peter Kramer, both have been models of scientific rigour.

A project in psychology is difficult to conceive without participants, therefore I thank all those, adults and children, that have participated in these studies but in particular NM that always undertook my numerous synaesthesia experiments with enthusiasm.

Anna Fertonani, for being a promising student and simply great to work with.

A special thank you to the lab technicians, Luca and Mauro, that were there for all my computer and technical emergencies; as well as Sandro Bettella, a real wizard with E-Prime.

All those with whom I shared hours of work, coffees, scientific (and non) discussions, harsh moments, but also and most, great after-hours and friendship! Just to name some: Marcella, Maria Grazia, Fabiola, Mario, Matteo, Giuliano, Marco, Pietro, Simone …

The NUMBRA group, teachers and students, with whom I shared great scientific moments that always had the power to renew my interest in cognitive sciences and generate new ideas of research. In particular, Petra, for sharing the rooms everywhere around the world and for her friendship.

My parents and family, that have always been present for help and support and, in spite of our diverging views, have respected and encourage my choices.

Last but not least, Daniele, for enduring three thesis, my moodiness, my trips around the world and for always and without fail being there.

CONTENTS

ENGLISH SUMMARY

A growing amount of evidence supports the hypothesis that humans are able, from the earliest age, to process numerical information in the absence of language. This work addresses the question of the nature of the internal representation for processing numerosities from three perspective: developmental, adults' skilled performance, and the peculiar case of synaesthesia.

In our studies with children we addressed the development of the mental representation for numbers. In the first experiment we showed that, before formal teaching, preschoolers possess multiple numerical representations that follow a specific developmental trend. Indeed, they first rely on an intuitive representation where numbers are distributed logarithmically and progressively, with numerical practice and increasing knowledge, they shift to a formal and linear representation. Moreover, preschool children can exhibit both types of representations according to the familiarity with the context.

In the second study, we tested the hypothesis that non-numerical sequences may also rely on a similar representation and follow the same developmental pattern. By studying children from the last year of kindergarten to 3rd grade we observed that numerical and non-numerical sequences have different mental representations. Indeed, only the numerical sequence shows the classical effects that support the hypothesis of a logarithmic representation. Moreover, we observed that children start to learn linearity in the numerical domain and then generalize the principle to all ordinal sequences.

In our third study we investigated adults numerical representation of symbolic and non symbolic material. The aim was to test if the basic ability of discriminating between numerosities could explain higher level processes such as approximate calculation and symbolic number comparison. Indeed, if the preverbal approximate system of the numerical representation forms the basis of more complex numerical and mathematical knowledge, it should influence performance in other numerical tasks. Moreover, the crossing of symbolic and non-symbolic format of the stimuli for the approximate calculation task allowed us to qualify previous findings about the *operational momentum* effect in approximate arithmetic (i.e., the tendency to overestimate additions and underestimate subtractions). Indeed, we observed that the effect may be explained by the tendency to underestimate numerosities and that this bias is proportional to the set size.

In the last experiment we investigated the relation between colour and numerical representation in NM, a number-colour synaesthete. Results showed that, in spite of not reporting colours for numerosities, our synaesthete was subject to interference effects. From these results we suggest a new model that accounts for the implicit and explicit synaesthetic effects by suggesting the existence of primary and secondary synaesthetic connections ("pseudo-synaesthesia"). Our results and model questions previous work on bi-directional effects and the operational definition of synaesthesia.

ITALIAN SUMMARY

Un numero crescente di studi dimostra che gli esseri umani sono in grado di processare informazioni numeriche fin dai primi giorni di vita e molto prima dell’acquisizione del linguaggio. Questa tesi si propone di indagare la rappresentazione mentale dell’informazione numerica tramite tre approcci distinti: i) lo studio dello sviluppo cognitivo in bambini di età scolare e prescolare, ii) lo studio delle abilità di soggetti adulti normodotati, iii) lo studio di un particolare caso singolo di sinestesia.

Il primo studio sui bambini ha indagato lo sviluppo della rappresentazione mentale dei numeri. I risultati mostrano che, ancor prima dell’inizio dell’educazione formale, i bambini della scuola per l’infanzia possiedono molteplici rappresentazioni numeriche che si sviluppano seguendo uno specifico percorso. I bambini utilizzano inizialmente una rappresentazione intuitiva dove i numeri sono distribuiti in modo logaritmico e progressivamente, con l’apprendimento e la pratica con i numeri, viene sostituita da una rappresentazione formale basata su una distribuzione lineare. Inoltre, i bambini in età prescolare possono utilizzare entrambe le rappresentazioni a seconda della familiarità con il contesto.

Nel secondo studio è stata testata l’ipotesi che sequenze non numeriche possano anch’esse far riferimento ad una rappresentazione simile a quella numerica, sviluppandosi secondo lo stesso percorso. I risultati, ottenuti con bambini dall’ultimo anno della scuola per l’infanzia fino alla terza classe della scuola primaria, hanno mostrato che le sequenze numeriche e non numeriche hanno rappresentazioni mentali distinte. Infatti, solo la sequenza numerica mostra gli effetti classici che sostengono l’ipotesi di una rappresentazione logaritmica. Inoltre, si osserva che nei bambini il concetto di linearità nasce inizialmente in ambito numerico, e successivamente si generalizza a tutte le sequenze ordinate.

Nello studio con gli adulti è stata indagata la rappresentazione di materiale numerico simbolico e non simbolico. Lo scopo era di verificare se la capacità di discriminazione tra numerosità potesse spiegare competenze di più alto livello quali il calcolo approssimato e il confronto di numeri arabi. Infatti, se il sistema di rappresentazione numerica preverbale ed approssimativo costituisce le basi per l’apprendimento di concetti numerici e matematici più complessi esso dovrebbe influenzare le prestazioni osservate in altri compiti di tipo numerico. Inoltre, la presentazione di materiale sia simbolico che non simbolico nel compito di calcolo approssimativo ha permesso di ridefinire il fenomeno dell’ “*operational momentum*”

nell’aritmetica approssimata (ovvero la tendenza a sovrastimare il risultato di una addizione e a sottostimare il risultato di una sottrazione). Infatti, si è osservato che l’effetto può essere spiegato dalla naturale tendenza a sottostimare la numerosità di un insieme e che questo errore di stima è proporzionale alla grandezza numerica dell’insieme.

Infine, nell’ultimo studio si è investigata la relazione tra colore e rappresentazione numerica in NM, un sinesteta numero-colore. I risultati hanno mostrato un effetto interferenza per insiemi di pallini malgrado NM non riportasse esplicitamente nessun colore per le numerosità. A partire da questo risultato è stato elaborato un nuovo modello che spiegherebbe gli effetti impliciti ed espliciti osservati in studi sulla sinestesia. In questo modello si presenta l’ipotesi di connessioni sinestetiche di primo e secondo grado (“pseudo-sinestesia”). I risultati ed il modello pongono nuove domande sia sulla reale bi-direzionalità dell’interferenza sinestetica che sulla formulazione di una definizione operativa del fenomeno.

Chapter 1: General Introduction

Ⅰ: **Numbers, Numerosities and Numerical Processing.**

What is a number? What is the information that numbers like “1, 2, 3…” convey? What information does our brain process to judge the size of two numbers or of two sets of objects? More generally, the question that has been occupying researchers in numerical cognition is to understand what underpins humans’ ability of estimating, counting and manipulating numerical information. Indeed the field of research is wide and it would be far from possible to answer all these questions in this work but yet some questions will be addressed, others will be answered and even more will be asked. Given that without questions there would be no science and research.

In this first chapter, the core theoretical knowledge of numerical representation will be outlined. First, the principal models of numerical cognition will be described and we will focus on those main effects and psychophysical laws that characterize numerical processing. Then two processes that are known to rely uniquely on numerical representations will be described: subitizing and estimation.

In the second section of this chapter, the developmental perspective of numerical cognition will be reviewed. First infants abilities with small and large numerosities will be discussed and will be followed by the description of preschoolers’ performance with both symbolic and non symbolic material. Finally, the third and last section of this chapter will be dedicated to adults behaviour mainly in non symbolic numerical tasks.

I.1. Models of numerosity processing.

Several models have been suggested to account for the ability to process numerical information. However only two models, that we will consider in detail, integrate symbolic and non-symbolic processing of numerical quantities and, even more importantly, consider that the ability of manipulating numerosities[1] is present prior to language and is shared with non-human species. This assumption is critical since it allows to investigate the abilities of both infants and children that have not yet started formal schooling.

Besides these models of approximate numerical representation, other models postulate an exact numerical representation (Verguts, Fias, & Stevens, 2005; Verguts & Fias, 2004; Zorzi & Butterworth, 1999; Zorzi, Stoianov, & Umiltà, 2005; Zorzi, Stoianov, Becker, Umilta, & Butterworth, 2008). The numerosity code model assumes a linear representation of numerosity, in a similar way to a "thermometer" representation (Zorzi et al., 2008). Each numerosity set is represented by a corresponding number of nodes which contains the smaller sub-sets. This model successfully explains the distance and size effect in number comparison as well as correctly simulating the distance-priming effect which is not explained by non linear numerical representations. Moreover, the authors suggest that this model should coexist with an approximate representation since cardinality is not the only type of mental representation of numbers. In the following section, only approximate representations will be described in details since they constitute the basements of the present work.

I.1.1. The Preverbal Counting Model.

The model proposed by Gallistel and Gelman (1992) is an extension of the animal model proposed by Meck and Church (1983). According to the authors, numbers would be represented in an analogical manner thanks to what they call magnitudes. In fact, this model postulates that all numerical judgments rely on a serial process that accumulates an equal amount of activation for each unit to be counted into what could be described as a container.

The comparison between two numbers or two numerosities would consist in comparing two amounts of activation. For humans, the acquisition of number words would allow the creation of a bidirectional mapping with magnitudes[2]. Each numerosity would be stored in

[1] Numerosity defines the quantity represented by a given number or the numerical value of a set of items (i.e., its cardinality). Thus it refers to a quantity that may be numerically determined.

[2] The magnitude corresponds to the activation of a numerical quantity on an internal and analogical representation.

memory and compared to the actual magnitude activated. Through this process, a number (written or spoken) would be mapped into a magnitude and vice-versa (Figure I.1.).

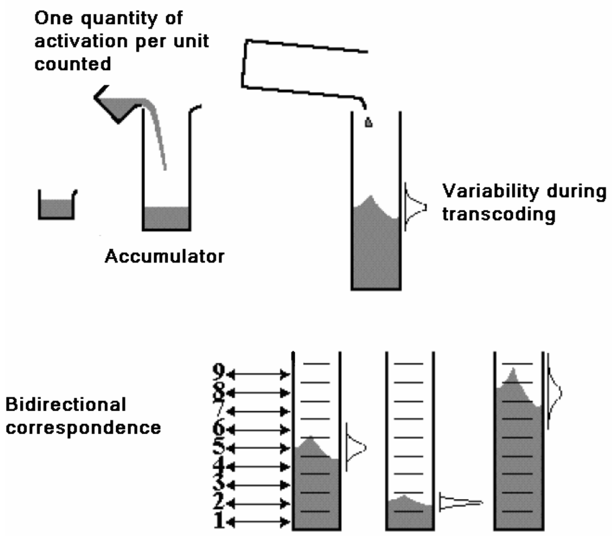

Figure I.1.: The preverbal counting model. Items are serially counted by the accumulator, that is, quantities are incremented one by one, as a cup would be poured into a graduated recipient; the result of the count is read out in memory where it has been stored, but memory is noisy and therefore leads to different estimates of the number counts on different occasions. The amount of noise in memory is proportional to the numerical quantity being counted (scalar variability: the variability in the estimates is proportional to the mean of the distribution of estimates). Reproduced from (Gallistel & Gelman, 2000).

An important axiom of the model is that the mapping between the codes (verbal and internal) would be subject to variability. The preverbal counting mechanism would lead to increasingly imprecise estimates as numerosity increases. As more and more counts are accumulated, there would be more and more chances of error keeping the exact count. Therefore, in this model, numerosities would be represented on a linear scale (same quantity added for each additional item counted and same distance between two neighbouring numerosities) but representations would get increasingly fuzzier due to increase in noise in the translation process. This would thus lead to increasing overlap of representations as numerosities increase.

I.1.2. The Triple Code and the Log-Gaussian model.

The Log-Gaussian model has been presented in detail by Dehaene and Changeux (1993) and is part of a broader model of numerical cognition called the Triple Code Model (Dehaene, 1992; Dehaene & Cohen, 1995). The Log-Gaussian code is what the authors consider the core component of numerical comprehension. The other components that constitute the Triple Code allow to translate from the symbolic format (number words and digits) to the non symbolic format and therefore access the semantic of numbers. These components are also responsible of a number of other tasks such as counting, parity judgment, calculation etc.

Briefly, the Triple Code Model is composed, as its name suggests, by three main components, each based on a distinct representation (or code) that is used for input and output (Figure I.2.). According to this model, therefore, the number of codes allowed to process numerical information are limited to three. Moreover, each component is tied to specific input and output procedures allowing the peripheral processing of each code separately. These procedures, although specific to each component, would be shared with other cognitive functions as for example reading and writing.

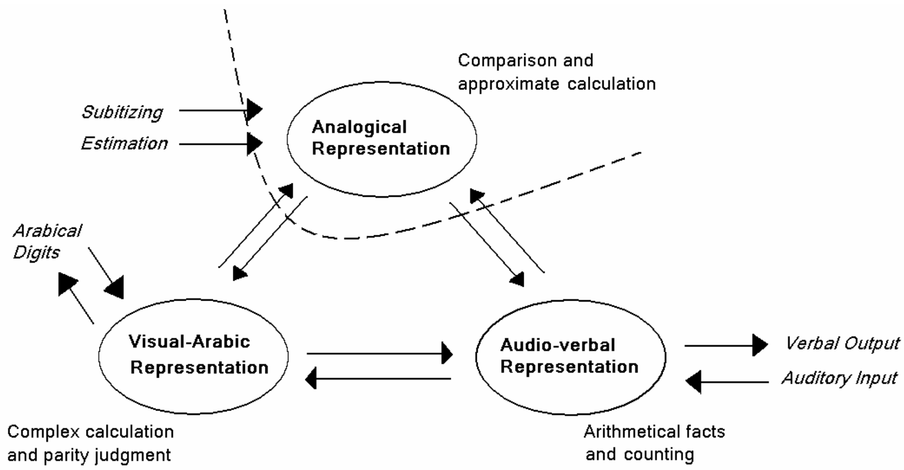

Figure I.2.: The Triple Code Model: each representation is associated to input and output processes and is also responsible for specific tasks (adapted from Dehaene, 1992).

The first component is the Visual-Arabic format which allows to represent numerical notations as sequence of digits in a visuo-spatial internal space. To this component, the main associated tasks would be parity judgment and complex calculation. The second component is the audio-verbal representation. In this system, numbers are represented as a sequence of words syntactically organized and would be similar to verbal representations. This component would therefore be responsible for the ability of counting and reciting simple arithmetical problems (arithmetical facts). The last component, the semantic component is the analogical representation of numerical quantities. This component is conceptualised as a mental number line (Dehaene, 1992).

The Log-Gaussian Model is a model of the mental number line that postulates a parallel numerosity mechanism and a compressive internal representation with fixed (Gaussian) noise. The model proposes that numerosity is represented on a log scale, explaining therefore that smaller numerosities are represented in a more precise way, whereas larger ones are more compressed and therefore less easy to discriminate. Two main differences distinguish the Log-Gaussian Model from the Preverbal Counting Model: first, the numerosity detection is parallel in the former and serial in the latter, second the noise is an intrinsic feature of the representation in the former, whereas it is the result of the accumulation process in the latter.

For the Log-Gaussian Model, each activated numerosity would yield a fixed Gaussian activation on the line and the logarithmic compression would account for the overlap with increasing numerosities (Figure I.3.). Therefore, for a given numerical distance, two large numbers would show a larger overlap of activation compared to two smaller numbers (i.e. 1 and 4 compared to 4 and 8).

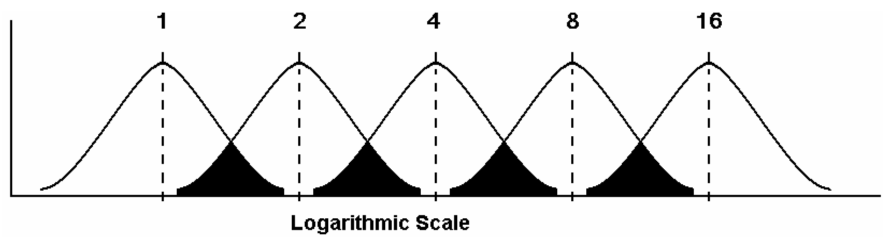

Figure I.3.: Representation of the Log-Guassian Model. Numerosities are represented on an internal logarithmic scale with fixed Gaussian noise. Larger numerosities are closer on the line therefore each Gaussian distribution overlaps more accounting for the decreasing discriminability.

The width of the Gaussian distribution determines the precision of the underlying representation. Indeed, the *internal Weber Fraction* is used to refer to the Gaussian's width which may be different for each single individual and might explain interpersonal differences. A more precise discrimination between numerosities would be described by a smaller width of the Gaussian (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004). In the following section the *internal Weber Fraction* will be discussed in more details.

I.2. Psychophysical laws of numerosity and the Weber fraction.

Before introducing the psychophysical laws of numerical processing, it is essential to describe a couple of very robust effects that have been found in numerical cognition. These are very important since any model of numerical cognition should be able to explain them and fit the behavioural data.

I.2.1. Distance and Size effects.

In adult research two main effects have shaped models on numerical cognition. In a number comparison task, the *distance effect* corresponds to the decreasing response time and increasing accuracy as the distance between the numbers increases. Moyer and Landauer (1967) were the first to observe this effect when presenting single digits and they showed that it followed the same psychophysical laws as perceptual comparison. Indeed, the same effect was previously described by Johnson (1939) in a perceptual task where participants were required to compare the length of bars. This result is very important since it shows that participants had to access a semantic representation to make the comparison as two numbers are not more likely to be physically different just because they are numerically farther apart. This also suggested that the semantic representation of numbers is analogical. After Moyer and Landauer, many other studies have confirmed the effect extending it to other material as for instance double digits (Dehaene, Dupoux, & Mehler, 1990) and sets of dots (Buckley & Gillman, 1974). Even more intriguing is the observation that this effect appears even when numerical information is irrelevant to the task (Dehaene & Akhavein, 1995) suggesting that the semantic activation of numbers is automatic. Automaticity is also revealed by other effects, discussed further in this chapter, such as the SNARC effect (Spatial Association of Response Codes; Dehaene, Bossini, & Giraux, 1993) and the numerical Stroop effect (Girelli, Lucangeli, & Butterworth, 2000; Mussolin & Noel, 2007).

The *size effect* is also typically found in comparisons tasks. This is the observation that for a given numerical distance, reaction time increases and accuracy decreases when numbers are larger (Banks, Milton, & Fortunee, 1976; Dehaene, 1989; Moyer & Landauer, 1967). For example it will take more time to discriminate between 8 and 9 than between 1 and 2. This result is critical because it suggest that not only the exact distance between numbers but also their size strongly influences the discrimination process. Moreover, Buckley and Gillman (Buckley & Gillman, 1974) have compared the performance on symbolic (digits) and non-symbolic comparisons of numerosities from 1 to 9 and have shown that reaction times where extremely similar, again suggesting an analogical representation of numerical magnitude.

As it already appears, both effects are explained by the models presented in the previous section (I.1.1.). In the Preverbal Counting Model, it is the scalar noise in the mapping process that accounts for the increasing overlap between numerosities, whereas in the Log-Gaussian Model it is the nature of the numerical representation itself that accounts for the observed effects.

I.2.2. The Law of Proportionality and the Weber fraction.

The above mentioned effects, *distance and size*, have been interpreted as the signature of an analogical representation of numerical quantities and as being the result of a same principle: the Law of Proportionality. This law states that the confusion between two numbers or numerosities is related to the proportion between the two values. Indeed, the more the ratio is close to 1 the greater the confusion. For instance, the degree of confusion may be calculated using percentage of accuracy. Therefore accuracy will drop when the ratio approaches 1.

This law explains the *distance effect* since the ratio approaches 1 when the distance between two numbers decreases. Lets take the following pairs: (4,10) and (4,8). Indeed, the ratio of the first pair is larger than the second (10/4 = 2.5 and 8/4 = 2). Moreover, this law also explains the *size effect* since, when the numerical distance is kept constant, the ratio also approaches 1 with larger numerical values. For instance, the ratio of pairs (20,30) and (120,130) is closer to 1 for the larger pair (1,5 and 1,08 respectively) although the numerical distance is constant.

The Weber Law is closely related to the Law of Proportionality. Indeed, the *just noticeable difference*[3] between two stimulations is lawfully related to the initial stimulus magnitude. In other words, the law states that the *just noticeable difference* is a constant proportion of the original stimulus value (ΔIntensity/Intensity of reference = constant). This law is applied to various sensory features (brightness, loudness, mass, line length, etc.) and numerosity is also one of them.

The *behavioural Weber fraction* is therefore the value of the constant in the equation allowing to estimate the performance in situations where the ratio between two sensory experiences may be calculated[4]. In the Log-Gaussian model, where numerosities follow a logarithmic distribution and the Gaussian activation of any given numerosity is constant, the activation overlap between two numerosities accounts for the ratio effect. Therefore, by means of a mathematical transformation of the *behavioural Weber fraction*, it is possible to formalize the *internal Weber fraction* (Izard & Dehaene, 2007; Piazza et al., 2004). The latter would therefore correspond to the width of the Gaussian activation on the internal numerical representation. It would indicate the minimal distance on the numerical representation for two numerosities to be considered as different in 75% of the cases. Therefore, it determines the precision in subjects' performance: a smaller Gaussian width corresponds to less overlap in numerosity representation and thus a more precise discrimination ability.

I.3. Subitizing and numerical estimation.

In this section we will discuss two numerical processes that should only rely on the internal numerical representation. The first, called subitizing, is the ability to enumerate extremely quickly and accurately sets of 1 to 3, maximum 4, visually presented objects. This phenomenon is classically observed when participants are asked to enumerate as fast and accurately as possible sets of dots ranging from 1 to 10. Reaction times are generally flat and error rates low for numerosities 1 to 3 and both indices of performance progressively increase

[3] The just noticeable difference is the minimum amount by which stimulus intensity must be changed in order to produce a noticeable variation in sensory experience.

[4] For sake of simplicity we will only discuss the *Weber Law* in the field of numerical cognition throughout the present work.

for larger sets. In the small range, the increase in reaction time is usually about 50ms whereas above the 1-4 range the increase in reaction time is about of 200ms for each additional item (Trick & Pylyshyn, 1994; Mandler, 1982; Atkinson, Campbell, & Francis, 1976). Beyond the subitizing range, if accuracy is high and reaction times slow, the process involved is considered to be counting.

The second process is observed when a participant is required to enumerate rapidly a large amount of stimuli without using the counting strategy. As numerosity increases, numerical estimation judgments move away form the real value. This characteristic is called *scalar variability* and is typical of estimation processes (Gallistel & Gelman, 1992; Izard & Dehaene, 2007; Whalen, Gallistel, & Gelman, 1999).

Both models described in section I.1.1. may account for these phenomena. The Preverbal Counting Model, with a *scalar variability* in the mapping to memory of the activated magnitude predicts exact responses for the small range and approximate ones for larger numerosities. Indeed, the noise in the mapping grows with numerosity and is negligible for numerosities up to 3 and then progressively increases. For the Log-Gaussian Model, it is the compressed nature of the internal number line that determines the uncertainty for large numerosities but exact enumeration for small ones. In fact, for small numerosities the activation does not overlap with adjacent values whereas the overlap increases with larger numerosities (see figure I.3.).

Interestingly, Izard and Dehaene (2007), have observed that participants tend to underestimate very large numerosities when no feedback is given. In their study, they presented sets of up to 100 dots for a duration of 150ms. Participants where asked to carry out several estimations for each numerosity in order to calculate mean response and deviations for each. The first observation was that the estimation increased with numerical size of the set even though answers were always highly underestimated. Furthermore, responses followed scalar variability which means that their coefficient of variation (standard deviation/mean response) was independent from the numerosity to estimate.

II：Numerical Representations in Childhood.

In section I.1., two important models of numerical cognition have been described and both assume that numerical abilities are present before language acquisition. In this section will review developmental studies of numerical cognition that give support to either model along with the main controversies in the field. The first part, describes a number of studies highlighting infants' numerical abilities and the second part shows how these abilities and the underlying numerical representation evolve in children before and after the first years of primary school. I will also briefly discuss the acquisition of number symbols, as digits and as written or spoken words.

II.1. Numerical Processing in Infants.

Before understanding how a child learns to count and calculate, it is interesting to understand how the numerical representation evolves and on what initial processes these abilities rely. It is undeniable that numerical processing in adulthood becomes completely automatic. To understand this development, several researchers have studied children's and infants' abilities to solve numerical tasks.

It is important to bear in mind that the studies and results exposed hereafter are still object of controversy and have only given clues to the innate skills of infants. Indeed, a first object of debate is the nature of the underlying representation. Some support the assumption that numerical processing is innate and specific, others postulate that there is a general mechanism for magnitude processing not specific to numerosities but also shared by other quantifiable dimensions as, for example, time and space.

Research on numerical abilities in infancy is also divided according to the numerosity tested during the experiment. A first stream of research has focused on infants ability to discriminate between small sets of objects (up to 3), whereas a second group of studies has used large numerosities.

Studying infants also involves the use of a specific methodology. The most common paradigms are habituation and violations of expectations. The former consists in repeating the same stimulation, for example an image, until the child's interest drops and only then a new stimulus is presented. The basic assumption is that babies are attracted by novelty. Therefore, the baby will lose interest for repetitive stimulation of the same information and if she regains attention it indicates that she has noticed the change of information in the novel stimulus. The other paradigm, violation of expectations, tests the ability of infants to make assumptions about the world presenting situations where implicit rules are violated. If the baby has expectation about the consequence of an event she will be aroused if the situation is inconsistent to her expectations.

II.1.1. Processing of small numerosities.

Among the first to study infants numerical ability, Starkey and Cooper (Starkey & Cooper, 1980), have shown that, using the habituation paradigm, 4-months-old infants were able to discriminate between two and three dots in a line. This result was replicated for both older and younger infants. Antell and Keating (1983) tested infants a few days old and Strauss and Curtis (Strauss & Curtis, 1981) tested infants ranging from 10 to 12 months of age. This finding was also replicated with paradigms where stimuli were presented sequentially or in movement (Canfield & Smith, 1996; Wynn, 1996; van Loosbroek & Smitsman, 1990). For instance, Wynn presented dynamic displays where 6-months-old infants were familiarized to puppets jumping 2 or 3 times and were then tested with both numerosities. Results showed that infants were looking longer at the testing sequence when the number of jumps differed form the habituation sequence (Figure I.4.). The authors concluded that numerical ability is not limited to certain entities as physical objects but could be general and abstract. In this case, infants could enumerate actions.

Bijeljac-Babic and colleagues (Bijeljac-Babic, Bertoncini, & Mehler, 1993) have shown that syllables were among the entities that could be enumerated whereas others have shown intermodal numerical correspondence (Starkey, Spelke, & Gelman, 1983, 1990; Moore, Benenson, Reznick, Peretson, & Kagan, 1987). Starkey and collaborators (1990) presented to 6 to 8-months-old infants two images, one containing two objects and the other three. After about a second, a sound with a sequence of either two or three tones was played. The result was that infants preferred looking at the image with the same amount of objects as sounds heard.

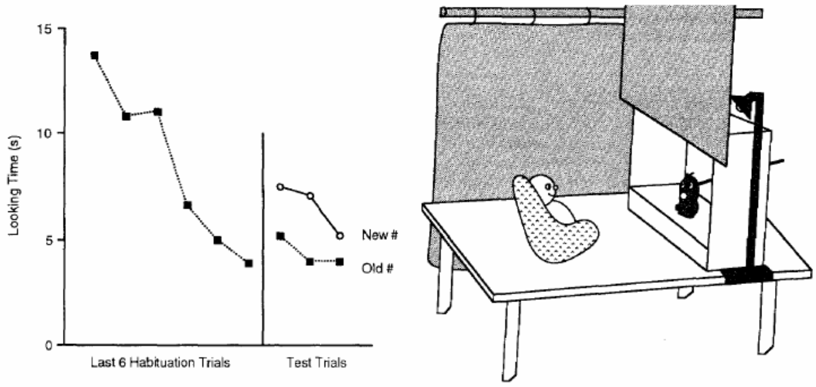

Figure I.4.: Experimental setting and results form an habituation paradigm (Wynn, 1996). On the left side of the figure is a classical response curve for looking time during habituation and during test trial. New stimuli in test trials arouse looking time compared to old ones. On the right side is a prototypical experimental setting for testing infants (reproduced from Wynn, 1996).

Finally, beyond these discrimination abilities, 5-months-old infants seem to posses basic calculation abilities. Koechlin, Dehaene and Mehler (1997) have used the violation of expectations paradigm to test infants predictions on addition and subtraction. Infants watched a little theatre where one or two puppets were presented. A curtain dropped to hide the set and a visible hand introduced or removed a puppet from behind the curtain. When the curtain was pulled away, the set could either be congruent or incongruent with expectations (1 puppet + 1 puppet = 2 puppets or 1 puppet). When the set mismatched expectations, infants looked longer as if they could not understand what happened. It has to be noted that spatial location was also controlled.

Unfortunately, although these results are very attractive, Mix, Huttnelocher and Levine (2002) criticised many studies for important methodological flaws. Indeed, they showed that perceptual properties highly co-varied with numerosity and suggested that a general quantification process could yield the same results as a specific numerical module. What is therefore the real cue used by infants for discriminating between sets? A number or researchers tested the hypothesis of a perceptual sensitivity instead of a numerical ability (Clearfield & Mix, 1999, 2001; Feigenson, Carey, & Spelke, 2002).

Feigenson, Carey and Spelke (2002) tested the perceptual hypothesis and controlled for surface area and number. Indeed, they found that when surface area and number were not correlated, infants dishabituated to area rather than to number (Figure I.5). The authors concluded that infants lack sensitivity to number but are more sensitive to continuous extent dimensions as area.

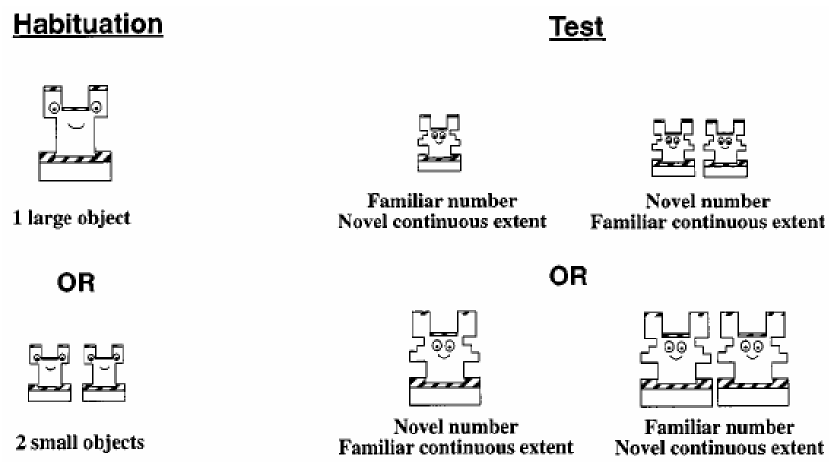

Figure I.5.: Objects used in Feigenson, Carey and Spelke (2002) study in which number was pitted against continuous extent. Here front surface area was controlled between habituation and test trials. In this experimental condition infants failed to dishabituate to number (reproduced from Feigenson et al., 2002).

Clearfield (2004) also replicated Wynn's experiment of dynamic displays but she controlled for rate of jump during habituation and across test trials. With this manipulation, infants failed to show any preference suggesting that rate changes can override infants' response to number. Furthermore, the author tested if the amount of time the puppet spent jumping instead of number was the explanation for infants' behaviour. Indeed, infants' dishabituated to displays that changed in time instead of number.

Finally, to prove intermodal numerical correspondence experiments wrong, Moore and collaborators (Moore et al., 1987) replicated Starkey's studies (Starkey et al., 1983; Starkey et al., 1990) and obtained opposite results. Indeed, their 7-months-old infants preferred looking at the display with a different number of objects. Mix, Levine and Huttenlocher (1997) also replicated Starkey's studies with no success.

The failure to replicate results could be explained by important methodological differences as time between habituation and test and also the length of blocks. Starkey and collaborators argued that it required several trials of practice in a block for infants to start discovering the correspondence between auditory and visual stimuli which were fewer in Moore and al's experiment. Moreover, in Mix et al's experiment trials of different numerosities were randomized, whereas in Starkey's studies several trials kept the same numerosity. This could have distracted infants' ability to figure out the correspondence between modalities.

A recent study presented by Feigenson (2005) seems to give a clear demonstration that infants are able to use numerosity when no other visual cue gives more salient information. In the first experiment 7-months-old were able to compute numerosity when the objects in the array contrasted for colour, pattern and texture, whereas in the second experiment results showed that, with these heterogeneous objects, infants no longer represented array's total continuous extent but relied on number. Feigenson thus concludes that infants extract numerical information when stimuli are heterogeneous and use perceptual cues if objects are identical.

II.1.2. Processing of large numerosities.

Already at 6-months of age, infants show discrimination abilities between large numerosities as long as the ratio is larege enough (Xu & Spelke, 2000). Infants were habituated to images of either 8 or 16 dots which were also controlled for size and dispersion. The test was either of 8 or 16 dots for which size was familiar and total surface was not outside the range seen during habituation. With a ½ ratio infants were able to make discriminations between the 2 numerosities but when the authors replicated the experiment with 8 and 12 dots they failed to discriminate between sets (⅔ ratio).

Lipton and Spelke (2003) showed that infants' ability to discriminate large sets extends to auditory stimuli with the same ½ ratio effect. Moreover, the authors also demonstrated that discrimination improved with age since, at 9-months infants discriminate between sets that have a ⅔ ratio (8 and 12 sounds) but fail when they are comparing sets with a ratio of $^4/_5$ (8 and 10 sounds). This result was recently confirmed by Xu and Arriga with 10-months-old infants and visuo-spatial arrays (2007).

Some research also investigated the principle of ordinality and basic approximate calculation skills for larger numerosities. A study carried out by Brannon (2002) investigated the development of ordinal numerical knowledge. She presented to 9 and 11-months-old infants sequences of stimuli that either increased or decreased in numerosity (Figure I.6.). Only the older group was able to discriminate between the sequences indicating that only by that age infants possess the ability to appreciate the greater than and less than relations between numerical values.

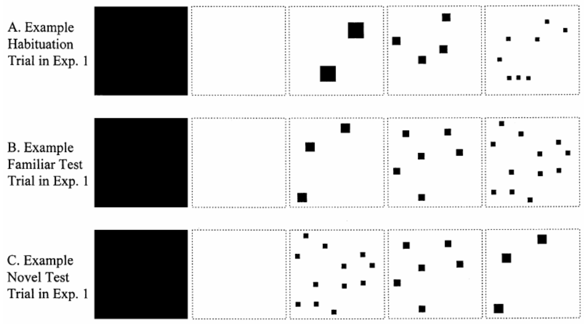

Figure I.6.: Sequential stimuli presented in Brannon's study to investigate ordinal numerical knowledge (reproduced form Brannon, 2002).

McCrink and Wynn (2004) investigated basic calculation skills in infants. Nine months old infants were shown a movie where a number of objects moved and continuously changed dimension and shape until they disappeared behind a patch. Once hidden, either a new set would go add itself to the first one (addition condition) or a subset would leave from behind the patch (subtraction condition). At the end of the movie the patch disappeared leaving the outcome visible. Numerosities employed were only 5 and 10 for both the terms of the problems and the results. Analysis indicated that infants were surprised and looked longer at incorrect outcomes compared to correct ones.

The same authors (McCrink & Wynn, 2007) have recently tested the ability of discriminating ratios in 6-months-old infants. This is, recognising the ratio between two types of visual objects inside the same stimulus. Each habituation stimulus was constituted of blue and yellow dots. Infants were habituated with ratios of ½, ⅓ or ¼ and test stimulus had ratios of ½ and ¼. If the test differed by a factor of 2 (e.g. habituation with ½ ratio and testing on ¼) then infants succeeded, otherwise they failed to recognize a change in ratio (factor of 1.5: habituation with ⅓ ratio and testing on ¼).

II.1.3. Summary of infants abilities.

The research on numerical cognition in infancy has shown that humans are endowed from the earliest days of life of specific numerical skills. Infants seem able to discriminate both small and large sets as long as the ratio between the two is large enough to be perceived. This could explain why some research on small sets has failed to show discrimination abilities between sets of 4 and 6 objects since the ratio was too small for the age tested (Antell & Keating, 1983). Moreover, infants also possess innate expectations of addition and subtraction for both small and large numerosities.

It is probably due to these skills that children learn formal numerical knowledge and later learn to manipulate numerical symbols. This topic is the object of the following section.

II.2. Numerical Processing in Children.

Numerical skills in children are more diverse than infants'. Indeed, children learn progressively to use number words and with instruction they also learn numerical symbols. Moreover, as we have shown in the previous section, numerical discrimination improves with age and therefore also the numerical representation. This development probably continues during the following years through the first years of formal schooling.

All these stages require specific cognitive modifications that need to be addressed to obtain a picture of children's numerical abilities. This is the aim of the next section. Since the populations studied in the two following experiments ranges from preschoolers to third graders, the abilities of older children will not be described as it is not relevant to the aim of the present work. Only a brief introduction to calculation abilities will be offered because our studies are mainly focused on the development of the internal representation of numbers.

II.2.1. Non symbolic processing.

Recognizing equivalence across different sets is a critical numerical accomplishment of early childhood. It represents the ability of understanding that different sets, containing assorted types of items, are equivalent if the number of items is the same. Three dogs is a set equivalent to three apples or three cars and it is even equivalent to a set composed of a fork, a spoon and a knife. In a first study, Mix, Huttenlocher and Levine (1996) showed that preschoolers were able to recognize equivalence between sets if they were similar (i.e., black disks and black dots) but were impaired if the sets were intermodal (i.e., sounds and dots). However, a subgroup of children that had proficient counting abilities could also match dissimilar sets. This result suggest that conventional counting skills may play an important role when similarity is low. Later, Mix (1999) studied the same ability in 3 to 4-years-old children only with visual stimuli that varied in similarity (dots, shells and random objects). Results indicated that the ability to match numerically equivalent sets improves with age. At age 3 only identical comparisons were successful, then progressively by 4 and ½ children were able to match those sets that were composed of heterogeneous items. Moreover, children that had better knowledge of conventional count words had also better performance in recognizing numerical equality. Long before, Siegel (1968) had shown the same result with children aged between 4 and 5. They performed significantly worse when the sets were heterogeneous than when the sets were homogeneous. Only the older group of children performed well in both conditions. Unfortunately no correlation with other numerical skills were performed leaving unclear the reason of matching improvement.

On comparisons of large numerosities, Rousselle, Palmers and Noël (2004) presented 3-year-old children with a task where sets were accurately controlled for perceptual cues and ratios. In the most perceptually controlled condition, children failed to use numerical cues to compare sets. Moreover, they also tested children's verbal counting skills and correlated the performance in the two tasks. Results showed a relation between the development of numerically-based judgments and some cardinality knowledge.

Using a numerosity comparison task, Huntley-Fenner and Cannon (2000) investigated 3 to 5-years-old children's performances on pairs of arrays that varied by either ½ or ⅔ ratios. Some perceptual variables as density and length of arrays were controlled. The authors analyzed the proportion of accurate responses and they observed that items with a ⅔ ratio were harder and that errors varied systematically with ratios, indicating that performance was consistent with an analogue magnitude representation. Moreover, performance in the discrimination task correlated with the ability of reciting number words but not with the ability of counting sets of objects.

A recent study carried out in Finland (Hannula & Lehtinen, 2005), investigated the Spontaneous Focusing on Numerosity (SFON) of children aged between 3 ½ and 6-years-old. The task was striking in simplicity. The child was asked to imitate (without further numerical instructions) the experimenter in the actions she was realizing. The actions consisted in feeding a puppet a number of food pieces. The experimenter first showed how she fed the puppet and clearly took a piece of food at a time (up to 3 pieces) and introduced them in the puppet's mouth. Then it was the child's turn. The experimenter took note of the number of pieces the child gave and if any type of comment was made on numerosity. This task allowed the experimenter to notice if the child was able to focus on the numerical aspect of the task. Spontaneous focusing also correlated with other mathematical skills such as counting. The authors conclude that within a child's existing mathematical competence, there is a separate process which refers to it's personal tendency to focus on numerosity.

As for non-symbolic arithmetic, Barth and colleagues tested the ability of preschoolers and adults to add and subtract large numerosities (Barth et al., 2006). Perceptual cues were controlled and proposed outcome sets varied according to four ratios from the correct outcome. Both adults and children succeeded above chance and were sensitive to the ratio: accuracy decreased with a ratio closer to one. In another study , Barth and colleagues (Barth, La Mont, Lipton, & Spelke, 2005), showed that 5-years-olds are able to perform calculations both when presented as visual non-symbolic material and when stimuli are presented across modalities. Performance was correlated with comparison abilities but not with symbolic versions of the addition tasks. The authors conclude that numerical quantity representations are computationally functional and may provide a foundation for formal mathematics. According to the authors, abstract knowledge of number and addition precedes, and guides, language-based instruction in mathematics.

II.2.2. Symbolic processing.

After infancy, the ability to perceive sets as discrete objects is likely present and it is reasonable to assume that children pull together formal concepts of numerosity from such beginnings. A theory proposed by Gelman states that nonverbal numerical reasoning is the starting point to learn both number words to count and the rules for how to use them. Numerical concepts have in this theory an ontogenetic origin and a neural basis that are independent of language (Gelman, 2006; Zur & Gelman, 2004; Gelman & Butterworth, 2005).

A. Counting.

What is counting? Nothing easier to a numerate adult than counting. But learning how to do it properly requires many years and a number of cognitive steps. Gelman (Gelman, 2006; Gelman & Meck, 1983) describes the different rules that a child needs to learn to make a proficient use of number words. Indeed, it is just not enough to know the words. Each number word has to be in a one-to-one correspondence with the item to be counted, words must have a stable order and finally, the last word represents the cardinal value of the set. According to the authors' model, the child has to learn the number words and learn to map them onto the internal numerical representation in order to create a memory for each numerosity (Gallistel & Gelman, 1992).

By the age of 3½ children already understand how the counting system determines numerosity and have acquired the cardinal meaning of all the number words within their counting range (Wynn, 1990). Remarkably they start very early in the counting stage to understand the specific numerosity of number words. In fact, even if they master only number words "one" or "two" they already understand that they refer to a specific number of objects. From this, children progressively acquire the meaning of each following number until they generalize the principle that any following number in the number word sequence is one item more than the previous (Wynn, 1992; Margolis & Laurence, 2007; Rips, Asmuth, & Bloomfield, 2006); for a discussion: (Rips, Asmuth, & Bloomfield, 2007; Gelman & Butterworth, 2005). It takes about a year for the child to reach a complete understanding of the counting system. Moreover, Mix (Mix, 2002) has shown in a single case longitudinal study that during the second and third year of life, a child is involved in a variety of one-to-one activities (give one to mummy, one to daddy…) embedded in social and linguistic contexts that helps her to construct number concepts.

While Gelman and collaborators suggest an innate knowledge of the counting principles others argue that the count list is learned just as the letters of the alphabet, without attributing any significance to the order. The knowledge of counting principles would be constructed by attempting to make sense of the number words themselves (Le Corre & Carey, 2008; Le Corre & Carey, 2007). Le Corre and Carey, put forward the "*enriched parallel individuation*" hypothesis. According to this hypothesis, numerical representations are based on the combination of two processes: parallel individuation and set-based quantification. The former is a system that represents individuals by creating working memory models in which each individual in a set is represented by a unique mental symbol and is able to store only a limited number of individuals in parallel (3-4). This system contains no symbols for numbers but they are represented implicitly through the criteria that maintain one-to-one correspondence between working memory and individuals in the world. The latter process, set-based quantification, is the root of the meaning of all natural language quantifiers, as for example, in English, the singular determiner “a”, the quantifiers “some” and the plural marker “-s”. The authors thus explain that, children would construct the meaning of number “one” as a situation where only an individual is represented in working memory. “Two” would be the situation where two individuals are represented in working memory and are in addition linked to linguistic markers of plural. This would be the type of associations that children would learn for all numbers up to 4. With time, the mapping of small numerosities to models of individuals in working memory could lead to the understanding that adding one corresponds to the next word in the counting list and therefore lead to the acquisition of the counting principles. In this process, an innate numerical representation would not play a central role but in view of the present work it does not exclude its existence. What constitutes a critical difference in this model is that the acquisition of the exact numerical representation and the concept of infinity, which differentiates humans from the other species, are gained by a bootstrapping process based on the parallel individuation system (3-4 elements). Conversely, for those models that postulate the existence of a magnitude representation, the exact numerical representation is acquired by mapping number words onto magnitudes.

B. Symbolic comparison.

Studies on symbolic comparison usually aim at understanding the automaticity of numerical processing and the activation of the underlying magnitude. Several studies have investigated magnitude comparison in children and have shown that they are sensitive to the same effects as those observed in adults (i.e., size and distance effects). Several researchers have shown that children exhibit a distance effect, they are faster and make less errors when the numbers are numerically further apart (Duncan & Mc Farland, 1980; Sekuler & Mierkiewicz, 1977). In Duncan and McFarland’s study, children form kindergarten through fifth grade were asked to make a same/different judgment on single digit pairs. This task could easily be solved only by visual cues but, as young as 6, children were influenced by numerical distance indicating that Arabic digits were processed all the way to the semantic level.

To further investigate the automatic processing of numbers, Girelli, Lucangeli and Butterworth (2000), used two Stroop type paradigms. In their study, children from 1st, 3rd and 5th grade as well as adults had to compare, numerically or physically, pairs of digits (1-9). Both physical and numerical distances in the pair were manipulated. In one task, participants had to judge the physical dimension ignoring the numerical one and in the other task they had to attend to the numerical dimension ignoring the physical one (e.g., 2 vs. 4). The physical and numerical dimensions could be congruent (i.e., the largest number was also physically bigger) or incongruent (i.e., the numerical magnitude was opposite to the physical one). There was also a neutral condition in which one dimension was kept constant (the pair either varied in size or number but not both). Results indicated that the physical dimension interfered with and facilitated numerical judgment for all age groups whereas only the older children and adults were also influenced by numerical magnitude in the physical comparison task. This suggests a gradual achievement of automaticity in number processing.

Mussolin and Noël (2007) tested the possibility that numerical processing could just not be fast enough to show effects in younger children. Using the same paradigm as Girelli and al. (2000) and a population of 2nd, 3rd and 4th graders, they controlled for the time of processing for the numerical and physical dimensions. The two digits were presented at first with the same physical size and progressively changed. Results indicated that all age groups were subject to numerical interference in the physical judgment task but only as long as the distance between the two digits was large enough (i.e., of ½) and enough time to process numerical information was allowed.

A parity judgment task on single digits also assessed the automaticity of numerical information in children from 7-years-old to 13-years-old (Berch, Foley, Hill, & Ryan, 1999). Indeed the authors noticed that from 4th grade (9-years-old), parity information is directly retrieved from memory without the use of calculation strategies and as early as 3rd grade, children exhibited a SNARC effect (Spatial Association of Response Codes: in a reaction time task, small numbers are responded to faster with the left hand and the reverse is true for large numbers, Dehaene et al., 1993). Once more, results indicate that, also for children, numerical information is automatically accessed although irrelevant for the task.

C. Internal numerical representation in children.

All studies mentioned above have described children's abilities in numerical cognition but only a few addressed the internal representation of numbers. Some authors have shown that ratio influences the ability to discriminate between sets and others have analyzed standard errors in relation to ratio, in both cases supporting the hypothesis of a compressed or noisy representation of numbers (see Chapter I, paragraph I.1.1. for models of numerosity processing). Only recent studies carried by Siegler and colleagues have investigated the development of such internal representation and its relation to other formal numerical knowledge (Siegler & Opfer, 2003; Siegler & Booth, 2004; Booth & Siegler, 2006; Opfer & Siegler, 2007).

Siegler and collaborators have conceptualized what they consider a pure numerical estimation task that reveals the underlying numerical representation. In a first study (Siegler & Opfer, 2003), the authors investigated second, fourth and sixth grade children as well as adults in two estimation tasks. In one task, number-to-position (NP) task, participants had to estimate the numerical position of given numbers on a line. Two different intervals were proposed: on the left end there was always 0 and on the right end there was either 100 or 1000. The second task, position-to-number (PN) task, was the complementary version of the former. Instead of estimating the position of a given number, participants were asked to estimate the number corresponding to a given position on the intervals 0-100 an 0-1000:. According to the authors, these tasks are advantageous because they allow a direct mapping between numerical and spatial representations thus embodying the core property of estimation. Moreover, the absence of specific measurement units makes these tasks particularly suitable for young children. The analysis showed that children's estimates changed over time. At the youngest age, children overestimated small numbers and compressed large numbers to the end of the scale (logarithmic shape of the estimates). This was however modulated by numerical context. Indeed, when the context was familiar (i.e., the smaller interval) positions of estimates were linear. Older children (grade six) and adults positioned numbers linearly on both intervals sizes. The authors stress that the logarithmic positioning on the larger scales was not the consequence of a general misunderstanding of numerical quantities since they could perfectly achieve a linear positioning on the smaller scale with numbers that were shared by both intervals. The logarithmic positioning for the larger interval would therefore be the consequence of an intuitive and logarithmic representation of numbers, used in unfamiliar situations. These results are taken as evidence for two main conclusions. First, children possess multiple numerical representations and second, different representations may apply according to the context.

To support their first study, Siegler and Booth (2004) replicated the experiment with a population of preschoolers, first and second grade pupils. The aim was twofold: first, to show parallels in the developmental sequence at different ages; second, to test the relation between number-line estimation and math achievement test scores. The development of numerical representations was replicated, preschoolers predominantly relied on logarithmic representations, progressively shifted to a linear one for familiar numerical contexts at grade one and became predominantly linear at grade two. Moreover, accuracy of estimates on the number-line task was correlated to math achievement scores (SAT-9).

Furthermore, Booth and Siegler (2006) have tested the generality of the developmental shift across different types of estimation. Performance in the number-line task was compared to other estimation problems such as approximate additions, numerosity estimation, measurement estimation. Results from a population of children from kindergarten to 3rd grade revealed that accuracy substantially improved across ages and for all tasks. Furthermore, the authors suggest that the reason for primary school children's poor estimation is the reliance on a logarithmic rather than linear representation of numerical magnitudes.

In a recently published paper, Opfer and Siegler (2007) studied the ability to improve estimates according to the type of feedback. Indeed, they tested the theoretical prediction that feedback providing the greatest discrepancy between the two representations would yield the greatest representational change. In this study, only second graders that relied on a logarithmic representation undertook the task and the experimenters manipulated the degree of discrepancy between the estimation made by the child and the feedback given. In other words, for each child the logarithmic fit of estimates was calculated and feedback was given on one estimate according to its discrepancy (high or low). Those children that were in the high discrepancy feedback group were the ones who showed the largest representational change. Strikingly, it occurred after a single feedback trial and for all the estimates of the interval. These results suggest that cognitive changes can occur over an entire representation with a minimal but significant feedback.

II.2.3. Summary of children's numerical abilities.

In this somewhat brief introduction of children's numerical ability we have shown that in the first years of life, a child is faced with a great deal of learning challenges. Theories on how they acquire the counting skills still diverge but it seems reasonable to state that an early numerical representation is present and guides future numerical learning. With time, the internal representation develops and becomes more precise. In parallel to this cognitive transformation, children learn to map numerical symbols onto numerosities and back until the processing of numerical information reaches automaticity. Finally, with formal teaching, children learn to use an adequate numerical representation where numbers are linearly represented for those numerical magnitudes that have been learned.

Lucangeli, Iannitti and Vettore (2007), have summarized in the following way the developmental stages that are achieved in the first years of life of a child. From birth to about 2 years of age, infants posses uniquely preverbal numerical competences that is subitizing and numerical expectations. In the next two years, children have to learn the number sequence and the one-to-one correspondence as well as the cardinality principle. Beyond this age and throughout school, they learn to write and read number words, they acquire formal computational procedures and, we can add to the authors' summary, they learn to rely upon a linear representation.

III : Numerical Representations in Adults.

Human adults are able to exactly represent a number just as they are able to approximate a numerical quantity. The first competence relies on a verbal system, the second is based on an approximate number system similar to a number line. The first section of this chapter offered a general overview of humans' numerical abilities and introduced the concepts of subitizing and estimation, as well as the various models that allow to explain these, but also other, numerical competences. In the present section, the focus is on the performance of skilled adults when faced with non symbolic tasks. That is, their ability of approximating and estimating numerical information. The tasks may be classified in two groups: those that require to estimate or compare numerosities and those that require participants to compute an operation (addition or subtraction) on sets of items.

III.1. Numerosity Comparison and Estimation.

Numerosity comparison is a task where two sets of items, usually dots, are presented with the instruction to either judge if the numerosity is the same or to choose the smaller/large set. Obviously, time limits are set to avoid counting although participants are highly discouraged from doing so. Estimation tasks are subdivided in those tasks requiring a verbal answer and those that require the production of a motor sequence (e.g., press a response key) corresponding to a given number or numerosity. This second paradigm has also been widely employed with animals allowing a comparison of response patterns.

Minturn and Reese (Minturn & Reese, 1951) have observed that estimation answers largely diverged with an estimated factor of 4 above and below the true answer. However, subsequent studies have reported mostly an underestimation (Indow & Ida, 1977; Krueger, 1982; Krueger, 1984). Moreover, underestimation is present since the first trial, increases during the experiment and the larger the numerosity the greater the variability in the response (Krueger, 1982). This variability has further been studied by Whalen and colleagues (Whalen et al., 1999) and described as the scalar variability property. Indeed, the coefficient of variation (CV = standard deviation/mean) is constant across numerosities. In another estimation experiment (Cordes, Gelman, Gallistel, & Whalen, 2001), participants had to press a response key either suppressing the possibility to count by repeating a sound or by counting as fast as they could while answering. The second condition reveal more accurate and less variable answers.

In addition, Krueger (1984) observed that when a reference stimuli is given, response patterns tend to be less underestimated and the coefficient of variation decreases for all subsequent stimuli. More recently, Lipton and Spelke (2005) have shown that in some conditions participants are able to properly estimate the numerosity of sets. In their study, two sets of dots where presented and the numerosity of the first set was given as reference for the estimation of the second. Average estimations for the second set perfectly matched real numerosity. Only one study (Izard & Dehaene, 2007) manipulated the accuracy of feedback. In one condition participants had to estimate without feedback numerosity up to 100 dots and in the other condition they were calibrated on numerosity 30 with either a correct or a wrong feedback (larger or smaller). Results indicated that when no calibration was given, participants underestimated systematically whereas feedback significantly improved performances. Moreover, the type of feedback influenced average estimation and extended to the whole range of numerosities tested. For both conditions, a scalar variability of estimates was observed; moreover, their mean followed a power function of numerosity.

It is important to highlight that perceptual variables influence the ability of estimating. Several authors have shown that the disposition of dots in the array play an important role. When dots are grouped and more dense, responses are more underestimated that when they are evenly and regularly spaced (Durgin, 1995; Allik, Tuulmets, & Vos, 1991; Allik & Tuulmets, 1991; Ginsburg, 1978). Moreover, the same effect has been observed with temporal presentations: evenly presented stimuli in time are over-estimated whereas irregularly presented stimuli are underestimated (Allik & Tuulmets, 1993).

Adults' performance in discrimination between numerosities mirrors the pattern observed in the comparison of symbolic numbers, with size and distance effects[5] (Buckley & Gillman, 1974; van Oeffelen & Vos, 1982). The amodal or abstract representation that underlies the ability to process numerosities has been studied by Barth in two experiments

[5] The distance effect is the observation that the smaller is the numerical distance between two numbers/numerosities the longer the reaction times and the higher the error rate. The size effect consists in an increased difficulty (error and RT) associated to the numerical size of the smaller number/numerosity (see section I for more details).

(Barth, Kanwisher, & Spelke, 2003; Barth et al., 2006) in which participants were asked to discriminate between sets of stimuli of either one format or between formats. Stimuli could be visual or auditory, sequential or simultaneous. Results showed that for all conditions, reaction times and accuracies are determined by the ratio between the numerosities to compare. This led the authors to conclude that numerical representation is independent from the format of the stimulus.

A slightly different paradigm was presented by Piazza and collaborators (2004). The aim was primarily to investigate neural correlates of numerosity in an fMRI study. Using an habituation paradigm where participants had to judge if a stimulus was numerically different from the previous ones, they observed that the only regions that responded to numerical distance were localized in the left and right intraparietal sulci (IPS, including the horizontal segment). Their work confirms and extends data observed both with single cell recording in animals and in neuroimaging experiments with humans.

Cantlon, Brannon, Carter and Pelphrey (2006), using a similar paradigm as the one used by Piazza and al. (2004), they addressed the question of the early development of brain areas for processing abstract numerical information. They found that in 4-years-old, the IPS responded similarly to adults for numerical changes. The authors concluded that the neural correlates of numerical cognition are active early in development prior to formal instruction and symbolic experience.

Overall, responses to estimation and comparisons tasks are subject to Weber’s law: judgments become increasingly less precise as numerosity increases and the variability increases proportionally to the mean response, such that numerosity discrimination is determined by the ratio between numbers (Gallistel & Gelman, 1992; Whalen et al., 1999; Cordes et al., 2001; Izard & Dehaene, 2007; Piazza et al., 2004).

III.2. Approximate Calculation

Approximate calculation tasks have been created to suit the necessity of testing implicit computational knowledge especially in children prior to language acquisition (see section II.1. and II.2.). Nevertheless, in one study (Barth et al., 2006), both adults and children were tested to observe similarities in performance in simple calculations (additions and subtractions). Three separate experiment were constructed for the adult population: in the first experiment, arrays of dots were presented visually and sequentially; in the second experiment, formats were mixed between visual and auditory; and in the last experiment, stimuli where presented in a movie format where dots moved on the screen reaching (addition) or leaving (subtraction) an occluding patch. All three experiments provided evidence that adults are able to add two arrays, or subtract one array from the other, and then compare the sum or difference to a third array. Performance also showed the classical ratio signature for large number representations.

This paradigm has also been used to test numerical abilities and knowledge in populations that lack words to express large numbers, as the Mundurucús found in the Amazonian forest of Brasil. Pica and collaborators (Pica, Lemer, Izard, & Dehaene, 2004) observed their performance in several numerical tasks including an addition task of large numerosities. For the computational task they used a similar task as the third experiment in Barth (2006) except that the occluder was replaced by a can were items fell inside or came out from it. Performance was above chance with a minimum of 80% correct responses and answers were solely affected by distance. In short, Mundurucús had no difficulty adding approximate numerosities and their precision was identical to the one observed in a group of French controls. The authors conclude that the data support the distinction between a verbally based numerical representation (very limited in the Mundurucú's language) used for exact number processing and an approximate system that is present early in development, independent of language proficiency and shared with other species (Nieder, 2004; Feigenson, Dehaene, & Spelke, 2004).

IV : <u>Introduction to the Studies</u>

After the general theoretical introduction in this chapter and more specifically the summary on numerical processing in infancy and childhood (section II), two experiments will give some insight into the development of numerical representation in space. In Experiment 1a, a developmental pattern suggested by Siegler and colleagues (Siegler & Opfer, 2003; Siegler & Booth, 2004; Booth & Siegler, 2006) is replicated with a population of younger children (4 to 6 years old). The task used is considered by Siegler and Opfer (2003) a pure numerical task. The Number-to-Position task, according to the authors, allows the investigation of children’s internal spatial representation of numbers.

In Experiment 1b, the performance of preschoolers to the Number-to-Position task was related to the ability in discriminating numerosities. Using the paradigm suggested by Piazza and collaborators (Piazza et al., 2004), the discrimination ability was evaluated with the *internal Weber fraction* for all children and correlated to the positioning performance in the Number-to-Position task.

Experiments 2a was designed to investigate the representation of numerical and non-numerical sequences using the Siegler et al. (2003, 2004) positioning task. Indeed, the developmental pattern observed with Number-to-Position task could be general to other sequences (e.g., letters). An alternative hypothesis to the one offered by Siegler, is that the task does not tap on a core numerical representation but highlights a common positioning bias due to the limited knowledge of a given sequence. The first items know would be spread apart whereas the other would be clustered together.

In Experiment 2b, the same children were tested on the mental bisection task used by Zorzi and collaborators (Zorzi et al., 2002; Zorzi, Priftis, Meneghello, Marenzi, & Umiltà, 2006) which also highlights the spatial representation of numerical and non-numerical sequences. Therefore, the Number-to-Position task and the mental bisection task were compared. Indeed, if both measure the same spatial representation of sequences, a positive correlation should be found.

After these developmental studies, Experiment 3 is focused on various adults’ numerical abilities. Participants numerosity discrimination ability was estimated using the *internal Weber fraction*. This value was correlated to other two numerical tasks: a double-digit number comparison and a task on approximate calculation. In the addition and subtraction approximation task, the symbolic and non-symbolic stimuli were crossed to observe any specific presentation format effect on what McCrink and collaborators have termed the *Operational Momentum* (McCrink et al., 2007). That is the tendency to overestimate additions and underestimate subtractions.

Numerical cognition is investigated from a different perspective in Experiment 4 with a single case study of a number-colour synaesthete. NM claims to perceive colours when viewing digits but not for sets of dots. To investigate the level at which the synaesthetic perception is triggered, two Stroop-like tasks were created: one using digits, to replicate previous findings, and the second using canonical and non canonical patterns of dots. In the discussion a new model for synaesthetic associations to numerical representations is offered.

Chapter 2: Development of Numerical Representations in Preschoolers

Ⅰ: **Experiment 1a: Numerical Representations in Preschoolers.**[1]

Abstract

Previous developmental studies of numerical estimation have shown that children performing the number-to-position task (Siegler & Opfer, 2003; Siegler & Booth, 2004) increasingly rely on formally appropriate, linear representations and decrease their use of intuitive, logarithmic ones. Here we investigate the development of numerical representations in a much younger population of preschoolers (from 3 ½ to 6 ½ y.o.) using 0-to-100 number lines and a novel set of 1-to-10 number lines. On the large interval, estimates became more accurate but also increasingly logarithmic with age. In contrast, estimates became more accurate and shifted from logarithmic to linear in the small number range (1-10) with increasing age, following the developmental trend previously reported with older children on 0-to-100 and 0-to-1000 number lines. Moreover, estimation accuracy was correlated with formal numerical knowledge measured by naming accuracy of one-digit Arabic numbers. Those results suggest that, the development of numerical estimation is built on a logarithmic coding of numbers – the hallmark of the approximate number system subserving the non-symbolic representation of numerosities – and is subsequently shaped by the acquisition of cultural practices with numbers.

[1] This experiment is an article in preparation: Berteletti, Lucangeli, Piazza, Dehaene and Zorzi.

I.1. Introduction.

Prior to language acquisition, infants and children are able to discriminate between sets of objects (Lipton & Spelke, 2003; Wynn, 1996; Xu & Spelke, 2000; Xu, 2003) or sequences of sounds (Bijeljac-Babic, Bertoncini, & Mehler, 1993) by relying only on numerical information. Around the age of three, children learn some counting words and understand how they refer to a distinct, unique numerosity (Wynn, 1996). Lipton and Spelke (2006) have shown that 5-years-old children understand that number words outside their counting range also refer to unique quantities and that a specific number word cease apply to a set when an item is removed from it. It is only when the item is reintegrated into the set that children again apply the original numerical label, suggesting that children possess a productive system for representing numbers before formal education. All these abilities, that appear early and are already present before language acquisition, provide evidence that numeracy is founded upon an early non-symbolic system of numerical representation (for reviews see Butterworth, 2005; Carey, 2001; Feigenson, Dehaene, & Spelke, 2004).

It has been proposed that the infant's sense of numbers, in the first year of life, is based on two "core systems" (Feigenson et al., 2004): (i) a *small number system* that is accurate for numbers up to 3; this is essentially the perceptual system for tracking objects. (ii) an *approximate* number system for representing larger numerosities; this system encodes approximate numerosities as analog magnitudes, which are usually thought of as overlapping distributions of activations on a mental number line that is logarithmically compressed (for review Dehaene, Piazza, Pinel, & Cohen, 2003,).

In subsequent years, children in our culture achieve a mental representation of number that goes beyond these core number systems in at least two different ways. First, numerate children and adults are able to go beyond approximate numerosities and can distinguish and represent exact numerosities greater than 3. Second, part of the adult concept of exact number implies a linear (rather than logarithmic) mapping between numbers and space, such that numbers can be used for measurement. It is a controversial matter as to how both advances are achieved, although it is clear that the child experience with counting and number words plays a major role (e.g., Le Corre & Carey, 2007). Notably, a developmental transition from logarithmic to linear representations has been documented in the seminal studies of Sigler and collaborators (Siegler & Opfer, 2003; Siegler & Booth, 2004).

Siegler and Opfer (2003) investigated the implicit numerical representations used by second, fourth and sixth grade pupils as well as adults in two numerical estimation tasks. In one task, individuals were asked to estimate the position of a given number on a "number line" with 0 at one end and either 100 or 1000 at the other end (*number-to-position task*). The other task was the complementary version of the former and consisted in estimating the number associated with a given position on the number line (*position-to-number task*). According to the authors, these tasks embody the core property of estimation because they require translation between numerical and spatial representations. Moreover, the absence of specific measurement units makes these tasks particularly suitable for young children.

Results of the estimation tasks indicated that with increasing age, children's estimates changed and shifted from a logarithmic to a linear representation. At the youngest age (grade two and four), children overestimated small numbers and compressed large numbers to the end of the scale (logarithmic positioning) when the context was not familiar (0-to-1000 number line), but they positioned numbers linearly when the context was well known (0-to-100 number line). Moreover, the oldest group (grade six) positioned numbers in a linear manner on both small and large scales, similar to adult participants (Siegler & Opfer, 2003). It is important to highlight that the logarithmic representation used by second and fourth graders on the larger interval (0-to-1000 number line) was not the consequence of a general misunderstanding of numerical quantities since both groups had a linear representation on the 0-to-100 number line. This provides evidence that the logarithmic fit for the larger scale is the consequence of an intuitive and logarithmic representation that is used when the context is unfamiliar. Finally, the authors showed that the context greatly influenced the numerical representation since the same numbers were treated differently according to the interval of reference, with identical numbers placed linearly on the small interval, and logarithmically on the large interval. Thus, multiple numerical representations coexist and that the choice among them changes with age and experience.

In a second study, Siegler and Booth (2004) replicated the experiment with a population of preschoolers, first graders and second graders (mean ages: 5.8, 6.9 and 7.8 years respectively). Results indicated that, from kindergarten to second grade, the developmental sequence for the 0-to-100 number line was equivalent to the developmental sequence shown with the larger 0-to-1000 number line by the older population of children in the study of Siegler and Opfer (2003). That is, the predominant use of a logarithmic representation was followed by the use of both linear and logarithmic representations depending on the scale, and finally by consistent reliance on a linear representation. Furthermore, mathematical achievement was found to correlate with the linearity of the estimates. In a subsequent study, Booth and Siegler (2006) found a correlation between performance in the number-to-position task and other numerical estimation tasks such as approximate addition, numerosity and measurement estimation.

The current study investigated numerical estimation in a population of even younger children. The aim was to assess children's ability to provide reliable estimates (whether logarithmic or linear) as early as 3.5 years old and to further characterize the developmental trend that leads to the emergence of a linear representation of numbers. We tested children aged from 3.5 to 6.5 years old on the number-to-position task. We used the 0-to-100 interval used by Siegler and Booth (2004) as well as a smaller interval 1-to-10 that was not used in previous studies. The latter interval contains numbers that should be familiar even to the youngest children tested. Moreover, the smallest interval started from 1 because children learn the counting sequence starting from one and the concept of zero is usually introduced later (Butterworth, 1999). Children were also tested on basic numerical knowledge to investigate its relation to their ability to estimate linearly, as was done in previous studies (Siegler & Booth, 2004; Booth & Siegler, 2006). This was done with a simple digit naming task, which, for very young children, has been found to be a strong predictor of other numerical tasks (Ho & Fuson, 1998; Huntley-Fenner & Cannon, 2000).

We predicted that performance of the youngest children would reveal a purely logarithmic representation and that the developmental pattern would show a shift from logarithmic to linear representation, but only for the smaller interval. Thus, the hypothesis that children possess multiple numerical representations (Siegler & Opfer, 2003) leads to the prediction that performance might become completely linear on the smaller interval but still remain logarithmic on the larger interval.

I.2. Method.

Participants

Forty-six children (21 females), recruited in two different kindergarten schools from north-eastern Italy, were divided in three age groups: the youngest group (n=11) had a mean age of 48 months (SD = 4), the middle group (n=16) had a mean age of 60 months (SD = 3) and the oldest group (n=19) had a mean age of 70 months (SD = 3).

Procedure

Two trained female teachers from each school met with the children individually during school hours in a quiet classroom for about half an hour. The familiarity with the teachers helped children to feel comfortable in the testing session. Children were first tested on basic numerical knowledge, and were asked to name Arabic digits from 0 to 9. Digits were randomly presented on separate cardboards of 5 x 5 cm and children were asked to name them aloud. For each digit correctly named, the child was given a score of one, for a maximum score of 10. No feedback was given during the task.

Numerical estimation was tested with the Number-to-Position task developed by Siegler and Opfer (2003). Note that this task does not require knowledge of measurement units. Children were presented with 25-cm long lines in the centre of white A4 sheets. Two different intervals were administered: 1-10 and 0-100. The ends of the lines were labelled on the left by either 1 or 0 and on the right by either 10 or 100. The number to be positioned was shown in the upper left corner of the sheet. All numbers except for 1, 5 and 10 had to be positioned on the smaller interval, whereas for the larger interval the numbers were 2, 3, 4, 6, 18, 25, 48, 67, 71, 86 (corresponding to sets A and B for the same interval used in Siegler & Opfer, 2003).

Order of the two intervals and order of items within each interval were randomized. Each line was seen separately from the others to avoid influence from previous positioning. Instructions given at the beginning were: “We will now play a game with the number lines. Look at this page, you see there is a line drawn here. I want you to tell me where some numbers are on this line. When you have decided where the number I will tell you is, I want you to make a mark with your pencil on this line.” To ensure that the child was well aware of the interval size, the experimenter would point to each item on the sheet while repeating for each item: “This line goes from 1(0) to 10(100). If here is 1 (0) and here is 10 (100), where would you position 5 (50)? The numbers to position were always named verbally by the experimenter.

The numbers 5 and 50 were used as practice trials for the small and large interval, respectively, and to check if the task was properly understood. Experimenters were allowed to rephrase the instructions as many times as needed without making suggestions about where to place the mark.

I.3. Results.

Number-to-position task

The analysis on the accuracy of children's estimates was computed using the percent absolute error of estimation for each child. This was calculated according to the following equation (Siegler & Booth, 2004):

percent absolute error = (estimate – estimated quantity) / scale of estimates

A one-way ANOVA on mean percent absolute error was computed for each interval with age as between-subjects factor. For both intervals, results indicated that the three groups were significantly different and the accuracy of estimation increased with age (1-10 interval: $F(2,43) = 6.14$, $p < .01$, $\eta^2 =.05$; 0-100 interval: $F(2,43) = 4.22$, $p < .05$, $\eta^2 =.01$). The youngest and the middle groups significantly differed from the oldest group on post hoc comparisons for the interval 1-10 (*ps* < .05). Percent absolute error for the youngest, middle and oldest group were 28%, 24% and 15% respectively. For interval 0-100, percent absolute error were 32%, 30% and 23% respectively and only the youngest group significantly differed from the oldest group on post hoc comparisons ($p < .05$). It is worth noting that for the 0-100 interval the accuracy of estimation for our oldest group is slightly better than the accuracy of the comparable age group studied by Siegler and Booth (2004; 27% in their study).

To analyze the pattern of estimates, the fit of linear and logarithmic functions were computed. These fits were first computed on group medians, and then for each individual child.

For group medians, the difference between models was tested with a paired-sample t-test on the distances between children's median estimate for each number and a) the predicted values according to the best linear model and b) the predicted values according to the best logarithmic model (see Figure II.1.). For the 1-10 interval, the model with the highest r-square was logarithmic for the youngest group (R^2 log = 87%, $p < .01$) but it did not significantly differ from the linear fit (R^2 lin = 84%, $p < .01$; $t(6) = -1.17$, $p \geq .05$). For the two older groups, the fit of the linear model was significantly better than the fit of the logarithmic model (intermediate group: R^2 lin = 95%, $p < .001$ vs. R^2 log = 89%, $p < .01$; $t(6) = 2.82$, $p < .05$; oldest group: R^2 lin = 97% $p < .001$ vs. R^2 log = 88%, $p < .01$; $t(6) = 4.05$, $p < .01$). For the 0-

100 interval, the best fitting model for the three groups was logarithmic, but the r-square value increased with group age (youngest: R^2 log = 59%, $p < .01$ vs. R^2 lin = 46%, $p < .05$, $t(9) = -1.2$, $p > .05$; intermediate: R^2 log = 85%, $p < .001$ vs. R^2 lin = 57%, $p < .05$, $t(9) = -3.15$, $p < .05$; and oldest: R^2 log = 94%, $p < .001$ vs. R^2 lin = 70%, $p < .01$, $t(9) = -3.62$, $p <. 01$). The absence of statistical difference between the two models for the youngest group might be explained by the large standard deviations (linear model μ = 15.6, *SD* = 12.2 vs. logarithmic model μ = 18.2, *SD* = 13.5). Alternatively, this lack of significant difference could be explained by the observation that during testing many children from the youngest group adopted non-numerical strategies to perform the task (e.g., they alternated between right-side and left-side marks).

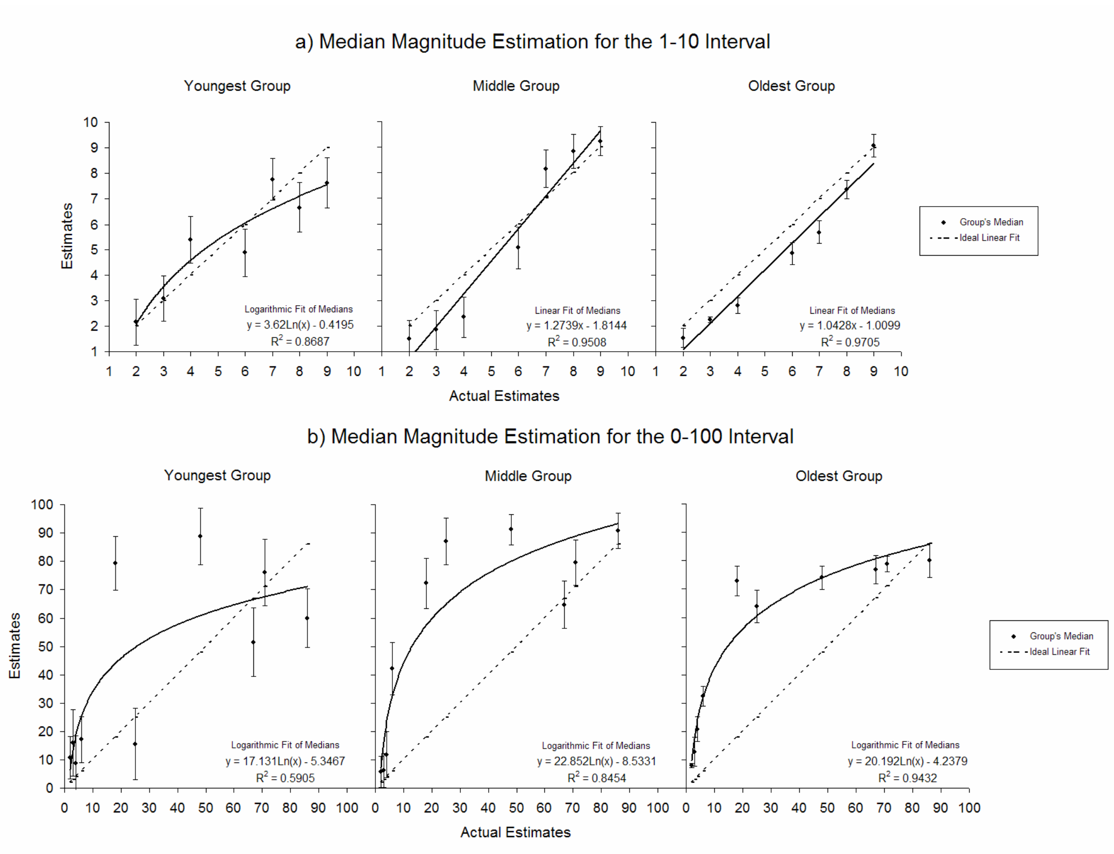

Figure II.1.: Best logarithmic or linear fit as a function of the interval and age group. For the small interval the type of fit evolves from logarithmic to linear with age and for the larger interval the precision of the logarithmic fit increases with age group.

Regression analyses were then performed on the data of individual children. The best fitting model between linear and logarithmic was attributed to each child, whenever significant. For example, the child was attributed a logarithmic representation for a given interval if the highest r-square was logarithmic. If both models failed to reach significance the child was classified as not having a representation for the interval considered[2]. For each interval, children were therefore classified as having a linear, logarithmic, or no representation (see Table II.1.). Spearman rank correlations were calculated between group (1= youngest, 2= intermediate and 3= oldest) and type of representation (1= no representation, 2= logarithmic and 3= linear). Results indicated that for the smaller interval the estimation tended to become linear with age (r_S = .49, p<.001, one-tailed test). For the bigger interval the logarithmic representation became more predominant with age, whereas the proportion of children that were unable to position numbers diminished (r_S = .33, p<.05, one-tailed test). These data highlight the developmental trend that with increasing age children learn to position numbers more accurately. When the numerical context is difficult, or unfamiliar, they rely on an intuitive, logarithmic representation, whereas when the numerical context is more familiar, they use a formal linear representation.

Table II.1.: Type of representation adopted by children as a function of group and task.

	Type of representation			
Task	**None**	**Logarithmic**	**Linear**	**Total**
1-10 Interval				
Youngest	5	4	2	11
Intermediate	5	3	8	16
Oldest	1	3	15	19
0-100 Interval				
Youngest	7	3	1	11
Intermediate	9	6	1	16
Oldest	4	13	2	19

Note. Cell values represent number of children.

This pattern is also supported by the analyses conducted on both intervals according to the type of representation (see Table II.2.). Indeed, the ability to position numbers on one interval is significantly correlated with the ability to position numbers on the other interval (r_S = .39, p < .005, one-tailed). In other words, the representation used on one interval was

[2] It is worth noting that those children who were not classified as having a linear or a logarithmic representation used non-numerical strategies to perform the task. For instance, some alternated between left and right marks on the lines.

dependent on the one used on the other interval. Children with a more precise representation on the 1-10 interval also had a better representation on the 0-100 interval.

Table II.2.: Type of representation adopted in both tasks by each child.

	0-100 Interval		
1-10 Interval	**None**	**Logarithmic**	**Linear**
None	9	1	1
Logarithmic	4	6	0
Linear	7	15	3

Note. Cell values represent number of children that adopt a given combination of representations as a function of task.

Finally, we investigated the relationship between digit naming and type of representation. The correlations between naming scores (from 0 to 10) and type of representation on the two intervals were significant (interval 0-10: $r_S = .46$, $p < .001$, one-tailed; interval 0-100: $r_S = .46$, $p < .001$, one-tailed). These results suggest that a better knowledge of Arabic numerals goes together with the use of more precise numerical representations – linear for the smaller interval and predominantly logarithmic for the larger interval – in the number to position task.

I.4. Conclusions.

The present study shows that an understanding of how numbers map onto space develops long before formal education begins. Preschoolers relied on a logarithmic representation when confronted with the 0-100 number range and the youngest children showed a trends towards a logarithmic representation even for the 1-10 interval. Thus, the use of a logarithmic representation before a linear one seems mandatory. In contrast, older children deployed a linear representation when confronted with the more familiar range of small numbers (1-10 interval) and the oldest group approximated very closely the ideal positioning. The dissociation between smaller and larger intervals is consistent with the results of Siegler and colleagues (Siegler & Opfer, 2003; Siegler & Booth, 2004) and reveals the coexistence of multiple representations. The youngest age groups in their studies relied on a linear representation for the 0-to-100 number line (our larger interval) whereas on the 0-to-1000 number line only the oldest group of children were able to position numbers linearly like adults. Together, these findings reveal a clear developmental trend with a progressive shift from logarithmic to linear representation.

Logarithmic coding of numbers (Dehaene et al., 2003) is a hallmark of the approximate number system subserving the non-symbolic representation of numerosities (Feigenson et al., 2004; see Dehaene & Changeux, 1993, for a computational model). The finding that the logarithmic fit over the three groups and the accuracies of estimates increased with age suggests a developmental pattern even for the logarithmic representation. Increasing precision of the logarithmic representation is consistent with the finding that the ability to discriminate the numerosity of two sets increases with age (e.g., Lipton & Spelke, 2003).

Lipton and Spelke (2005) have shown that preschool children map the number words within their counting range onto non-symbolic representations of numerosity, but that they show no such mapping for number words beyond that range. Indeed, the youngest children in our study showed a poor and inconsistent performance in the 0-100 interval, whereas they mastered the estimation task in the 1-10 interval. The precision of numerical estimation across all children in our study was correlated with their ability to name single-digit Arabic numbers. This finding highlights the role of the acquisition of numerical symbols in structuring the child’s understanding of numbers, although the exact path that leads from a logarithmic to a linear understanding remains to be understood (for theoretical suggestions, see Verguts & Fias, 2004; Dehaene 2007)

II : Experiment 1b: Numerical Representations and the Weber Fraction.

II.1. Introduction.

In Experiment 1a, results clearly showed that children as young as 3 and a half already start using an intuitive representation of numbers for familiar contexts such as the 1-10 interval. At the age of 4 and a half they already posses two representations for numbers according to the context. Indeed, for the familiar and smaller numerical context, representations are linear whereas when confronted to an unfamiliar context they rely on their intuitive representation. Finally, with the acquisition of more formal numerical knowledge they rely solely on a linear representation. This leads to the hypothesis that if the task is a translation of the internal numerical representation then it should also correlate to other estimates of the internal representation of numbers such as the individual Weber fraction. As introduced in Chapter I (paragraph I.1.2., part B), the Weber fraction allows to estimate the discrimination threshold between two numerosities. The *internal Weber fraction* (*w*) is the width of the Gaussian activation on the mental number line for any given numerosity that best fits the behaviour of the individual in a numerosity comparison task (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004).

The hypothesis tested is straight forward, A more precise internal representation (indexed by a lower Weber value) should correspond to a better performance in the Number-to-Position (NP) task because both describe the same characteristic of the internal representation. To this aim, the same preschool children form Experiment 1a also undertook a numerosity discrimination task.

Participants

The same 46 children (21 females) from experiment 1a also completed the discrimination task. They were divided in three age groups: the youngest group (n=11) had a mean age of 48 months (SD = 4), the middle group (n=16) had a mean age of 60 months (SD = 3) and the oldest group (n=19) had a mean age of 70 months (SD = 3).

Procedure

Two trained female teachers from each school met individually with the children during school hours in a quiet classroom for about half an hour. The familiarity with the teachers helped children to feel comfortable in the testing session. Children were first tested on the NP task and then were presented with a computer generated discrimination task.

The task consisted in two arrays of randomly spread black dots on a white circle presented laterally to a fixation cross. The answer was recorded by pressing on the response key located on the same side of the larger array. No time limit was set but it was asked not to count the number of dots. This task consisted in a modified version of the task presented by Piazza and collaborators (Piazza et al., 2004). On each trial, one of the two arrays was composed of either 16 or 32 dots (reference numerosity). The paired numerosities for the 16 dot reference contained 12, 13, 14, 15, 17, 18, 19 or 20 dots. For the 32 reference, numerosities for the second array were twice as lager as those for reference 16 (24, 26, 28, 30, 34, 36, 38, 40). The second reference was the double of the first in order to confirm that participants’ performance complied with Weber’s Law. Indeed, when the reference is doubled, the discrimination threshold should also be doubled.

Perceptual variables were also controlled. Half of stimuli were controlled for total area whereas the remaining half were controlled for size of dots.

- Stimuli controlled for total dots area (Figure II.2.a.): dots dimension for the non-reference arrays was fixed thus yielding different surfaces (dots area correlating with numerosity). Control for area was realized on the reference arrays. Different versions were realized in order to cover the total surface range occupied by the non-reference arrays.
- Stimuli controlled for size (Figure II.2.b.): dots total area was controlled for the non-reference arrays yielding different dot sizes. To keep area constant, different dot sizes were generated. The reference arrays were thus constructed to cover the range of non-reference arrays sizes.

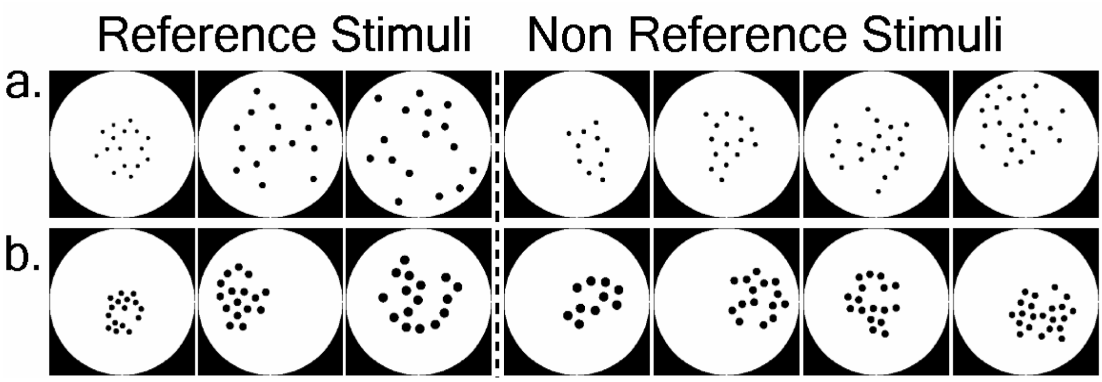

Figure II.2: Reference and non-reference stimuli employed in the numerosity discrimination task. The top row (a.) shows to an example of perceptual control for total area and the bottom row (b.) shows an example of perceptual control for stimulus size.

Considering participants' age, time and length of the task was critical. Therefore only a limited number of trials was run. Only 5 blocks with each of the 16 pairs were thus prepared.

II.2. Results.

The percentages of times a non-reference stimulus was considered as more numerous than the reference were calculated to plot each child's discrimination curve and calculate both logarithmic fit and the Weber value (for the mathematical model and computational details of the procedure please refer to (Piazza et al., 2004) and to the Supplemental Data provided by the authors available at: http://www.neuron.org/cgi/content/full/44/3/547/DC1/).

Response curves were analyzed separately for each group. Figure II.3. Shows the logarithmic fits for both reference numerosities (16 and 32) for each age group. When percentages are plotted on a linear scale, the amplitude of the curve for the larger reference is wider than for the smaller. Indeed, Piazza and colleagues (2004) have shown that for adults the curve for reference 32 is twice as large as the one for reference 16. Moreover, when adults performance is plotted on a logarithmic scale the two curves are identical indicating the logarithmic characteristic of the internal representation. Indeed, children's performance in a log scale tends to become more similar for the two references (Figure II.3. second row).

Finally, the logarithmic fit increases with age and the Weber value (*w*) decreases. Coherently, since the Weber value indicates how precise is the internal representation, it can be concluded that with age it becomes more fine grained. For the younger group, two numbers that differ of about 84% (e.g., 5 and 9, or 100 and 184) are just within one standard deviation of the internal variability. For the oldest age group, two numbers within one standard deviation of internal variability differ of about 48% (e.g., 5 and 7, or 100 and 148).

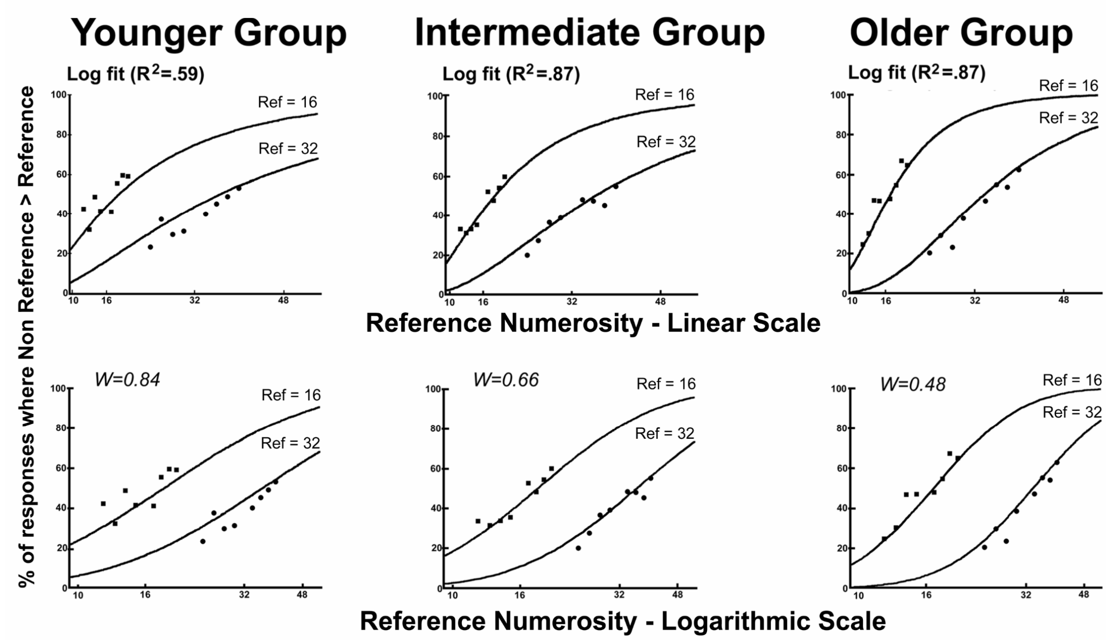

Figure II.3.: Numerosity discrimination task. Graphs represent percentages of trials in which participants responded that the non-reference numerosity was larger than the reference one. Performance is plotted as a function of the amount of numerical deviation. In the top row the scale is linear and curves for the two references are distorted and become progressively symmetrical on a logarithmic scale (bottom row).

To investigate the relation between the Weber value and the Number-to-Position task, each child's Weber value was calculated and correlated with several performance indices for the type of representation in the two Number-to-Position tasks as the slope and both the linear and logarithmic R^2 of individual regressions. None revealed significant.

II.3. Discussion.

The Numerosity Discrimination task revealed that children's internal numerical representation becomes progressively more fine grained with age and experience. This is by decreasing the variability of the internal activation and thus being able to discriminate between sets that have a smaller numerical difference. This is coherent with the results observed in other studies. In infants, for example, the ability to discriminate between sets goes form a ratio of ½ at 6-months of age to a ratio of ⅔ at 9-months (Lipton & Spelke, 2003).

It was expected that *w* would correlate with the representation used in the Number-to-Position task, since it is also a measure of the development of the internal representation. The more a child is familiar and has experience with numbers the more the representation becomes mature.

Two hypotheses could explain the absence of correlation. First, the Weber value in our population had a large variability. Instead, the Number-to-Position task seemed to generate a more coherent set of data for each age group. From a statistical point of view, this is a first limit for finding a correlation.

A second possible explanation is that the two tasks do not tap onto the same component. The Numerosity Discrimination task is highly analogical and does not involve any symbolic format whereas the Number-to-Position task does. Moreover, the latter presupposes the understanding of the symbolic format which might by itself draw upon a different process. Indeed, this task could actually not be a pure translation of an internal representation. The Number-to-Position task might only evaluate the mapping ability of children. A mature representation and knowledge of numbers would be necessary but not sufficient for performing such task. The ability to position numbers requires more formal knowledge than judging the numerosity of two sets.

The question of the specificity of the evaluation of the numerical representation assessed through the Number-to-Position task is therefore addressed with a follow up study presented hereafter (Study 2).

Chapter 3: Development of Numerical and Non-Numerical Sequences.

Ⅰ: Representation of Numerical and Non-numerical Sequences[1]

Abstract

This study investigated the representation of numerical and non-numerical sequences in children from kindergarten to 3rd grade. The development of the mental representations of numbers, letters, and months was studies with a positioning task where children had to position items on a line representing a given interval (Siegler & Opfer, 2003) as well as with a mental bisection task (Zorzi, Priftis, & Umiltà, 2002) where children had to estimate the midpoint of verbally presented intervals. The positioning task had never been used with non-numerical sequences, whereas the mental bisection task had never been used with children, Results of the positioning task showed a similar developmental pattern for both numerical and non-numerical sequences, with a shift from compressive (i.e., logarithmic) to linear positioning. In constrast, in the mental bisection of intervals only the numerical task revealed the typical compressive representation that characterizes the youngest children. The inconsistency in the data obtained with the two tasks is discussed in terms of the mapping between internal (mental) representation and external representational medium, which is a fundamental component in the positioning task but is not required in the mental bisection task. Finally, the results suggest that children learn the concept of linearity in the numerical domain first and progressively extend it to all ordinal sequences.

[1] This experiment is an article in preparation: Berteletti, Lucangeli, and Zorzi.

I.1. Introduction.

An influential theory of numerical representation in humans is based on the notion of a "mental number line". Numbers and numerosities would activate a certain position on the line allowing us to grasp the magnitude perceived (Dehaene, 1992; Dehaene, Piazza, Pinel, & Cohen, 2003). Essentially, this internal representation is assumed to be logarithmically compressed, whereby small numbers would be overrepresented and large numbers would be closer in mental space (Izard & Dehaene, 2007; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004).

This internal numerical representation would be present form the first days of life, long before language acquisition. Several studies have shown that infants and young children are able to make judgments that rely upon numerical information independently form other dimensions (Butterworth, 2005; Bijeljac-Babic, Bertoncini, & Mehler, 1993; Wynn, 1996; Xu & Arriga, 2007). These results have also suggested that the internal representation, although functional since the earliest months of life, might require several years and practice manipulating numbers to develop. For example, studies with infants have shown that numerical ratio between arrays of dots is a predictable variable of performance. At 6 months, a ratio of ½ is necessary for being able to discriminate between the sets whereas at 9 months the ratio drops to ⅔ (Lipton & Spelke, 2003).

During the first years of school children seem to posses multiple numerical representations (Siegler & Opfer, 2003; Siegler & Booth, 2004). Siengler and colleagues developed an estimation task where children had to position a given number on a line. The authors consider it a pure numerical estimation task since it uses numbers as inputs and does not require real-world knowledge of the entities for which properties are being estimated or of conventional measurement units. Moreover, to the contrary of approximate calculation, in their task no mathematical knowledge is required. Lines could go from 0 to 100 or 0 to 1000. Reference numbers were printed on both extreme ends. The analysis of children's estimates indicated that they could position numbers on the line according to 2 main models: logarithmically or linearly. Indeed, smaller children, having less familiarity with numbers, overestimated small numbers on the line whereas they underestimated large numbers yielding a logarithmic fit of the estimates. At the oldest age tested, children properly estimated numbers' positions, which were consequently fit by a linear model. The interesting result was that at intermediate ages, children would position numbers differently according to the interval. When the context was familiar and therefore easy, numbers were linearly positioned but when the context was unfamiliar or hard, children positioned numbers logarithmically. These results led to two main conclusions. First, children possess multiple representations at a given time and the linear representation is not directly generalized to all numbers until the interval becomes familiar. Second, the task was considered as a good tool to map the child's internal representation of numbers.

The developmental shift from an intuitive to a formal representation has been replicated in our previous study (Experiment 1a) with a population of younger children. Indeed, we have tested children from 3-years-old to 6-years-old with number lines 1 to 10 and 0 to 100. The youngest group showed a logarithmic representation for both lines whereas the intermediate group had both representations according to the line and the oldest reached linearity for both. Beside replicating the developmental pattern and showing that before formal education both representations may already coexist at a same time, this experiment has shown that children as young as 3 already possess an intuitive representation and it is functional as soon as they start learning about numbers. Indeed, their estimation ability correlated with performance on a simple number naming task.

In a follow-up study, carried out by Booth and Siegler (2006), the authors of the Number-to-Position task have shown that the estimation abilities on the number line were correlated to a number of other numerical estimation tasks. The aim was to find the same developmental trend, that is a shift from a logarithmic representation to a linear one, with other types of numerical estimation tasks and to assess if the change occurred during the same age period. The different tasks were approximate addition, numerosity estimation, line length in inches and the Number-to-Position task. Performance from kindergarten up to grade 4 showed that the developmental shift across different types of numerical estimation task was comparable (percentage of absolute error decreased with age in a similar proportion for all tasks). Moreover, results showed that the main source of poor estimation is the reliance on a logarithmic representation, because the average fit of the linear function for all tasks improved with age.

All these data converge to the conclusion that the Number-to-Position task is a good numerical estimation task. It has been shown to map the development of children's internal numerical representation and it correlates well with other numerical estimation tasks. But two points have been partially overlooked by the authors. Indeed, no other non-numerical developmental change has been studied and correlated with their task. In fact, it could be argued that the developmental shift observed would not be specific to the development of the sole numerical representation but would be a more general development of all cognitive abilities. Children from preschool to 3rd or 4th grade learn a great deal of information and many fields of knowledge are taught during school hours. Therefore, the results could be general and not specific. Moreover, the transition from an intuitive representation to a linear one could be the consequence of the familiarization with a metric system where the mapping with space is linear. Therefore, it could influence the way a sequence is represented in space.

Although several effects observed in numerical cognition support the logarithmic internal representation of numbers (size and distance effects, Moyer & Landauer, 1967) other sources could explain children's performance. Children could overestimate small numbers simply because they have a limited knowledge of the numerical sequence. That is, those numbers that are spread wide apart on the line could be the only ones they know and on the contrary, those numbers that are outside their counting range are just clustered together towards the end of the line. This pattern could thus be replicated with non-numerical sequences as the alphabet, the months, and the days of the week.

In Experiment 2a we tested children from the last year of kindergarten, 1st grade, 2nd grade and 3rd grade in several positioning tasks. That is, we presented the two classical Number-to-Position tasks as previously done by Siegler and collaborators but also various non-numerical sequences. One was the alphabet line with "A" and "Z" at the extreme ends, a second one was the months line with "January" and "December" at the ends, and the third one was the week line starting with "Monday" and ending with "Sunday". Moreover, to allow a direct comparison with the numerical representation we also prepared equivalent number lines for each non-numerical one. For example, the correspondent numerical line for the week line was marked 1 on the left end and 7 on the right end. These lines would thus make the comparison between the two series straightforward.

In addition to the numerical and non-numerical lines, we tested the same group of children on bisection of verbally presented numerical and non-numerical intervals (Experiment 2b). This task was presented by Zorzi and collaborators (Zorzi, Priftis, & Umiltà, 2002; Zorzi, Priftis, Meneghello, Marenzi, & Umiltà, 2006) to hemi-spatial neglect patients. These patients, after a right parietal lesion, show a spatial deficit for the left-side stimuli. A common test to visual line, these patients systematically misplace the mark to the right as if the part of line on the left was not processed (Halligan & Marshall, 1988; Halligan & Marshall, 1989; Marshall & Halligan, 1989). From this behavioural pattern, the authors tested the assumption that the internal numerical representation resembles a left to right oriented line. To do so, they presented orally several numerical intervals for which patients had to state the midpoint. They observed that the answers tended to be misplaced to the right (e.g., for the interval 1-7 they would answer 6 instead of 4) as if the interval corresponded to a real physical line. This result demonstrates the spatial nature of the mental number line and its isomorphism to a physical line. The authors have tested a second group of patients with numerical and non-numerical intervals. The hypothesis was that if all sequences, therefore not only numerical ones, would be spatially coded in the same way as the numerical one, patients should misplace the midpoint of non-numerical intervals in a similar way. The results to the non-numerical intervals showed very different patterns. When bisecting letter intervals, patients showed a rightward shift of the midpoint compared to controls but it was not modulated by the length of the interval. Instead, when bisecting month intervals, the pattern was different and patients tended to underestimate the midpoint (leftward bias). Therefore results indicate that non-numerical sequences are not mapped onto a spatial representation that has the same characteristics as the numerical representation. The authors suggest, in light of other results, that letters and months would follow a different spatial organization. For example, months could be circularly represented since they have the additional characteristic of being cyclic over time. Moreover, to support the special coding for months and letters, Gevers and colleagues (Gevers, Reynvoet, & Fias, 2004; Gevers, Reynvoet, & Fias, 2003) have observed a spatial coding similar to the SNARC effect (Space-Number Association of Response Codes, Dehaene, Bossini, & Giraux, 1993) for both non-numerical sequences. The association between number and space is considered a strong effect for demonstrating the special coding of the numerical representation. Therefore, if a similar results is observed with non-numerical sequences, these sequences must also have some type of spatial representation. Finally, also neural correlates support these results since both numerical and non-numerical representations activate the same region, namely the horizontal segment of the intraparietal sulcus (Fias, Lammertyn, Caessens, & Orban, 2007) and the consolidation of new ordinal sequences is correlated to the activation in the angular gyrus usually involved in simple arithmetic problems (Van Opstal, Verguts, Orban, & Fias, 2007).

With children three types of intervals were administered: numerical, alphabetical and months. The expected pattern of behaviour would be an increase of precision in estimating the midpoint with age for all types of sequences, but only for numbers the misplacement of the midpoint should follow determined characteristics. Indeed, if numbers to the right end of the interval are compressed and those to the left are expanded, the subjective midpoint should be smaller number than the true one and the underestimation should be stronger for longer intervals. Additionally, performance in the verbal interval bisection task and in the Number-to-Position task should correlate if both measure the same representation.

Finally, children were tested on their knowledge of the different sequences they were tested on. They had to count as far as they could, recite the letters of the alphabet, the months of the year, and the days of the week.

I.1.1. Experiment 2a.

A. Method

<u>Participants</u>

A total of 136 children from 14 different schools of north eastern Italy ranging from the last year of kindergarten to 3rd grade took part in the study. There were 51 preschoolers (27 girls) with mean age of 68 months (standard deviation (*SD*) = 6 months), 28 first graders (15 girls) with mean age of 83 months (*SD* = 4), 35 second graders (18 girls) with mean age of 95 months (*SD* = 4), and 22 third grader (9 girls) with mean age of 105 months (*SD* = 5).

<u>Procedure</u>

Trained teachers from each school met with the children individually during school hours in a quiet classroom for about half an hour. The familiarity with the teachers helped children to feel comfortable in the testing session. Order of experimental tasks was randomly presented to children. Tasks were presented as games to attract children’s attention. For all tasks, no time limit was given and items or questions could be repeated if asked. Children that felt tired or bored were free to stop the study at any time.

Sequences knowledge

All children were tested on their minimal knowledge on numbers, letters, months, and days of the week. They were asked to count as far as they knew for the number sequence (e.g., “Do you know the numbers? Try to tell me all the numbers you know”). If the child reached 30 with success he was stopped. For the other sequences they were asked to give all the letters/month/days they knew (e.g., “Do you know the letters? Try to tell me all the letters you know”).

Number-to-Position task (NP task)

The evaluation of the numerical estimation was done as presented in Siegler and Opfer (2003). Children were presented with 25-cm long lines in the centre of A4 paper. Two different lines were administered: 0-100 and 0-1000. The ends of the lines were labelled on the left by 0 and on the right by either 100 or 1000. In the upper left corner of the sheet was the number to be positioned. The number was not directly above the line avoiding any positioning influence. Numbers to be positioned for the 0-100 line were: 2, 3, 4, 6, 18, 25, 48, 67, 71, 86; and for the 0-1000 line: 4, 6, 18, 25, 71, 86, 230, 390, 780, 810 (corresponding to sets A and B for the same lines used in Siegler & Opfer, 2003). Numbers were completely randomized within each interval. Each line was seen separately from the others to avoid influence from previous positioning. Instructions given at the beginning were: “We will now play a game with the number lines. Look at this page, you see there is a line drawn here. I want you to tell me where some numbers are on this line. When you have decided where the number I will tell you is, I want you to make a mark with your pencil on this line.” To ensure that the child was well aware of the interval size, the experimenter would point to each item on the sheet while repeating for each item: “This line goes from 0 to 100 (1000). If here is 0 and here is 100 (1000), where would you position 50 (500)? The numbers to position were always named verbally by the experimenter for the child.

For both intervals, there was a practice trial that used the numbers 50 and 500 for the smaller and larger interval respectively. It was then possible to check if the task was properly understood. Experimenters were allowed to rephrase the instructions as many times as required without making suggestions of where to place the mark.

Non-numerical lines task

The alphabet, month and week lines as well as the corresponding numerical lines were presented in the same way as the Number-to-Position task. The stimuli used for the Letter-to-Position (LP) task were: “B, E, H, L, N, P, S, V”. For the Month-to-Position (MP) task, children had to position: “February, April, July, September and November”. Finally, for the Day-to-Position (DP) task, every day name except “Thursday” had to be positioned and the line started on “Monday”. The corresponding numerical lines were 1-21 for the Italian alphabet, 1-12 for the months, and 1-7 for the week. We considered to start from 1 instead of

0 since it represented more accurately non-numerical lines that all start with the first element of the sequence. The numbers to position were those corresponding to the non-numerical items (e.g., items “Tuesday”, “B”, and “February” were replaced by number 2).

B. Results.

Sequence knowledge

Mean correct responses (M) and standard deviation (*SD*) were calculated for all recited sequences. The maximum score was 30 for numbers, 21 for letters (according to the Italian alphabet; letters from other alphabets were not considered), 12for months and 7 for days. Two scores were initially calculated. One was the overall number of items reported from that sequence without repetitions and in any order, the other score considered only items given in the correct order, without repetitions and with a maximum gap of 2 in-between items (e.g. “a, b, e, f…” was considered as an acceptable sequence but not “a, b, f…”). In some cases children would give parts of correct sequences intermixed with random items, therefore the score corresponded to the longest correct piece of sequence (e.g., “a, b, c, z, v, d, e, f, g, h” was scored 5 for the sequence part starting with “d” and ending with “h”).

Correlations between the two types of scores were very high therefore only the second and more strict type of scoring was kept. In the following table mean scores and *SD*s are summarized for each class and for the four sequences (Table III.1).

Table III.1.: Mean scores and standard deviations for the Sequences task.

Class	**Numbers (30)**		**Letters (21)**		**Months (12)**		**Days (7)**	
	M	**SD**	**M**	**SD**	**M**	**SD**	**M**	**SD**
Preschool	21.9	8.9	5.9*	6.9	1.9*	3.4	3.4	2.9
1st grade	29.4	1.9	16.8	6.4	9.4	3.8	6.9	0.3
2nd grade	29.5	1.9	19.3	4.6	11.3	2.1	6.9	0.5
3rd grade	30	0	20.3	1.6	11.9	0.5	7	0
Total	26.7	6.7	13.9	8.4	7.5	5.3	5.6	2.4

Note: In parenthesis by each sequence are maximum scores. ***** One participant did not complete the task.

Separate one way analysis of variances (ANOVA) on scores for each sequence were calculated introducing class as a factor. For all sequences, class was significant

(Numbers: $F_{(3,132)} = 19$, $p < .001$; Letters: $F_{(3,131)} = 56$, $p < .001$; Months: $F_{(3,131)} = 102$, $p < .001$; Days: $F_{(3,131)} = 41$, $p < .001$). Post-hoc comparisons highlighted that the significant improvement occurred between preschool and the primary school for all sequences (preschool versus all primary grades: *ps* < .001). Moreover, only for months a significant improvement occurred between 1st and 3rd grade ($p < .023$).

One-tailed Pearson correlation analyses were also conducted between sequences and with class. The a priori hypothesis was that improvement should occur with level of instruction for all sequences. For simplicity, correlation results are displayed in Table III.2. All tasks positively correlate with each other and with class.

Table III.2.: One tailed correlations for all sequences and class.

	Numbers	**Letters**	**Months**	**Days**
Class	.48	.69	.77	.59
Numbers	/	.60	.55	.63
Letters	/	/	.71	.60
Months	/	/	/	.69

Note: all correlations are significant at $p < .001$ one-tailed.

Number-to-Position task

For the Number-to-Position task, analysis were conducted according to the method recommended by Siegler (Siegler & Opfer, 2003; Siegler & Booth, 2004). Analysis on accuracy of children's estimates was computed using the percent absolute error of estimation for each participant. This was calculated as follows:

percent absolute error = (estimate – real value)/line of estimates.

A mixed ANOVA on mean percent absolute error (corrected with the 2*arcsin√(percentage/100) formula) was computed with class as a between-subject factor and type of interval (0-100 and 0-1000) as within-subject factor. Results indicated that both variables introduced were significant (class: $F_{(3,127)} = 37$, $p < .001$; type of interval: $F_{(3,127)} = 348$, $p < .001$). For the 0-100 line, the mean percent absolute error was 17% ($SD = 10\%$) and for the 0-1000, the mean was 32% ($SD = 12\%$). Post-hoc comparison indicated that the estimation progress was significant through all classes (all *ps* < .005) except between 1st and 2nd grade that did not reach significance. Mean percentage absolute error for both lines, 0-100 and 0-1000 were 27% and 38% for preschoolers, 14% and 36% for 1st

graders, 12% and 28% for 2nd graders, and 8% and 20% for 3rd graders. These percentages closely resemble those obtained by Siegler and collaborators in various studies (Booth & Siegler, 2006; Siegler & Booth, 2004) and also our oldest group in Experiment 1a had a mean percentage absolute error of 23% on the corresponding 0-100 line. The interaction also revealed significant, highlighting that the main accuracy change arises at different times according to the line ($F_{(3,127)}$ = 9.47, p < .001). Repeated contrasts indicated a significant improvement of accuracy from preschool to 1st grade and from 2nd to 3rd for the 0-100 line. In contrast, for the 0-1000 line the accuracy improves mostly from 1st to 2nd and from 2nd to 3rd grade (Figure III.1.). Only five children did not complete enough items for these tasks to be included in the analysis.

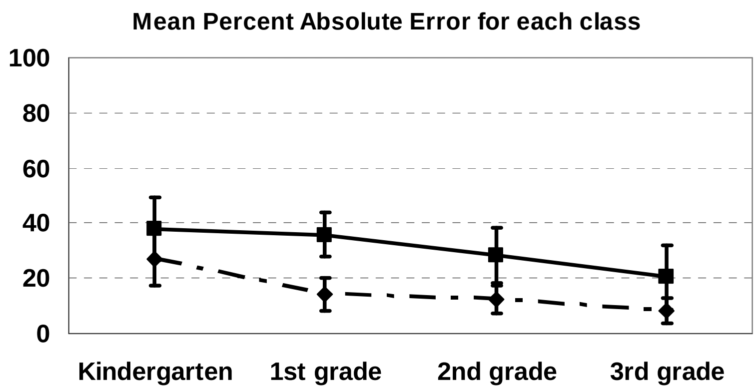

Figure III.1.: Mean percent sbsolute error for each class in the two Number-to-Position tasks. The dashed line corresponds to the 0-100 line and the solid line to the 0-1000 line. The error is smaller for the smaller line and the greatest improvement occurs earlier for the smaller line (preschool to 1st grade) and later for the 0-1000 line (after 1st grade).

To analyse the pattern of estimates, the fit of logarithmic and linear functions were computed on group medians first and then for each individual child. For group medians, the difference between models was tested with two t-tests on the distances between children's median estimates for each number and the predicted values of both linear and logarithmic models (i.e., t-test on the residuals of both models). For preschoolers, both lines were best represented by a logarithmic model (**0-100**: $t_{(9)}$ = 2.86, p < .05, R^2 = .93; **0-1000**: $t_{(9)}$ = 2.56, p < .05, R^2 = .83). For 1st and 2nd graders the t-test was not significant between the linear and logarithmic representation for the 0-100 line but both R^2 were very high (1st graders: linear R^2 = .88, logarithmic R^2 = .98; 2nd graders: linear R^2 = .95, logarithmic R^2 = .96). This indicates that overall children from 1st and 2nd grade have almost completely achieved a linear representation. On the 0-1000 line, performance of both age groups were best represented by the logarithmic function (1st graders: $t_{(9)} = 3.41$, $p < .01$, $R^2 = .93$; 2nd graders: $t_{(9)} = 3.72$, $p < .01$, $R^2 = .99$). Only 3rd graders were clearly linear for the smaller line and started to show some linearity also on the larger line (**0-100**: $t_{(9)} = -3.21$, $p < .05$, $R^2 = .99$; **0-1000**: $t_{(9)} = 2.56$, $\boldsymbol{p > .05}$, linear $R^2 = .85$ and logarithmic $R^2 = .94$). Graphs of median estimates and the best fitting model are presented in Figure III.2.

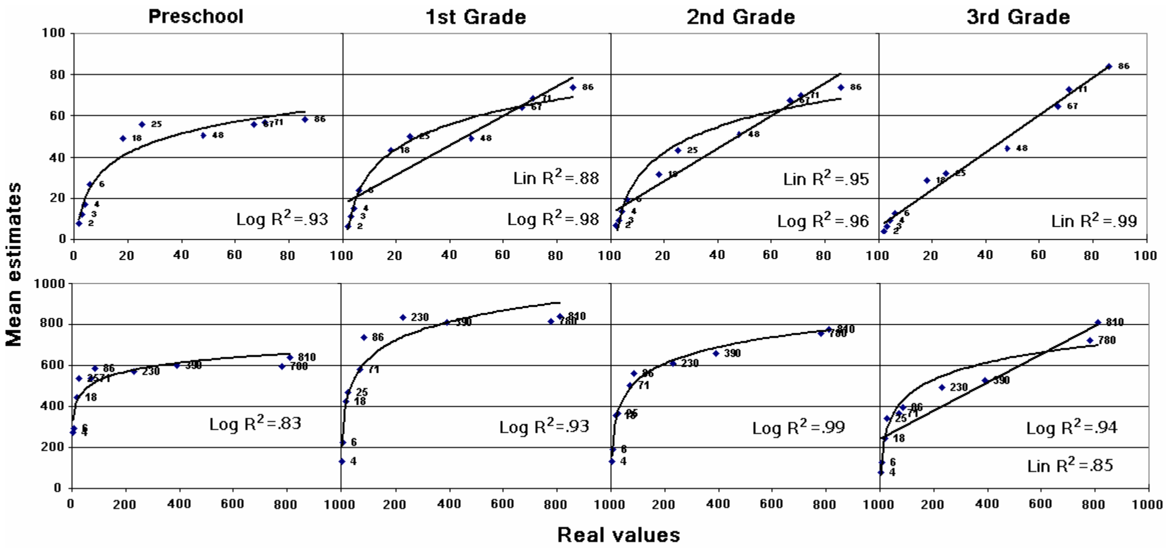

Figure III.2.: Best logarithmic and linear models for each class in the Number-to-Position task. Mean estimates are plotted against real numerical values. The graphs for the 0-100 line are in the top row and those for the 0-1000 are in the bottom row. Groups that did not show a significant difference between the type of positioning have both models plotted.

Fitting individual children's estimates allows to further understand the results found on group medians, and also to grasp developmental patterns. The best fitting model between linear and logarithmic was attributed to each child whenever one was significant. If for example, both models were significant but the logarithmic R^2 was the highest then the child was attributed a logarithmic representation. In the case where both were not significant, the child was considered not to be able to position numbers and probably solved the task using a non-numerical strategy (e.g., alternating left and right marks from trial to trial). For each line, children could be classified as having linear, logarithmic or no representation.

Spearman's rank correlations were computed between class (preschool, 1st, 2nd, and 3rd grade) and type of representation (no representation, logarithmic and linear) separately for each line. Moreover, another correlation was also computed on the representations for both lines. Additionally, a chi-square test was performed to test the random distribution of representations. Since the a priori assumption is that the older the child, the better the familiarity with numbers, all tests were one-tailed. All correlations were significant as well as the chi-square tests (**class & representation on the 0-100 line**: $r_s = .60$, $p < .001$, $\chi^2_{(6)} = 73$, $p < .001$; **class & representation on the 0-1000 line**: $r_s = .49$, $p < .001$, $\chi^2_{(6)} = 43$, $p < .001$; **representation on 0-100 & 0-1000 lines**: $r_s = .50$, $p < .001$, $\chi^2_{(4)} = 70$, $p < .001$). Tables III.3. and III.4. report the percentages of each type of representation as a function of class and type of line.

Table III.3.: Percentages of children adopting a specific representation on each numerical line.

	Type of representation 0-100*			Type of representation 0-1000**		
Class	**None**	**Logarithmic**	**Linear**	**None**	**Logarithmic**	**Linear**
Preschool	50	40	10	53	41	6
1st Grade	4	82	14	11	89	0
2nd Grade	0	62	38	6	82	12
3rd Grade	5	18	77	4	73	23

Note: * Three preschoolers and one 2nd grade child did not complete enough items to fit the models.
** Four preschoolers and two 2nd grade children did not complete enough items to fit the models.

Table III.4.: Percentages of children adopting a given combination of representations on the two numerical line.

	0-100 line		
0-1000 line	**None**	**Logarithmic**	**Linear**
None	68	19	13
Logarithmic	3	66	31
Linear	18	9	73

Note: overall 7 children did not complete enough items on one or both tasks to fit the models.

Non-numerical lines task

For simplicity each non-numerical line with its numerically equivalent line will be presented separately: alphabet (LP), months (MP) and week (DP). All non-numerical lines and their corresponding numerical lines were analyzed following the procedure used for the Number-to-Position task.

Alphabet line (LP) and 1-21 line

A mixed ANOVA on mean percent absolute error (corrected with the 2*arcsin√(percentage/100) formula) was computed with class as a between-subject factor and type of line as a within-subject factor. The main effect class and type of line were significant ($F_{(3,128)}$ = 1666, p < .001; $F_{(3,128)}$ = 6.07, p < .05), highlighting both an improvement in the precision of estimating with age and schooling, and the better performance for the 1-21 line compared to the LP. Mean percent absolute error for the alphabet was 17% (SD = 12%) and for the 1-21 line was 14% (SD = 11%). Post-hoc comparisons revealed that the main difference occurred between preschool and the other 3 classes (all ps < .001). Moreover, only for preschoolers there was a significant difference in the precision of estimating the items on the two lines (27% for letters against 22% for numbers). For the other classes the difference was not significant (1st and 2nd graders: 12% and 10 % respectively for letters and for numbers; 3rd graders: 10% for both types of items). Four children were not included in the analysis because they did not complete enough items of each line.

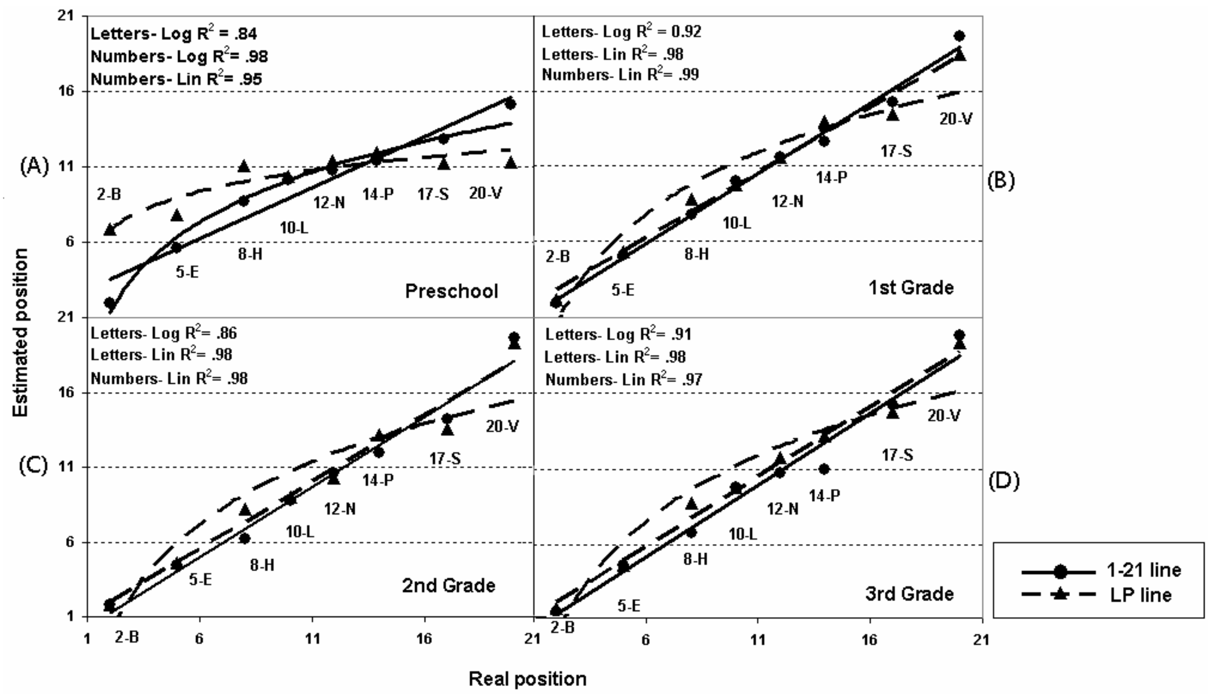

Figure III.3.: Best logarithmic and linear models for each class in the LP and its numerical equivalent. Mean estimates are plotted against real numerical values. The models fit on the letter estimates are represented with dashed lines and triangle whereas the models fit on the number estimates are represented with solid lines and dots.

To analyse the pattern of estimates, the fit of logarithmic and linear functions were computed on group medians first and then for each individual child as done for the Number-to-Position task. Preschoolers performance on the LP were best represented by a logarithmic function ($t_{(7)}$ = 2.77, p < .05) whereas for the numerical counterpart the difference between the two types of representations was not significant indicating that with such a small interval children from preschool already start to master a linear positioning. Indeed, in our previous study (Experiment 1a) we have shown that children of the same age were perfectly linear on a 1-10 line. In 1st, 2nd and 3rd grade, the way children positioned letters was equally well fit by a linear and a logarithmic model while numbers were predominantly positioned linearly (1st grade: $t_{(7)} = 2.88$, $p < .05$; 2nd grade: $t_{(7)} = -2.75$, $p < .05$; 3rd grade: $t_{(7)} = -3.26$, $p < .05$; Figure III.3.).

A type of representation was attributed to each child using the same criteria as those used for the Number-to-Position task allowing to further understand developmental patterns. Indeed, one could argue that mean estimates could not truly represent individual performance. Therefore, for each children and for each line (alphabet and numerical), a linear, logarithmic or no representation could be assigned (Table III.5.).

Table III.5.: Percentages of children adopting a specific representation on each type of line (LP and 1-21 line).

	Type of representation LP*			**Type of representation 1-21****		
Class	**None**	**Logarithmic**	**Linear**	**None**	**Logarithmic**	**Linear**
Preschool	66	17	17	49	24	27
1st Grade	11	25	64	4	25	71
2nd Grade	6	20	74	0	11	89
3rd Grade	5	27	68	4	14	82

Note: * Five children form preschool did not complete enough items to calculate the models.
** Two children from preschool did not complete enough items to calculate the models.

One tailed Spearman's rank correlations were computed between class and type of representation separately for each line as well as between the representations to both lines. Additionally, a chi-square test was performed to test the random distribution of representations. All correlations were significant as well as the chi-square test (**class & representation on the LP**: $r_s = .54$, $p < .001$, $\chi^2_{(6)} = 54$, $p < .001$; **class & representation on the 1-21 line**: $r_s = .55$, $p < .001$, $\chi^2_{(6)} = 54$, $p < .001$; **representation on LP & 1-21 line**: $r_s = .62$, $p < .001$, $\chi^2_{(4)} = 69$, $p < .001$).

Months line (MP) and 1-12 line

A mixed ANOVA on mean percent absolute error (corrected with the 2*arcsin√(percentage/100) formula) was computed with class as a between-subject factor and type of line as a within-subject factor. Both main effects were significant as well as the interaction. The older the class the better the accuracy in positioning months and numbers ($F_{(3,126)}$ = 1523, p < .001). Overall performance was more accurate for the numerical line ($F_{(3,126)}$ = 5.36, p < .005). Post-hoc comparisons again revealed a significant difference between preschoolers and the three primary classes. The significant interaction ($F_{(3,126)}$ = 5.29, p < .005) indicated a larger improvement in positioning months compared to positioning numbers. Preschoolers were more accurate for numbers than months but by 1st through 3rd grade children were more accurate in positioning months. In separate analysis, only 1st and 3rd graders were significantly more accurate in positioning months compared to numbers ($F_{(1,26)}$ = 6.69, p < .05; $F_{(1,21)}$ = 10.95, p < .005) although the analysis on preschoolers was also close to significance ($F_{(1,45)}$ = 3.7, p = .061). Six children overall did not complete a sufficient amount of items to be included in the analysis.

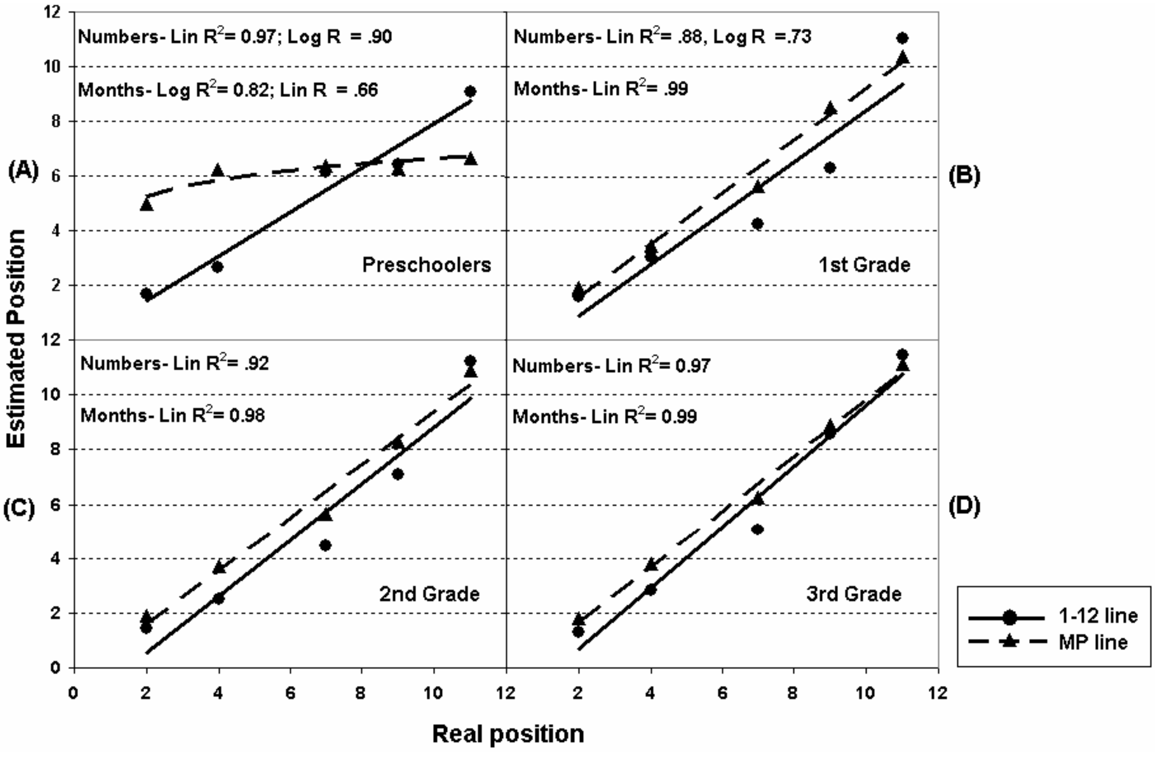

Figure III.4.: Best logarithmic and linear models for each class on the MP and its numerical equivalent. Mean estimates are plotted against real numerical values. The models fit on the months estimates are represented with dashed lines and triangles whereas the models fit on the number estimates are represented with solid lines and dots. For clarity of the graph, only the model with the highest R^2 is shown although not always significantly different from the other R^2.

Subsequently, the fit of logarithmic and linear functions were computed on group medians first and then for each individual child. The t-test on the logarithmic and linear

models for preschoolers' performance, on both months and equivalent numerical line, did not reach significance indicating that the data could be explained just as well by both models (Figure III.4.). Nevertheless, a slight better R^2 was found for the logarithmic fit when positioning months and a better linear R^2 was found for the equivalent numerical line (Months: logarithmic $R^2 = .82$; linear $R^2 = .66$; 1-12 line: logarithmic $R^2 = .90$; linear $R^2 = .97$) First graders positioned months linearly whereas both models explained equally well the numerical estimates (Months: $t_{(4)} = -3.74$, $p < .05$). For the two older classes, both tasks were best fit by a linear model (**2nd graders**: months $t_{(4)} = -3.41$, $p < .05$: 1-12 line $t_{(4)} = -2.83$, $p < .05$; **3rd graders:** months $t_{(4)} = -3.72$, $p < .05$; 1-12 line $t_{(4)} = -3.7$, $p < .05$).
Again, each child was attributed a representation for the MP and the 1-12 line. Therefore, for each line, children could have a linear, logarithmic or no representation (Table III.6). One tailed Spearman's rank correlations were computed between class and type of representation separately for each line as well as between the representations for both lines. Additionally, a chi-square test was performed to test the random distribution of representations. All correlations and the chi-square tests were significant (**class & representation on the MP**: $r_s = .37$, $p < .001$, $\chi^2_{(6)} = 28$, $p < .001$; **class & representation on the 1-12 line**: $r_s = .56$, $p < .001$, $\chi^2_{(6)} = 67$, $p < .001$; **representation on MP & 1-12 line**: $r_s = .45$, $p < .001$, $\chi^2_{(4)} = 31$, $p < .001$).

Table III.6.: Percentages of children adopting a specific representation on each type of line (MP and 1-12 line).

	Type of representation MP*			Type of representation 1-12**		
Class	**None**	**Logarithmic**	**Linear**	**None**	**Logarithmic**	**Linear**
Preschool	58	6	36	83	2	15
1st Grade	15	11	74	22	14	64
2nd Grade	17	6	77	3	26	71
3rd Grade	14	14	72	18	9	73

Note: * Three children form preschool and one from 1st grade did not complete enough items to calculate the models. ** Five children from preschool did not complete enough items to calculate the models.

Week line (DP) and 1-7 line

The mixed ANOVA on mean percent absolute error (corrected with the 2*arcsin√(percentage/100) formula) was computed with class as between-subject factor and type of line as within-subject factor. Only main effects reached significance (class: $F_{(3,129)} = 2239$, $p < .001$; type of line: $F_{(1,129)} = 70$, $p < .001$). Overall, mean percent absolute error for positioning the days of the week was 16% (SD= 11) and 24% (SD = 12%) for numbers. Children were therefore more accurate in positioning days of the week than numbers. Once more, post-hoc comparisons showed a marked improvement between preschool and the other 3 primary classes (all *ps* < .005). Only three children had do be removed from the pool of participants for not positioning enough items on the lines.

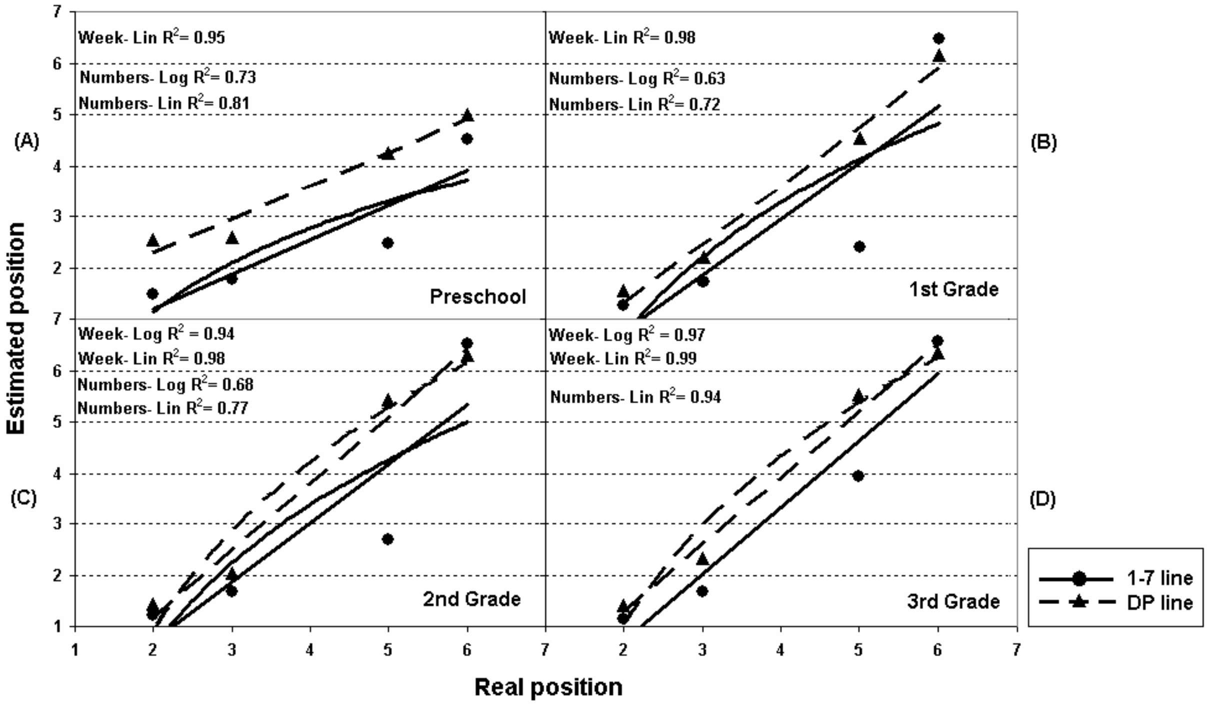

Figure III.5.: Best logarithmic and linear models for each class on the DP and its numerical equivalent. Mean estimates are plotted against real numerical values. The models fit on the week estimates are represented with dashed lines and triangles whereas the models fit on the number estimates are represented with solid lines and dots.

Logarithmic and linear fits over median estimates per class resulted in contradictory results. On the one hand, only in 3rd grade, children were significantly more linear in positioning numbers on the 1-7 line ($t_{(4)}$ = -3.91, p < .05). For the three younger classes the positioning was fit equally well by both linear and logarithmic models even if the R^2 was highest for the linear model. On the other hand, days of the week were fit significantly better by the linear representation for the two younger classes (preschool: $t_{(4)}$ = -4.22, p < .05; 1st grade: $t_{(4)}$ = -4.9, p < .05) whereas both models fit equally well median estimates made by the older group. An explanation for the awkward result on the numerical line is that the numerical interval of 1-7 is very unusual. As a matter of fact, children at that age have already shown linearity on longer lines as the 0-100 in this experiment and on the 1-10 line in our Experiment 1a. Indeed, in the first years of school, children learn using with a 10-base reference. Placed in a task were the reference was 7, children were impaired in "resizing" their line. The other lines, 1-21 and 1-12, were much closer to familiar lines compared to the 1-7 line. This is also evident when visualizing the position of item number "5" (Figure III.5.). In general, all items for the numerical line have been underestimated but this item has been systematically and more strongly underestimated distorting model fits in all classes with the exception of 3rd grade. This particular underestimation and non linearity of fits is contradictory with the ability of 1st, 2nd and 3rd graders to estimate the positions of numbers linearly on larger lines (i.e., 0-100, 1-21 and 1-12).

To further support this hypothesis, models made on single children showed an inability to position numbers across all groups. Up to 3rd grade, more than 40% of children were not able to position numbers adequately and were categorized as having no representation (Table III.7.). Moreover, the linear representation is more or less absent until 3rd grade where a big jump appears. Indeed, 50% percent of the children position numbers linearly highlighting an all-or-none pattern.

Since the 1-7 line was misrepresentative of children abilities, the one tailed Spearman's rank correlations was computed only between class and type of representation in the DP. Additionally, a chi-square test was performed to test the random distribution of representations. Both the correlations and the chi-square test were significant (**class & representation on the DP**: $r_s = .45$, $p < .001$, $\chi^2_{(6)} = 36$, $p < .001$).

Table III.7.: Percentages of children adopting a specific representation on each type of line (DP and 1-7 line).

	Type of representation DP*			Type of representation 1-7**		
Class	**None**	**Logarithmic**	**Linear**	**None**	**Logarithmic**	**Linear**
Preschool	67	6	27	84	6	10
1st Grade	43	0	57	64	11	25
2nd Grade	17	6	77	77	20	3
3rd Grade	9	18	73	41	9	50

Note: * Three children form preschool did not complete enough items to calculate the models.

** One child from preschool did not complete enough items to calculate the models.

Regression analyses with the Sequences knowledge tasks and type of representation on the different lines (numerical and non-numerical).

To understand the role of specific sequence knowledge in the ability to position numbers, letters, months and days of the week, numerous regressions were performed introducing both class and score on the appropriate sequence as predictors. Indeed, the better a child knows the sequence of numbers the better he might position numbers on the lines. This was the assumption made by Siegler and collaborators (Siegler & Opfer, 2003; Siegler & Booth, 2004) when they claim that the familiarity with the numerical context influences children's ability to position numbers either intuitively or linearly. In our analysis, class was introduced to partial out the general effect of schooling. Table III.8. presents the percentages of variance explained by each predictor introduced in single regressions.

Table III.8.: Percentages of variance explained for each regression, with class as first predictor and score on the different sequences as second predictor.

1st Predictor	**0-100**	**0-1000**	**LP**	**1-21**	**MP**	**1-12**	**DP**
Class	36%	20%	28%	29%	14%	30%	20%
2nd Predictor							
Numbers score	39%	26%		39%		37%	
Alphabet score			38%				
Months score					22%		
Days score							26%
Δ of R^2 explained	3%	6%	10%	10%	9%	7%	6%
N participants	130	128	129	132	127	129	132

Note: In each row are represented the first and second predictors, in each column the percentage of variance explained for each task. The penultimate row indicates the percentage of unique variance explained by the second predictor and the last row the number of participants introduced in the analysis. All predictors were significant with p-values < .01.

Although percentage of unique variance explained by the second predictor was small, it was never the less always significant. This indicates that even if a core part of the variance is explained by a general improvement due to years of instruction, the specific knowledge of each sequence influences the ability to position items on the line in a formal and more appropriate way. It is interesting to highlight how children implicitly tend to position non-numerical sequences linearly even if this is not required. In fact, nothing specifies that the distance between "A" and "C" is twice the one between "D" and "E". The same is true for months and days of the week. At this point we could offer two possible, but opposite, explanations: first, children generalize the linearity principle to all sequences whether or not they are numerical as long as the context is familiar enough, and second, logarithmic and linear positioning are not the consequence of a numerical representation but a general phenomenon: the more items of the sequence are known, the more they are equally spread onto the line.

Table III.9.: Percentages of variance explained for each regression, with class as first predictor, score on the appropriate non-numerical sequence as second predictor and finally type of representation for the equivalent numerical line.

1st Predictor	**LP**	**MP**	**DP**
Class	28%	14%	20%
2nd Predictor			
Alphabet score	38%		
Months score		22%	
Days score			26%
Δ of R^2 explained	10%	9%	6%
3rd Predictor			
Representation 1-21	48%		
Representation 1-12		26%	
Number score			29%
Δ of R^2 explained	10%	4%	3%
N participants	129	127	132

Note: In each row are represented the first, second, and third predictors, in each column the percentage of variance explained for each task. The penultimate row indicates the percentage of unique variance explained by the third predictor and the last row the number of participants introduced in the analysis. All predictors were significant with *p*-values < .05. For the DP the representation for the 1-7 line was not representative, therefore the predictor was substituted with the score in the Number Knowledge task.

To discern between these two explanations, another set of regressions was computed to establish if the type of representation on the equivalent numerical line would explain a significant part of the ability in positioning items on non-numerical lines. Indeed, if the type of representation for a non-numerical sequence is not the consequence of a generalisation of the numerical representation then it should only be explained by abilities in that specific field and not by numerical abilities. As in the previous set of regressions, class was the first predictor introduced, followed by the knowledge of the non-numerical sequence and finally by the type of representation for the numerically equivalent line. Only for the WL the third predictor was the score in the Numerical knowledge task since the 1-7 numerical line was biased and not representative of children's performance (Table III.9.).

Percentages of unique variance explained by the third predictor were significant, indicating that the ability to position numbers can still predict the ability to position non-numerical items even though the effect of specific sequence knowledge was partialled out. This result supports the possible generalization hypothesis where children learn linearity for numbers first and then generalize the principle to other sequences.

C. Discussion Experiment 2a.

To summarize the data, we have observed that the general knowledge for sequences improves with class and the greatest improvement occurs between kindergarten and 1st grade. This result is in agreement with the beginning of formal teaching at school. In the Number-to-Position (NP) task, the developmental pattern observed by Siegler and collaborators (Siegler & Opfer, 2003; Siegler & Booth, 2004) and in Experiment 1a has been replicated. Children initially overestimate small numbers and progressively shift to an equidistant positioning of estimates. For the 0-100 line, preschoolers where either unable to position numbers or were best represented by a logarithmic positioning. First and 2nd graders were in-between the logarithmic and the linear positioning whereas the majority of 3rd graders were clearly linear. For the 0-1000 line, from kindergarten to 2nd grade, positions of estimates was best described by a logarithmic representation and only by 3rd grade some ability to position numbers linearly appeared. Additionally, we directly tested improvement in both lines and observed that the strongest increase in precision for the smaller line (0-100) occurred in 1st grade whereas for the larger line (0-1000) it occurred from 2nd grade on.

Comparisons between the Letter-to-Position (LP) task and the 1-21 line showed that, although accuracy increased for both lines, performance in positioning numbers was higher. Analysis on individual representations for the LP task indicated that the majority of preschoolers did not possess enough knowledge to position letters. From 1st to 3rd grade, even if a fourth of the children overestimated the first letters of the alphabet, the majority were able to position them linearly. This rapid shift, from kindergarten to 1st grade, could be explained by the fact that the experiment was carried out at the end of the school year. Consequently, 1st graders, compared to preschoolers, had an intensive practice on letters. This explanation is supported by the significant increase in score on the letter sequence task (from 6 to 17 letters named in correct order by the end of 1st grade). For the equivalent numerical line, the classical developmental trend was observed although some preschoolers were already able to position numbers linearly. Indeed, children of the same age group in Experiment 1a were able to position numbers linearly on a 1-10 line.

Performance in the Months-to-Position (MP) task and in the 1-12 line was also characterized by a general improvement as a function of class. Numbers were positioned linearly starting from kindergarten but for months there was an all-or-none pattern. Preschoolers did not reach a logarithmic positioning whereas from 1st grade on they all positioned months linearly. Again, this shift could be the consequence of the late testing period that gave pupils almost a year of practice.

The Days-to-Position task and the numerical equivalent line (1-7) showed a peculiar pattern of performance. Although the greatest improvement occurred in 1st grade, conversely to the other tasks, the non-numerical sequence was solved more accurately. Performance in the 1-7 line was poor in general and only 3rd graders were able to position numbers linearly. The other classes were not even characterized by a logarithmic fit of the estimates. This result is contradictory with the better performance observed in tasks with larger intervals. A reasonable explanation is that numbers are always thought in school using a base-10 reference system. The particular interval used probably induced children in error by confusing their scheme of reference. In support, 3rd graders were either unable to position numbers (41%) or had a linear positioning (50%) with almost no in-between performance. Moreover, in 2nd grade only 3% of the children positioned numbers linearly. Probably, only by 3rd grade children have had enough practice to master numbers independently form the teaching reference and are able to overcome the confusion. Furthermore, children have a better performance in the longer intervals that are closer to the base-10 reference system. This result is nevertheless interesting because it highlights the importance of familiarity with the context for achieving the task and knowing the numbers is not sufficient to be able to position them according to a linear representation.

Finally, the ability to position items was a function of school years and specific knowledge. It appeared that once the number of years in school was partialled out, the knowledge of a specific sequence significantly explained the ability in positioning items in the corresponding line task. Interestingly, non-numerical lines were positioned linearly although nothing in the sequence itself suggests it (i.e., the distance between A and B needs not to be the same as between C and D, as well as B is not larger than A). Therefore it could be that children generalize the linearity concept to all sequences. Indeed, the ability to position numbers in the equivalent line still explained a part of unique variance once class and specific knowledge were partialled out.

Overall, in this experiment we have observed that logarithmic and linear positioning of estimates is not restricted to numerical sequences. Moreover, the ability to position numbers predicts the ability to position non-numerical sequences supporting the assumption that children generalize the linearity principle to all ordinal sequences.

I.1.2. Experiment 2b.

A. Method.

Participants and Procedure

The same participants of Experiment 2a also undertook the tasks from Experiment 2b. Moreover, the same procedure was followed.

Mental bisection of numerical intervals

Stimuli and procedure strictly followed those of Zorzi and collaborators (Zorzi et al., 2002). From their stimuli we only used the forward intervals considering the backward task too challenging for young children. The 48 forward number intervals were randomly presented verbally to the each child. Lengths of intervals were: three (e.g., 1-3), five (e.g., 1-5), seven (e.g., 1-7), and nine (e.g., 1-9). Each interval length was presented in the units range (e.g., 1-5), the tens (e.g., 11-15), and the twenties (e.g., 21-25). Children were asked to say what number was in the middle of each number interval (i.e., “What number is in the middle between 1 and 9?”). In case children used their fingers to solve the problem, the experimenter kindly discouraged them from doing so.

Mental bisection of non-numerical intervals

Also for the non-numerical intervals only the forward items from Zorzi and collaborators (Zorzi et al., 2006) study were used again considering the backward task too challenging. Twenty-two verbally presented letter intervals (e.g., L-T) were randomly and verbally presented. The length of the interval was three (e.g., L-N), five (e.g., L-P), seven (e.g., L-R), and nine (e.g., L-T). Participants were asked to say what letter occupied the midpoint position for the interval. (i.e., “What letter is in the middle between P and T?” Correct answer: “R”).

For months intervals, 16 forward stimuli were used and randomly presented. Interval lengths were three (e.g., April-June), five (e.g., April-August), seven (e.g., April-October), or nine (e.g., April-December). Participants had to name the month that was the midpoint of each interval (i.e., “What month is in the middle between April and August?” Correct answer: “June”).

B. Results.

Analysis on response types in the mental bisection of intervals

For the mental bisection tasks the first analysis aimed at understanding children’s response pattern. Indeed, compared to adults, children had a much wider range of responses including answers out of the given interval or even answers from other domains (e.g., respond “3” for the interval “April-August”). Therefore, answers were first classified in 5 distinct categories: correct answers (CA), interval answers (IA), scale answers (SA), aberrant answers and finally no answer. Since the percentages for the two last classifications were overall very low, they have been merged into the other answers (OA) category. IA included all the answers that were not correct but were inside the interval given for that trial (e.g., “May” for the interval “April-August” when the correct answer was “June”). SA were all the answers that fell inside the total range covered by all the items in that domain. For example, number intervals included numbers going from 1 to 29, therefore, a SA could have been “19” for interval 11-17. Finally, aberrant answers were all answers outside the total covered range and even those answers from another domain (e.g., respond “1000” or “B” for interval 1-9). Percentages of each type of answer were calculated individually for each child and corrected with the $2*\arcsin\sqrt{(percentage/100)}$ formula.

This classification allows to qualitatively understand the development of the ability to solve the task. Analyzing percentages of absolute error would have reduced the amount of information from the task. Indeed, we predict that in later classes and with familiarity with the domains investigated, children would give more CA and error patterns would shift from OA or SA to IA.

Numerical intervals

Separate analysis of variance were computed for each type of answer (CA, IA, SA, OA) introducing class (preschool, 1st, 2nd, and 3rd grade), interval length (3, 5, 7, and 9) and numerical size (units, tens and twenties). Figure III.6. displays percentages of each type of answer as a function of class.

For CA, the three main effects were significant (class: $F_{(3,144)} = 233$, $p < .001$; interval length: $F_{(3,144)} = 626$, $p < .001$; numerical size: $F_{(2,144)} = 8.79$, $p < .001$). Percentages of CA increased with class (**preschool** $\mu = 17\%$, *SD* = 14%; **1st grade** $\mu = 48\%$, *SD* = 30%; **2nd grade** $\mu = 59\%$, *SD* = 25%; **3rd grade** $\mu = 62\%$, *SD* = 26%). With increasing interval length and numerical size, percentages of CA decreased (**interval 3** $\mu = 71\%$, *SD* = 26%; **5** $\mu = 32\%$, *SD* = 17%; **7** $\mu = 21\%$, *SD* = 14%; **9** $\mu = 22\%$, *SD* = 15%; **units** $\mu = 50\%$, *SD* = 30%; **tens** $\mu = 45\%$, *SD* = 31%; **twenties** $\mu = 25\%$, *SD* = 20%). All two-way interactions were significant. The numerical size of intervals influenced more preschoolers compared to 3rd graders ($F_{(6,144)} = 2.96$, $p < .05$). Performances were better for the units intervals but by 3rd grade, percentages of CA were similar throughout all numerical sizes. The interaction length of the interval by class was significant ($F_{(9,144)} = 9.92$, $p < .001$). For the smallest interval accuracy increased more with class compared to the other three. In 1st grade performance already reached 80% of CA. The other intervals also improved with class but not so strongly and never reached 50% of CA. The interaction length of interval by numerical size was also significant ($F_{(6,144)} = 2.23$, $p < .05$). The best solved interval was the one with a length of three in the units range (almost 80% accuracy). All the other combinations did not reach 50% of accuracy. Finally, the three way interaction class by interval length by numerical size was also significant ($F_{(18,144)} = 2.51$, $p < .001$). Overall, this interaction describes a gradual improvement of accuracy for interval length and numerical size throughout classes.

For IA the same ANOVA was computed. All three main effects were significant indicating a slight increase in IA with class ($F_{(3,144)} = 22$, $p < .001$; $\mu = 20\%$, *SD* = 16%; $\mu = 35\%$, *SD* = 29%; $\mu = 29\%$, *SD* = 26%; $\mu = 30\%$, *SD* = 27% respectively from preschool to 3rd grade) and interval length ($F_{(3,144)} = 835$, $p < .001$; $\mu = 3\%$, *SD* = 3%; $\mu = 44\%$, *SD* = 12%; $\mu = 55\%$, *SD* = 14%; $\mu = 54\%$, *SD* = 16% respectively from interval 3 to 9), while a slight decrease for numerical size ($F_{(2,144)} = 9.8$, $p < .001$; $\mu = 31\%$, *SD* = 27%; $\mu = 28\%$, *SD* = 26%;

μ = 26%, *SD* = 24% respectively for units, tens and twenties). Only the class by interval length interaction reached significance showing an almost absolute absence of IA for the length of 3 because only the two extremes of the interval could count as such answers, and an improvement with class for the interval of length 5 ($F_{(9,144)} = 15$, $p < .001$). Moreover, the two way interaction numerical size by class was close to significance ($F_{(6,144)} = 15$, $p = .052$): preschoolers gave more IA for intervals composed of units but this difference disappears in the older children. The three way interaction was also close to significance ($F_{(18,144)} = 1.65$, $p < .055$).

Class and numerical size reached significance also for the SA (**class**: $F_{(3,144)} = 47$, $p < .001$; μ = 24%, *SD* = 7%; μ = 7%, *SD* = 5%; μ = 12%, *SD* = 6%; μ = 5%, *SD* = 5%; respectively from preschool to 3rd grade; **numerical size**: $F_{(2,144)} = 23$, $p < .001$; μ = 9%, *SD* = 10%; μ = 12%, *SD* = 9%; μ = 15%, *SD* = 9% respectively for units, tens and twenties) showing a decrease with class for answers outside the interval. More interestingly, the numerical size effect shows that the larger the numerical size of the interval the greater the tendency to give SA responses, probably due to a less extensive knowledge of the larger numbers. Indeed, if a child does not know the sequence of numbers between 25 and 29 he might just guess a number that is twenty something (e.g., 21). Moreover, the interaction numerical size by interval length was significant ($F_{(9,144)} = 2{,}27$, $p < .05$). That is, the percentage of SA decreased with interval length only for the smallest numerical size (i.e., units), showing a better knowledge for the smallest numbers.

The percentage of OA significantly decreased with class ($F_{(3,144)} = 344$, $p < .001$; μ = 39%, *SD* = 12%; μ = 11%, *SD* = 4%; μ = 0%, *SD* = 0%; μ = 3%, *SD* = 3% respectively from preschool to 3rd grade). Additionally, the larger the numerical size of the interval the higher the percentages of OA ($F_{(2,144)} = 3.87$, $p < .05$; μ = 10%, *SD* = 11%; μ = 14%, *SD* = 19%; μ = 15%, *SD* = 18% respectively for units to twenties). But this effect was modulated by the class, because only preschoolers showed a real advantage for units compared to larger numerical sizes ($F_{(6,144)} = 2.66$, $p < .05$). This shows a progressive acquisition of numbers with instructions.

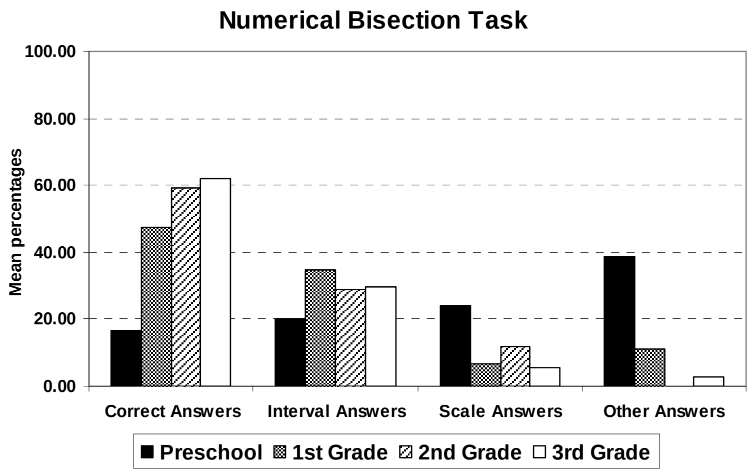

Figure III.6.: Mean percentages for each type of answer as a function of class in the Numerical Bisection Task.

Letters intervals

Separate ANOVAs for each type of answer were computed with class and interval length as independent variables. Figure III.7. displays percentages of each type of answer as a function of class.

CA increased with class ($F_{(3,72)}$ = 124, p < .001; µ = 6%, *SD* = 5%; µ = 25%, *SD* = 22%; µ = 42%, *SD* = 28%; µ = 44%, *SD* = 23% respectively from preschool to 3rd grade) and decreased with interval length ($F_{(3,72)}$ = 186, p < .001; µ = 55%, *SD* = 28%; µ = 22%, *SD* = 14%; µ = 15%, *SD* = 10%; µ = 10%, *SD* = 9% from interval length 3 to 9). The interaction class by length of interval was significant ($F_{(9,72)}$ = 11, p < .001). The smallest interval was also the one that had the strongest increase with class followed in order of length by the other 3 intervals.

All effects and interaction were also significant for IA. With class and with interval length the percentage of IA increased significantly (**class**: $F_{(3,72)}$ = 39, p < .001; µ = 16%, *SD* = 9%; µ = 33%, *SD* = 24%; µ = 40%, *SD* = 30%; µ = 42%, *SD* = 28% respectively from preschool to 3rd grade; **interval length**: $F_{(3,72)}$ = 287, p < .001; µ = 3%, *SD* = 4%; µ = 40%, *SD* = 15%; µ = 51%, *SD* = 19%; µ = 58%, *SD* = 21% from interval length 3 to 9). Furthermore, IA to the smallest interval were constant (only two possible IA), whereas, for all the other intervals they increased with class ($F_{(9,72)}$ = 12, p < .001).

SA to the letters intervals slightly decreased with class showing an improved knowledge of letters ($F_{(3,72)}$ = 13, p < .001; µ = 14%, *SD* = 6%; µ = 15%, *SD* = 9%; µ = 7%, *SD* = 5%; µ = 10%, *SD* = 7% respectively from preschool to 3rd grade). As well as a decrease with interval length ($F_{(3,72)}$ = 27, p < .001; µ = 16%, *SD* = 7%; µ = 12%, *SD* = 7%; µ = 9%, *SD* = 5%; µ = 3%, *SD* = 3% from interval length 3 to 9). Indeed, the larger the interval the less the possible answers outside the scale. Finally, <u>OA</u> significantly decreased with class ($F_{(3,72)} = 256$, $p < .001$; µ = 64%, *SD* = 4%; µ = 26%, *SD* = 7%; µ = 11%, *SD* = 45%; µ = 5%, *SD* = 4% respectively from preschool to 3rd grade).

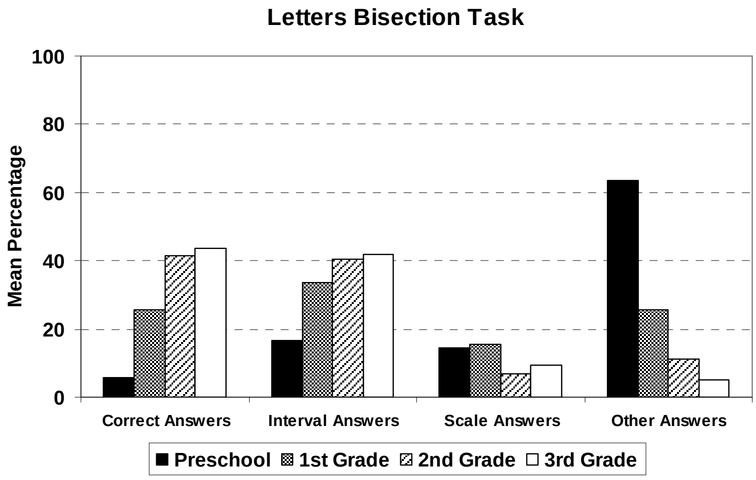

Figure III.7.: Mean percentages for each type of answer as a function of class in the Letters Bisection Task.

Months Intervals

Separate ANOVAs for each type of answer were computed with class and interval length as independent variables. Figure III.8. displays percentages of each type of answer as a function of class.

Children became more accurate <u>(CA)</u> with class ($F_{(3,48)} = 56$, $p < .001$; µ = 8%, *SD* = 5%; µ = 42%, *SD* = 29%; µ = 58%, *SD* = 30%; µ = 61%, *SD* = 28% respectively from preschool to 3rd grade) and less accurate with interval length ($F_{(3,48)} = 129$, $p < .001$; µ=66%, *SD* = 33%; µ = 28%, *SD* = 16%; µ = 19%, *SD* = 16%; µ = 17%, *SD* = 16% from interval length 3 to 9). The interaction highlighted an increase in performance more salient for the smallest interval length with class whereas it was less marked for all other intervals ($F_{(9,48)} = 9$, $p < .001$).

<u>IA</u> showed a slight increase with class ($F_{(3,48)} = 15$, $p < .001$; µ = 14%, *SD* = 9%; µ = 29%, *SD* = 26%; µ = 33%, *SD* = 32%; µ = 32%, *SD* = 28% respectively from preschool to 3rd grade) but greatly increased with interval length ($F_{(3,48)} = 197$, $p < .001$; µ=3%, *SD* = 3%; µ = 39%, *SD* = 18%; µ = 52%, *SD* = 19%; µ = 56%, *SD* = 21% from interval length 3 to 9). For the interaction between the two variables, again the percentages of IA were constant for all classes for the small interval (3) and increased for the other intervals with class ($F_{(9,48)} = 11$, $p < .001$).

Percentages for the SA were very low since the total range covered by the items was of nine over the twelve possible months. Even so, both main effects were significant indicating first that with class, children decreased in the percentage of answers outside the interval ($F_{(3,48)}$ = 4.88, $p < .005$; µ = 15%, *SD* = 6%; µ = 9%, *SD* = 6%; µ = 7%, *SD* = 7%; µ = 4%, *SD* = 9% respectively from preschool to 3rd grade) and second, the percentage of SA also decreased with interval length ($F_{(3,48)}$ = 3.69, $p < .001$; µ = 10%, *SD* = 8%; µ = 11%, *SD* = 9%; µ = 6%, *SD* = 5%; µ = 0%, *SD* = 0% from interval length 3 to 9). This second effect is even more obvious since the possible outside answers were only of three for interval item "April-December", possible SA: January, February and March).

The remaining type of answers (OA) decreased with class and slightly increased with interval length (**class**: $F_{(3,48)}$ = 141, $p < .001$; µ = 63%, *SD* = 4%; µ = 19%, *SD* = 4%; µ = 2%, *SD* = 3%; µ = 4%, *SD* = 4% respectively from preschool to 3rd grade; **interval length**: $F_{(3,48)}$ = 3.3, $p < .05$; µ = 21%, *SD* = 25%; µ = 22%, *SD* = 26%; µ = 23%, *SD* = 25%; µ = 27%, *SD* = 27% from interval length 3 to 9).

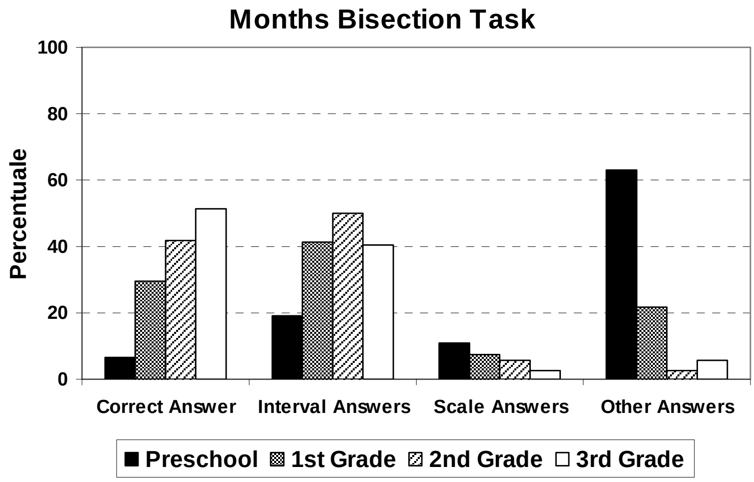

Figure III.8.: Mean percentages for each type of answer as a function of class in the Months Bisection Task.

Representations to the mental bisection of intervals

Further analyses aimed at understanding the type of representation used to solve the task. Zorzi and collaborators (Zorzi et al., 2002; Zorzi et al., 2006) observed that patients' bisected mental intervals in the same way as they bisected physical lines. The subjective midpoint was systematically overestimated and it was modulated by length of the interval. For the non-numerical intervals, Zorzi and collaborators observed different patterns. For the letters intervals, patients' performance was not modulated by the length of the interval and for the months intervals, they tended to underestimate the true midpoint. These results suggested that these sequences are not represented on a similar internal space as the numerical representation.

According to these results, children were expected to bisect numerical and non-numerical intervals in different ways. If numbers are distributed logarithmically onto a mental line, children should systematically underestimate the true midpoint. Indeed, smaller numbers are more spread apart than larger ones. Moreover, as a consequence the larger the interval the stronger the underestimation. With a progressively better representation of numbers, the underestimation should disappear. For the other sequences no such pattern was expected. Finally, precision was also supposed to increase with class for all sequences.

Because children gave an heterogeneous range of answers, it was necessary for the following analysis to consider only those answers that were in the interval range. Indeed, considering a number too far outside the range of the interval would have biased the overall result. Moreover, a second criteria of 70% of accuracy was set in order to consider only those children that properly understood task as well as ensuring a sufficient number of responses for the analysis. In all following analyses, degrees of freedom were corrected with Huynh-Feldt method whenever the assumption of sphericity of data was violated.

Numerical intervals

At a group level, mean differences between observed and correct (dO-C) answers were calculated for each class. A mixed ANOVA with interval length as within-subject variable, class as between-subject variable and decade as covariate showed a main effect of class ($F_{(3,7)} = 5.89$, $p < .05$) and an interaction class by interval ($F_{(2.8,19.8)} = 3.42$, $p < .05$; Figure III.9.). The larger the interval the larger the leftward deviation from the true midpoint but this was modulated by the class. In other words, the tendency to answer a smaller number than the true midpoint is stronger for larger intervals but this deviation attenuates with class (µdO-C: -1.2, -0.41, -0.15 and 0.04 respectively from preschool to 3rd grade). Preschoolers were the group that had the strongest underestimation bias with interval length. Older children were progressively more accurate and less influenced by interval length.

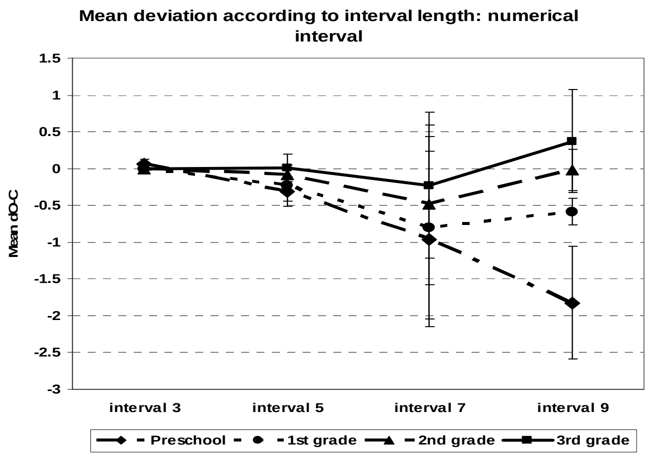

Figure III.9.: Mean difference between observed and correct answer (dO-C) per class for each interval across numerical size. A negative value indicates that the observed answer was smaller than the correct answer.

Letters intervals

To obtain a adequate number of data points for the analysis, the mean dO-C were calculated for each subject and then introduced in a mixed ANOVA. Therefore, for the letter interval analysis, the within-subject variable was interval length and the between-subject variable was class (Figure III.10). Results showed no significance for class but only for interval ($F_{(1.9,117.2)} = 7.71$, $p < .001$). Longer intervals had an overall stronger leftward bias (mdO-C: -.03, -.13, -.51, -.8). No interaction reached significance.

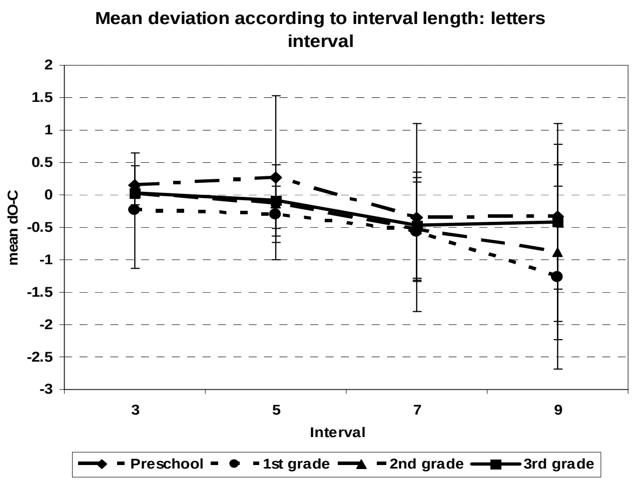

Figure III.10.: Mean difference between observed and correct answer (dO-C) to the Letter Bisection task per class for each interval length. A negative value indicates that the observed answer was smaller than the correct answer.

Months intervals

The same analysis as for the letters intervals was carried out with the data for the months intervals. The mixed ANOVA with interval length as within-subject variable and class as between-subject variable did not reach significance (Figure III.11).

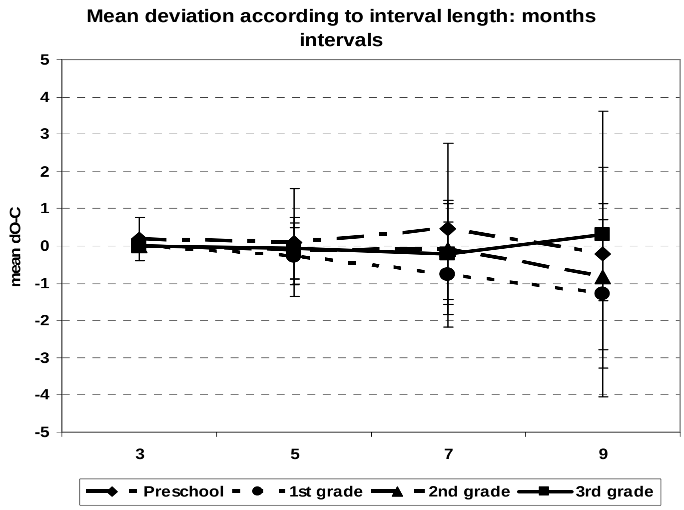

Figure III.11.: Mean difference between observed and correct answer (dO-C) to the Months intervals task per class for each interval length. A negative value indicates that the observed answer was smaller than the correct answer.

C. Discussion Experiment 2b.

Common conclusions can be drawn following the analysis on type of answers for each interval. First, the number of correct answers increases with the class and decrease with the length of the interval. Indeed, the older children had more years of practice for each type of domain and the larger the interval the harder the task. Moreover, the error patterns were described by a large number of answers outside the interval at the youngest age group that progressively became interval answers. This pattern shows the knowledge improvement in the domain. For example, a child that does not master well numbers might be tempted to guess the answer when faced with an interval outside his range of knowledge (e.g., “28” for interval “21-25”). Moreover, for the numerical intervals, the larger the numerical size of the interval the lower the accuracy. Indeed, children first started to master units, then tens and finally also twenties. Finally, to the contrary of the numerical domain which is virtually infinite, the alphabet and the months had less scale answers especially with increasing length of the interval. For example, when a child was presented with the interval “April-December” there were only three possibilities of falling outside the interval. For the numerical bisection task answers could go well beyond the maximum number presented (i.e., 29, some children answered 70, 100 and 1000).

In the analysis on the representation used in the mental bisection of intervals we observed different patterns according to the type of sequence. The numerical bisection task shows a clear underestimation as a function of class and interval length. Preschoolers underestimated the subjective midpoint more strongly than the older groups. Moreover, the degree of underestimation was stronger with longer intervals. This result reflects the hypothesis of a logarithmic distribution of numbers on the mental number line that would progressively improve with practice. Moreover, this pattern parallels the numerical representations observed with the Number-to-Position task. Indeed, the more immature is a child’s numerical representation the more the midpoint of an interval is underestimated.

The letter interval bisection shows a general underestimation that is influenced by the length of the interval but does not seem to attenuate with practice. Interestingly, the adult control group from the study carried out by Zorzi and colleagues (2006) shows a tendency to underestimate that is modulated by interval length although it is not significant. Finally, performance in the bisection of month intervals does not seem to change with practice nor with interval length.

Overall, the patterns observed in the three bisection tasks confirm the claim made by Zorzi and collaborators (2006) that each sequence seems to have its own type of bisection bias and therefore its own mental representation. Only the performance in the numerical bisection task follows the clear characteristics of a logarithmic distribution and significant changes with practice.

I.1.3. Combined analysis for Experiments 2a and 2b

In addition to the previous findings, children’s performance to numerical intervals, was expected to be a function of their internal representation. The type of positioning adopted in the Number-to-Position task should be related to the degree of underestimation in the mental bisection task. For the other sequences, although performance to the positioning task seemed to follow the developmental pattern found for the Number-to-Position task, the correlation with the bisection task was not expected. As observed, the type of bias was different for the non-numerical sequences compared to the numerical one.

Numerical intervals, Number-to-Position and number sequence

To test for a relation between the performance in the mental bisection of numerical intervals and performance in other tasks, regressions for each individual were computed. Introducing mean dO-C for each interval at each numerical size as dependent variable and interval length as predictor. Standardized betas (β) were extracted and used as dependent variable in a step-wise regression where class, counting score and the type of representations in the Number-to-Position task (both 0-100 and 0-1000 lines) were introduced as predictors. The only predictors that had a significant contribution were class and counting score ($F_{(2,78)}$ = 8.97, $p < .001$; $R^2 = .19$). The type of representation for both Number-to-Position intervals did not explain a significant part of unique variance.

Letters intervals, Letter-to-Position and alphabet sequence

For each individual, the regression slope (β) was calculated on mean dO-C and interval length. This index of the deviation tendency from the true midpoint was introduced into a stepwise regression with score in the alphabet task, type of representation in the Letter-to-Position task and class as predictors. None of these reached significance.

Months intervals, Month-to-Position and moths sequence

Individual regression slopes (β) were calculated on mean dO-C and interval length as deviation index. The stepwise regression with score in the months sequence, representation to the Months-to-Position taks and class as predictors did not reach significance.

Finally, a correlation analysis between all mental bisection tasks (number, letters and months) was run on individual slopes to observe for a common individual deviation tendency. However, none reached significance.

Percentage absolute error analysis for the mental bisection and the positioning tasks.

A last correlational analysis was performed to test if the precision of estimation was similar across tasks, as previously done by Booth and Siegler (they correlated the Number-to-Position task to other 3 estimation tasks; 2006). The mean absolute percentage of errors was calculated for both the positioning tasks and the mental bisection of intervals. Simple correlations on the percentage of errors for the number interval bisection and for the 0-100 NP task were significant (r = .469, p < .001, one-tailed). The correlations for the non-numerical sequences were also significant (letter intervals and Letter-to-Position: r = .563, p < .001, one-tailed; for months intervals and the Months-to-Position: r = . 323, p < .005, one-tailed). However when class was introduced as covariate, none of the above correlations remained significant nor did the correlation between the 0-100 and the 0-1000 lines. This covariate was not introduced in Booth and Siegler analysis which could have also explained the results obtained in their study.

In summary, results to the combined analysis showed that the bisection bias was not correlated to the way children positioned items on the line tasks. Moreover, only the mental bisection of numerical intervals was explained by class and counting score indicating a decrease in the bias with practice. Finally, the absolute percentage of error for the two types of tasks correlated although it was completely accounted by the class (years of school).

I.2. General discussion.

The first aim of this study was to test the developmental pattern observed by Siegler and collaborators (Siegler & Opfer, 2003; Siegler & Booth, 2004) with non-numerical sequences. Indeed, it was suggested that the logarithmic positioning of numbers on the numerical intervals was the consequence of a limited knowledge of the numerical sequence and with increasing knowledge the numbers would be positioned linearly. If this were true, than the same developmental pattern should have appeared with non-numerical patterns. Therefore, children from kindergarten to 3rd grade were tested on both numerical and non-numerical positioning tasks. Results in the Number-to-Position task replicated previous findings, that is, a shift from an immature representation to a linear one. This was also true for the Letter-to-Position, the Month-to-Position and the Days-to-Position task. Children in the youngest groups were not able to position items according to any model (logarithmic or linear) but by 3rd grade were able to position them linearly. The only difference was that the logarithmic representation was less noticeable at individual level for probably two main reasons. First, the intervals were much smaller than the numerical ones and second, the testing period occurred at the end of the school year allowing extensive practice to 1st graders compared to preschoolers. Additionally, representations used in the Number-to-Position task were explained by both class and specific knowledge (counting score). For the non-numerical sequence the same was true but an additional unique variance was explained by the type of representation in the Number-to-Position task. Therefore, this leads to the suggestion that the type of representation used with non-numerical sequences is a consequence of the ability in the numerical domain. Children would probably learn the linearity principle in the numerical domain and extend the concept to other non-numerical sequences.

Finally, we have observed that linearity is not just a consequence of number knowledge. When children were faced with a small and unusual interval, that is an interval inside their range of numerical knowledge but with an unfamiliar limit (i.e., 1-7), even the older children experienced difficulty in positioning numbers. Indeed, for being able to accomplish the task, children had to overcome the predominant base-10 reference scheme used in all teaching methods and predominant in the metric system (the International System of Units).

The second aim of this experiment was to compare performance in two tasks that tap onto the same representation. Zorzi and collaborators (Zorzi et al., 2002; Zorzi et al., 2006) demonstrated that the mental representation of numbers resembles a physical line and influences the performance on a mental bisection task. Therefore, if children possess an immature representation of numbers this should be observed in their answers. The initial hypothesis was that an immature representation overrepresents small numerosities and underrepresent large ones resulting in an underestimation performance in the mental bisection task. For the non-numerical interval, no specific bias was expected.

Overall, performance in the bisection task of numerical and non-numerical intervals improved with class and was modulated by the length of the interval. A progressive shift of answers from outside to inside the given interval was observed highlighting an increasing knowledge of the sequence. Indeed, if the sequence is not well mastered it is harder to guess an item inside the interval.

Moreover, percentage of error in estimating the midpoint of the interval was correlated to percentage of error in placing the items in the positioning tasks. The more they were precise in positioning an item on a line the more they were precise in estimating the true midpoint in the mental bisection task. However, once the class was partialled out from the analysis, the correlation between the two disappeared. This should be considered when correlating the Number-to-Position task with other estimation tasks. In their work, Booth and Siegler (2006) did not consider this variable when testing children from kindergarten to 4th grade.

An interesting result is found when analyzing the type of representation used in the mental bisection task. Children showed different patterns according to the type of interval to bisect. Numerical intervals were more underestimated by younger children but there was a strong modulation of interval length. This result complies with the hypothesis of a logarithmic distribution in the mental representation of numbers. Conversely, letters only showed a general underestimation tendency that was not modulated by class and months did not show any specific deviation bias. Individual performance in the bisection of the different intervals were not correlated indicating that children did not posses the same bias for all sequences and only performance in the numerical bisection were correlated with class but more importantly with counting score.

In light of these results, it seems that children generalize linearity to all sequences once it is achieved in the numerical domain. This is supported by the different deviation patterns for the different sequences in the mental bisection task, by the unique variance explained in the regression analysis of the positioning tasks and by the fact that non-numerical sequences show an all-or-none pattern in the positioning task. This result is in line with previous findings of spatial representation for non-numerical sequences. Indeed, if early in childhood a child learns to generalize the linear representation to all ordered sequences it is reasonable to expect SNARC-like effects (Gevers et al., 2004; Gevers et al., 2003) or even activation in the same brain regions (Fias et al., 2007; Van Opstal et al., 2007)

Moreover, it appears that the mental bisection task is a more straightforward task compared to the Number-to-Position task since it does not require a mapping between the mental and the visual representations showing more clearly the difference between sequences. The mapping onto a visual line probably induces the use of strategies tied to the numerical or metric system.

Chapter 4: Numerosity Processing in Adults.

Ｉ: **Experiment 3: Numerosity discrimination underpinning approximate calculation**

Abstract

The relation between the ability to discriminate numerosities, as described by the Weber fraction, with more complex abilities was investigated. Indeed, if the Weber value describes the quality of individuals' numerical representation it should be related to the accuracy observed in more complex numerical tasks requiring its activation. Therefore, sixteen adults were tested on three tasks: numerosity discrimination, number comparison and approximate calculation. In the latter task, participants viewed two operands (as well as the operator) and then had to choose between two alternative results. The format of operands and results was manipulated by crossing symbolic (digits) and non-symbolic (dots) notations, with the exception of the digit-digit condition that was not presented to avoid exact calculation. The proposed results were always composed of the correct solution and one distracter (ratio of ± 1.5 and 1.25). Results showed that the formats of both problem and answers influenced the overall estimation accuracy. A classical distance effect was found and the closest distracters were harder to reject than the distant ones. Moreover, a general operational momentum effect was found although this effect was modulated by the format combination. When the problem was presented in dots and the answer in digits, the tendency was to underestimate for both operations whereas in the reversed format combination (digits to dots), the general tendency was to overestimate the correct answer. The classical operational momentum, overestimating when adding and underestimating when subtracting, was only found in the dots-dots format combination. These results are discussed according to previous findings showing that large collections of dots are generally underestimated (Ginsburg, 1978; Allik & Tuulmets, 1991; Durgin, 1995) and an alternative account, not mutually exclusive with the operational momentum, is considered. Finally, the relation between the Weber fraction and the other tasks is discussed.

I.1. Introduction.

Numerous studies have demonstrated that adults and infants are endowed of various numerical abilities. Since the earliest months of life, humans are able to discriminate sets of objects only relying on numerical information (Feigenson, 2005; Xu & Spelke, 2000; Lipton & Spelke, 2003; Xu & Arriga, 2007) as well as anticipating numerical outcomes of non symbolic approximate operations (Koechlin, Dehaene, & Mehler, 1997; Barth, La Mont, Lipton, & Spelke, 2005; Barth et al., 2006; McCrink, Dehaene, & Dehaene-Lambertz, 2007; McCrink & Wynn, 2004). Performance, at all ages, in numerical tasks is described by a number of characteristics that all follow Weber's law: first, the ratio between numerosities influences discrimination performances; second, the larger the numerosity the greater the variability of responses with a constant coefficient of variation (CV = Standard deviation / mean).

Dehaene, among others, has suggested the hypothesis that the numerical representation would be similar to a mental line where numerosities follow a logarithmic distribution (Dehaene, 1989; Dehaene, 1992; Izard & Dehaene, 2007). More specifically, the representation of numerosity was formally described by a log-Gaussian model (Izard & Dehaene, 2007; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004), where numbers are placed on the line according to a log scale and correspond to a Gaussian distribution with fixed variability. Therefore, the compressive nature of the number line would account for the increasing overlap of activation between larger numerosities accounting for the ratio effect. The precision of discrimination is directly measured by the ratio necessary to discriminate two numerosities. The usual performance threshold is 75% of correct discrimination for a given numerical distance. Then, mathematically, it is possible to estimate the *internal Weber fraction* (*w*) which refers to the width of the Gaussian distribution of activation on the internal number line. This value, determines the precision in subjects' performance: a smaller width of the Gaussian corresponds to less overlap in numerosity representation and thus a more precise discrimination ability.

Piazza and collaborators designed an habituation task where participants viewed a reference numerosity (16 or 32) for at least three consecutive trials followed by a fourth trial with either the same numerosity (25% of cases) or a deviant numerosity (ratio of ± 1.25, 1.5 or 2). Participants simply had to give a same/different judgement. Stimuli were constructed in order to avoid the use of non numerical cues as shape of items (triangles or dots), density, size and layout. Proportion of “different” answers followed a classical U-shaped function and it conformed to the predictions of the compressed number line model. Indeed, when plotted on a linear scale for numerosity, the performance curves were asymmetrical and twice as broad for reference 32 compared to reference 16. However, if the curves where plotted on a logarithmic scale they became symmetrical and Gaussian with a fixed width. In a second version of the task, participants had to judge if the fourth stimuli was larger or smaller than the previous ones. Again, results confirmed those obtained in the first task, performance followed a sigmoid curve and the width of the slope was twice as large for the larger reference. Moreover, the average curves for both tasks were centred on the habituation number, suggesting that participants could extract accurate numerical information. From a neuroanatomical point of view, they also observed that the only regions that responded to numerical distance were localized in the left and right intraparietal sulci (IPS, including the horizontal segment) confirming and extending data observed both with single cell recording in animals and in neuroimaging experiments with humans.

Cantlon, Brannon, Carter and Pelphrey (2006) replicated Piazza’s experiment and addressed the question of the early development of brain areas for processing abstract numerical information. By testing both adults and 4-years-old children they observed, besides replicating the adults findings, that the changes in activations for numerosity were similar in both groups. This data confirms that the neural correlate of numerical cognition is active early in development prior to formal instruction and symbolic experience.

Also results obtained in approximate calculation experiments with children and adults comply to Weber’s law (McCrink & Wynn, 2004; Barth et al., 2006; Barth et al., 2005; Pica, Lemer, Izard, & Dehaene, 2004). In the most recently published study on approximate calculation with adults (McCrink et al., 2007), the authors have shown that variability was found to increase with problem size and participants were more likely to consider a response as correct if the ratio between the proposed and true outcome was closer to one. In their experiment, participants viewed movies where a first sets of dots moved behind an occluder and a second set could join or leave the one already hidden (addition or subtraction). Participants had to predict the outcome and then compare it with a proposed solution. Three larger and three smaller incorrect solutions where created using three ratios (1.25, 1.5, 2). Interestingly, not only they found that the response pattern overall complied with Weber’s law but also fount that when solving additions the tendency was to overestimate the outcome and the reverse when performing subtractions. They termed this effect the *operational momentum* effect for its analogy with the *representational momentum*. The latter indicates a perceptual effect for which an observer remembers the final position of a moving target (implied or apparent) as shifted in the direction of the motion (Freyd & Finke, 1984; for a review see T.L. Hubbard, 2005). Moreover, this effect was also found with abstract concepts of represented motion as change in pitch (Freyd, Kelly, & DeKay, 1990). In addition to the observation that the internal numerical representation resembles a line, supported by the SNARC effect (Dehaene, Bossini, & Giraux, 1993) and neuropsychological data (Zorzi, Priftis, & Umiltà, 2002), the authors argue that realizing an addition or a subtraction would be comparable to a displacement on the mental number line in the direction suggested by the operation.

The work of McCrink and has been recently followed up by Knops, Viarouge and Dehaene (submitted). In their study, they investigated in more detail the psychophysical laws of approximate mental arithmetic with both symbolic and non symbolic problems. The goals were to understand the influence of numerical magnitude of the operands on the *operational momentum* and whether this effect holds for symbolic format. In two experiments participants had to estimate the answers of basic problems (additions and subtractions) by choosing among seven proposed outcomes. Overall, results complied with Weber's law. Indeed, participants were relatively accurate in solving the problems, showed a consistent bias to underestimate (stronger for non symbolic operations) and the classical scalar variability with a constant CV were also observed. Furthermore, they replicated the *operational momentum* effect, although it was weak for the symbolic format, and showed that it is influenced by the size of the operands: as the size of the outcome increases the effect becomes stronger.

All these studies lead to the conclusion that different numerical competences show the signature of Weber's law: ratio effect, scalar variability and constant CV. Therefore, comparison and approximation should rely on a same representation and individual performance in both tasks should correlate. Moreover, since these abilities are present early in development it is reasonable to assume that the representation on which they rely is also the basis of formal numerical knowledge and competences. In agreement to a common representation between non symbolic and symbolic processing of numerical information, distance and size effect have been found in many symbolic tasks such as number comparison (Dehaene, 1992; Dehaene & Akhavein, 1995; Verguts & De Moor, 2005; Reynvoet & Brysbaert, 1999; Girelli, Lucangeli, & Butterworth, 2000). To our knowledge, no study has yet demonstrated a direct relation between these two types of numerical abilities.

Therefore, the aim of the present study was to find a link between the various competences that would rely on a common representation. Indeed, as described in Piazza et al's study (2004), the interval Weber fraction describes the individuals' internal representation and therefore should be predictive of his/her performance in all numerical tasks.

In this study, we adopted a modified version of the numerosity discrimination task used in Piazza and collaborator's study in order to obtain a value of the Weber fraction for each participant (henceforth Weber value, *w*). Then, the same group of participants was administered both the corresponding number comparison task and an approximate addition and subtraction task. With these three tasks, both the relation with the symbolic processing of numerical information and the ability in approximating calculations should be highlighted if all competences rely on a common numerical representation. In our approximate calculation task we also manipulated the presentation format. Symbolic and non symbolic formats were crossed for the problem and the outcomes. The only condition with both problems and outcomes in symbolic format was excluded since it would have resulted in exact calculation.
Finally, up to date, the stability of the Weber value has not been established. Crucially, the Weber value has to prove stable for each participant to be considered a real index of individual numerical representation. Therefore, the numerosity discrimination task was administered twice in two separate sessions.

I.1.1. Method.

Participants

Seventeen university students (9 females) took part to the study and were paid for their collaboration.

General procedure

Each participant was tested individually in a dimly lit and silent laboratory. The experiment was subdivided in two sessions of about an hour each. Testing conditions were kept as similar as possible by setting the two sessions exactly one week apart at approximately the same hour. The approximation task was subdivided in two conditions, addition and subtraction, counterbalanced between participants. Both sessions started with the numerosity discrimination task and were followed by one of the two conditions of the approximation task. The number comparison task was administered only at the end of the first session. Moreover, participants were asked to report the number of hours of sleep and the quantity of coffee consumption before the first session and were required to monitor these parameters for the second session.

All tasks were presented on a 17 in. screen (1024x768 pixels) at 105 cm from the chin rest. E-Prime software (version 1.1.4.4., 2003) was used for stimuli presentation and responses were recorded using the PST Serial Response Box. For all tasks, participants used only two out of five buttons situated at the extreme ends with the ipsilateral index finger.

Numerosity comparison

Two versions of this task were constructed and counterbalanced between participants to ensure that the same stimuli were not seen twice. Each version was composed of 320 plus 10 practice trials subdivided in four blocks. Only practice trials received feedback on accuracy. Each trial started with two white disks (on black background) on each side of a fixation point for 1400 ms. Sets of black dots appeared in each white disk and disappeared when the answer was recorded. The initial screen appeared again for 100 ms (Figure III.2.). Instructions were to choose the most numerous set without counting. Both reaction time and accuracy were stressed.

This task consisted in a modified version of the task presented by Piazza and collaborators (Piazza et al., 2004). On each trial, one of the two arrays was composed of either 16 or 32 dots (reference numerosity). The paired numerosities for the 16 dots reference contained 12, 13, 14, 15, 17, 18, 19 or 20 dots. For the 32 dots reference, numerosities for the second array were twice as larger as those for reference 16 (24, 26, 28, 30, 34, 36, 38, 40). The second reference was the double of the first in order to confirm that participants' performances complied with Weber's Law. Indeed, when the reference is doubled, the discrimination threshold should also be doubled.

Perceptual variables were also controlled. Half of stimuli were controlled for total area whereas the remaining half were controlled for size of dots.

- Stimuli controlled for total dots area (Figure IV.1.a.): dots dimensions for the non-reference arrays was fixed thus yielding increasing areas with increasing numerosity. Control for dots area was realized on the reference arrays. Different versions were used in order to cover the total area range occupied by the non-reference arrays.
- Stimuli controlled for size (Figure IV.1.b.): dots total area was controlled for the non-reference arrays yielding different dot sizes. To keep surface constant in the reference stimuli, different dot sizes were generated. The reference arrays were thus constructed to cover the range of non-reference arrays sizes.

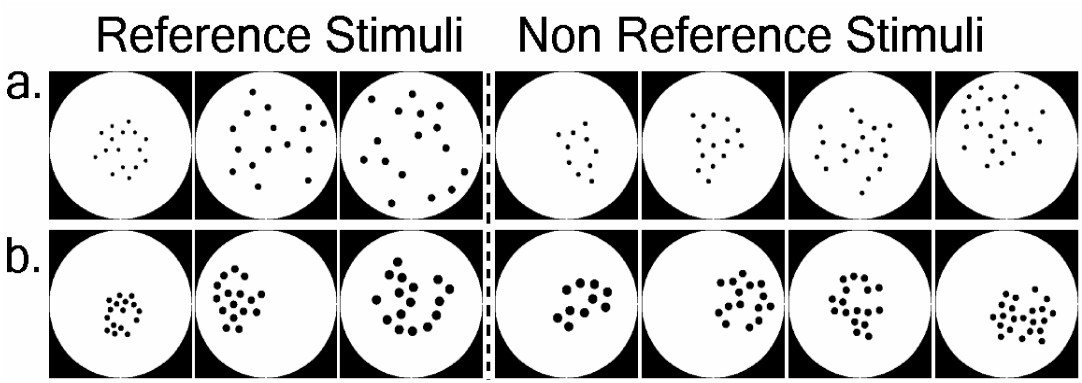

Figure IV.1.: Reference and non-reference stimuli employed in the numerosity discrimination task. The top row (a.) shows to an example of perceptual control for total area and the bottom row (b.) shows an example of perceptual control for stimulus size.

Finally, four different configurations of each stimulus where created and side of presentation was counterbalanced giving a total of 320 stimuli (2 references x 10 distances x 2 versions x 4 configurations x 2 presentation sides). Total size of stimulus on the screen occupied 5.5 degrees of visual angle.

Number comparison

This task was the symbolic version of the numerosity comparison task. The same 20 pairs were presented inside a white disk on a black background and the same procedure was followed (Figure IV.2.). The 160 trials (2 references, 10 distances, 2 presentation sides x 4 repetitions) were divided in four blocks and preceded by 6 practice trials. The double digits numbers occupied 7.5 cm on the screen, equivalent to 4 degrees of visual angle.

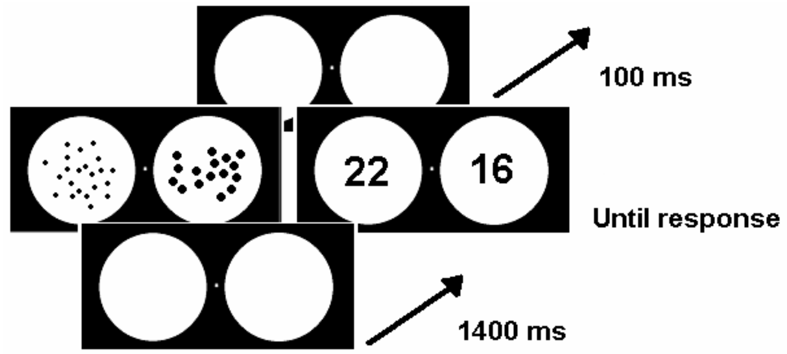

Figure IV.2.: Example of a trial for the symbolic (on the right) and non symbolic (on the left) comparison task.

Approximate addition and subtraction

Additions and subtractions were presented separately but the procedure was the same for both conditions. The tasks was to approximate the solution to the problem presented either as symbols (digits) or as dots and choose the correct answer among two possible outcomes. The proposed outcomes could also be in symbolic or non-symbolic format, although the symbolic to symbolic combination was excluded.

Problems used for this task were constructed using the same reference numbers as those for the numerosity comparison task guaranteeing that the magnitude of the sets were comparable throughout the tasks. Furthermore, answers for the two operations and the two references were controlled to allow a comparison between the different conditions. Five possible answers were selected making a total of 10 different problems per operation (Table IV.1). Distracters were generated by dividing and multiplying the correct answer either by 1.25 or by 1.5 ratios. Correct answers were therefore presented with the four possible distracters and counterbalanced for side of presentation. Overall, 240 trials per operation were administered (2 references, 5 problems, 4 distracters, 3 format combinations, 2 presentation sides).

Some precautions were taken for both controlling luminosity of the stimuli and ensuring that participants did not confuse the problem with the proposed outcomes. Problems were presented in black and possible outcomes in white inside a grey disk on a black background (Figure IV.3.). For this task, size of dots were kept constant throughout problems and outcomes. Total size of non-symbolic stimuli covered 5.5 degrees of visual angle whereas symbolic stimuli covered 4 degrees of visual angle.

Table IV.1.: Stimuli used for the approximate addition and subtraction task.

Problems		Correct Answer	Distracters			
Subtractions	Additions		1/1.25	1/1.5	1*1.25	1*1.5
58-16	16+26	42	28	34	53	63
60-16	16+28	44	29	35	55	66
61-16	16+29	45	30	36	56	68
62-16	16+30	46	31	37	58	69
63-16	16+31	47	31	38	59	71
74-32	32+10	42	28	34	53	63
76-32	32+12	44	29	35	55	66
77-32	32+13	45	30	36	56	68
78-32	32+14	46	31	37	58	69
79-32	32+15	47	31	38	59	71

Each trial started with the two empty grey disks for 1500 ms on each side of a fixation point. The problem appeared for 2000 ms with the red operation sing between the two disks. Again the initial empty disks appeared for 200 ms before the two possible choices filled the disks. The selection of the answer erased the screen (Figure IV.3.).

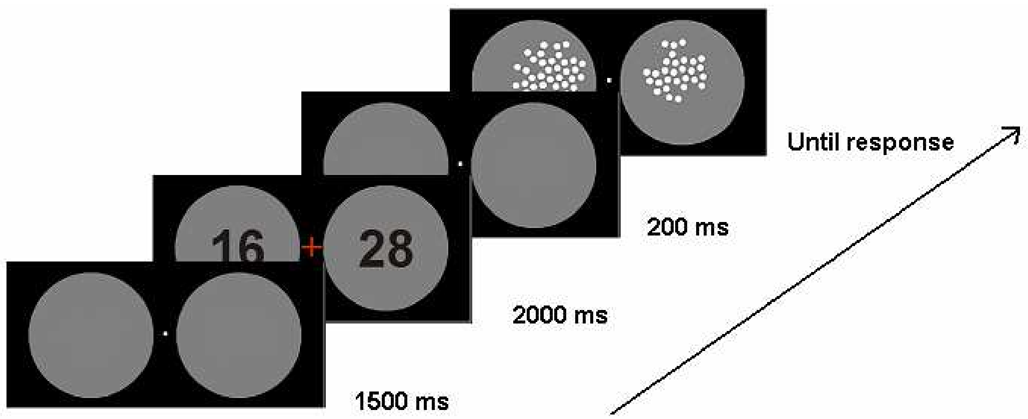

Figure III.3.: Example of an addition trial with the problem presented in the symbolic format and the outcomes in non symbolic format.

I.1.2. Results.

Numerosity comparison

The percentage of times a given set was chosen as numerically larger by each participant were averaged at group level. Mean percentages were then plotted on a linear and a logarithmic scale for numerosity (Figure IV.4.). When means are plotted on a linear scale, the slope of the sigmoid curve is steeper for reference 32 than for reference16. Conversely, the two curves become identical with a fixed amplitude on a logarithmic scale. The log-Gaussian model explains 99% of the variance and the Weber value (*w*) is of 0.16. This data conforms to Piazza and collaborator's (2004) results, who found a *w* of 0.17.

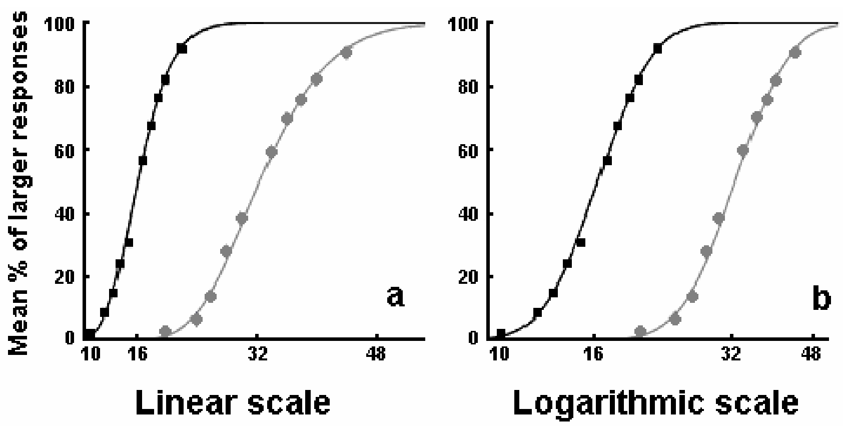

Figure IV.4.: Mean percentages of larger responses for each numerosity are plotted on a linear (a) and logarithmic (b) scale. The sigmoid curves become identical for the two references once plotted on a logarithmic scale. The y-axis represents percentage of times a given numerosity was considered larger than the reference.

For all participants a Weber value was also calculated for each version of the task. A correlation on participants' *w* did not reach significance (r = .392, p = .12). However when the values were introduced in a repeated measures Anova, group means did not statistically differ for the two sessions ($F_{(1,16)} = 1.2$, $p > .1$) indicating that overall the values fell inside the same range (Figure IV.5.). Mean and standard deviation (*SD*) of *w* for the first session was of 0.166 (0.05) and for the second session of 0.152 (0.04).

To analyse the between participants heterogeneity in achieving the task, a one-way Anova was run with subject as independent variable and *w* as dependent variable. A statistically marginal significance was obtained ($F_{(16,17)} = 2.22$, $p = .055$) suggesting different numerical abilities between participants. These differences could also reflect in more complex tasks.

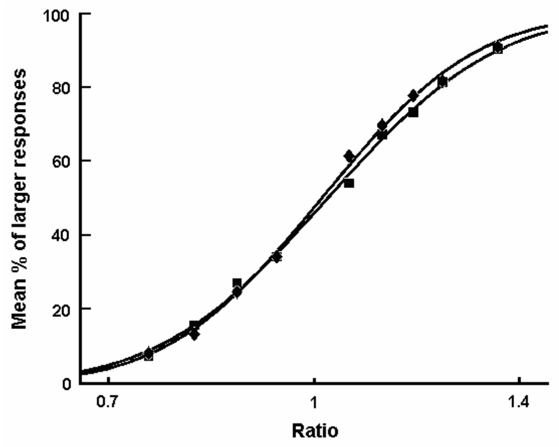

Figure IV.5.: Mean percentages of larger responses on the two sessions as a function of ratio. The y-axis represents percentage of times a given numerosity was considered larger than the reference.

In summary, the data observed with the numerosity comparison task replicates previous data and shows to be stable within a range of values. The correlation could have failed to reach significance because the pool of participants was relatively small as well as the number of trials. Increasing both parameters would increase statistical power and the correlation for the two sessions would presumably reach significance.

Number comparison

For the number comparison task, mean accuracy was of 97.8% (*SD* = 1.9) and therefore only reaction times (RT) were considered for the analysis. RTs were averaged across distances and then averaged across participants. Since the numbers were composed of two digits, it was not possible to apply a commonly used function that reorders any couple of one-digit numbers according to both size and distance of the pair (Welford function, Butterworth, Zorzi, Girelli, & Janckheere, 2001). Therefore, a modified version of this function was used: for each pair the ratio between the logarithms of the smaller and the larger value (logarithm MIN/logarithm MAX) was calculated. For example, the ratio of the logarithms is smaller for the pair 10-16 compared to the pair 16-22 since the second pair has an overall larger numerical size although both distances are of 6. Moreover, number pairs were classified according to the compatibility of the digits. If both digits of one number are larger/smaller than both digits of the other number in the pair, the pair was compatible. In the opposite case, the pair was incompatible (e.g., 16 vs. 22 are not compatible since for the decades 1 is smaller than 2 and for the units 6 is larger than 2). Compatibility has shown to influence performance in two-digit number comparison (Verguts & De Moor, 2005; Reynvoet & Brysbaert, 1999; Nuerk, Weger & Willmes, 2001). Finally, the two indices were introduced in a regression and the model explained a significant part of the variability ($F_{(2,39)} = 20$, $p < .001$, $R^2 = .528$) with the following equation:

Y = (88.195 compatibility) + (512.104*ratio of the logarithms) + 141.489*

Reaction times were therefore faster for pairs with compatible numbers, numerically small and with a large distance.

Approximate addition and comparison

Only accuracies were considered for the approximation task and percentages were normalized with the 2arcsin√(p) formula. All main effects and interactions violating the sphericity assumption were adjusted using the Huynh-Feldt correction method and multiple comparisons were corrected with the Bonferroni method. One participant was excluded from the pool for being more than 2 *SD* below group average for accuracy.

The data were submitted to a repeated measures ANOVA Operation (addition, subtraction) x Reference (16, 32) x Format combination (dots-dots, dots-digits, digits-dots[1]) x Ratio (1/1.5, 1/1.25, 1*1.25, 1*1.5). The order of task was introduced as a between subjects variable. Given that it did not influence performance it was not considered further.

The type of operation participants had to solve did not influence overall accuracy. Participants were 70,1% (*SD* = 4.5) accurate for the addition task and 67.8% (*SD* = 7.9) for the subtraction task. Format combination was significant ($F_{(2,30)} = 16$, $p < .001$) and post-hoc comparisons showed significant differences for the three format combinations (all *ps* < .05). The combination with the highest accuracy was dots-digits with 76.5%, followed by dots-dots with 70.4%, and the worst were the digits-dots trials with 63.5% of accuracy.

A main effect of ratio also highlighted the classic distance effect ($F_{(1.3,19.5)} = 5.3$, $p < .05$). Accuracy dropped when the ratio between the correct answer and the distracter was closer to one. For the distant ratios, 1/1.5 and 1*1.5, accuracies were respectively of 77% and 73% whereas they were of only 67% and 62% for the 1/1.25 and 1*1.25 ratios.

The two-way interaction Ratio x Operation was significant ($F_{(1.2,18.5)} = 5.2$, $p < .05$) revealing a greater accuracy for smaller ratios when adding and a greater accuracy for larger

[1] For simplicity, format combinations are abbreviated by indicating the format of the problem and outcomes as follows: both parts presented as dots will be labelled dots-dots combination; problems presented as dots and outcomes as digits will be labelled dots-digits; and the reverse combination will be labelled digits-dots.

ratios when subtracting. When participants performed additions, they always tended to choose the largest set. That is, the correct answer was preferentially chosen when it was presented with a smaller distracter (smaller ratio) compared to when it was presented with a larger distracter. For the subtractions the reverse was true: participants tended to select the smaller set as the correct answer. They chose the correct answer when the distracter was larger whereas they selected the distracter when it was the smallest set (Figure IV.6.a). This result is in agreement with the *Operational Momentum* described by McCrink and collaborators (2007). In their experiment they observed that when performing approximate additions participants tended to overestimate whereas when the operation was a subtraction, answers were systematically underestimated.

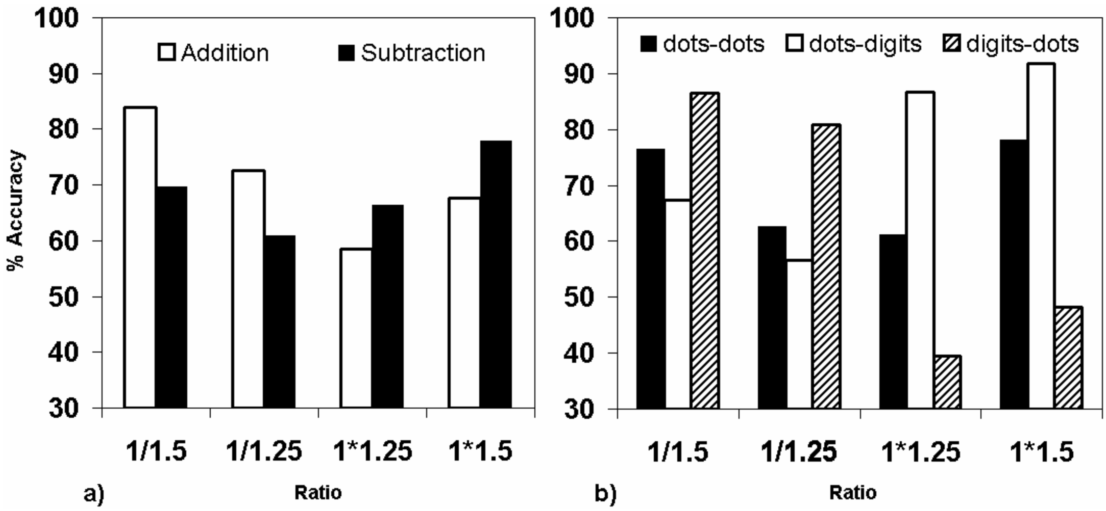

Figure IV.6.: The graph on the left side of the figure (a) represents percentages of correct responses as a function of operation and ratio. The graph on the right side (b) represents accuracies as a function of format combination and ratio.

The second significant double interaction was Format x Ratio ($F_{(2.5,37.5)} = 34$, $p < .001$; Figure IV.6.b). This interaction highlights that the tendency to underestimate or overestimate is influenced by the format of presentation. When both problem and outcomes are presented in dots, accuracies show a classical distance effect, whereas an asymmetry appears when formats are crossed. For the dots-digits combination, the choice is preferentially set on the smallest number whereas it is preferential for the largest set in the digits-dots condition. In other words, when the combination is dots-digits, accuracy is high when the correct answer is presented with the larger distracter. The reverse was true for the digits-dots format combination.

The interaction Operation x Reference was also significant ($F_{(1,15)}$ = 6.5, p < .05). For the addition task, accuracy was higher for problems with reference 32 whereas the opposite was true for subtractions. Accuracies for addition were 69% for reference 16 and 72% for reference 32. In the subtraction condition, accuracies were 69% and 67% for reference 16 and 32 respectively. This could be explained by the choice of operands for the experimental stimuli. Indeed, the second operand for the additions with reference 32 was smaller than the second operand for reference 16 (range 10 to 15 vs. 26 to 31). Indeed, underestimation is proportional to the numerosity, hence, for reference 32 the approximation was more accurate. Moreover, the same was true for the subtraction: problems where 16 had to be subtracted had a better accuracy since overall underestimation was smaller than for problems with 32.

Finally, the significant triple interaction ($F_{(3.9,59)}$ = 3.3, p < .05) qualifies the *operational momentum* effect according to format combination. As shown by the graphs in Figure IV.7. only the dots-dots combination clearly replicates the pattern observed by McCrink (2007). The other two combinations are less clear showing only a strong format effect. Therefore further ANOVAs were run separately for each combination. For both dots-dots and digits-dots format combinations, the ratio and the ratio by operation interaction were significant (dots-dots ratio: $F_{(3,264)}$ = 7.9, p < .001; dots-dots interaction: $F_{(3,264)}$ = 11, p < .001; digits-dots ratio: $F_{(3,264)}$ = 51, p < .001; digits-dots interaction: $F_{(3,264)}$ = 5.9, p < .001) confirming the tendency to overestimate when performing addition and underestimate when performing subtractions. For the third combination, dots-digits, the interaction failed to reach significance (ratio: $F_{(3,264)}$ = 54, p < .001).

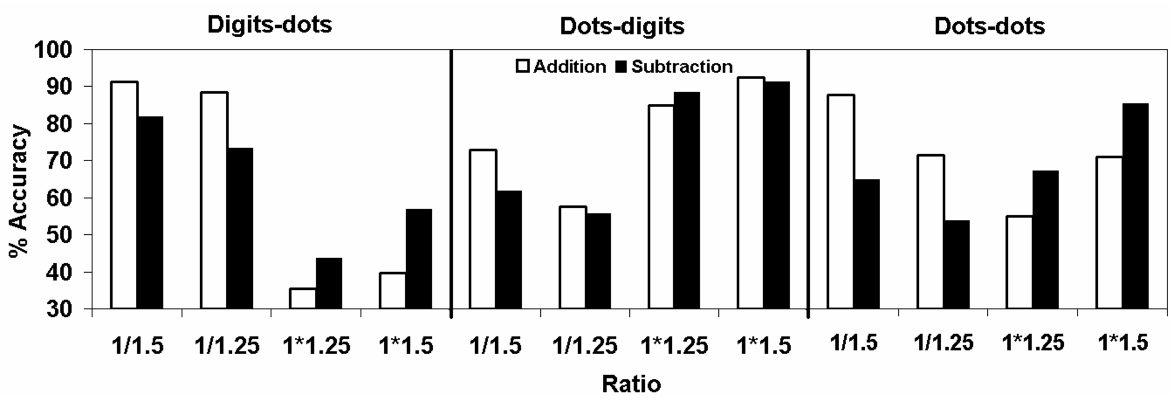

Figure IV.7.: Percentages of accuracies are represented as a function of format, operation and ratio. Only the dots-dots condition clearly replicates the operational momentum introduced by McCrink and collaborators (2007).

In summary, the analyses of approximate addition and subtraction task showed that when both operation and outcomes are presented as dots, performance is influenced by the type of operation that had to be computed. When subtracting, the tendency is to prefer a smaller outcome than the correct answer and when adding, the reverse is true. This corresponds to the definition of the *operational momentum* effect.

Performance was also influenced by the format since for mixed format combinations (digits and dots) the patterns of overestimation and underestimation were modified. Indeed, for the dots-digits combination, participants were more prone to of prefer a smaller outcome and the type of operation did not influence accuracy. Conversely, when participants were solving problems in the digits-dots combination, they preferentially chose the larger set of dots. Moreover, the type of operation influenced accuracy since participants overestimated less when subtracting digits compared to adding them.

These results will be discussed in more details in the conclusions section since they allow to offer an alternative explanation based on the general underestimation finding but without being mutually exclusive with the *operational momentum* hypothesis.

Relations between the Weber value and the other tasks.

The Weber value failed to correlate with the number comparison task. For the latter, regressions for each participant were computed introducing the same parameters as those used at the group level (compatibility and ratio of the logarithms). The slope value (β) associated to the ratio of the logarithms for each participant was then correlated to the Weber value (r = -.066, p > .05).

To evaluate a relation between the Weber value and the approximate addition and subtraction task, for each participant an index of the tendency to overestimate was calculated separately for each format combination and for both operations. According to McCrink and collaborators, when performing an addition there is a tendency to prefer the largest answer and the converse for the subtraction. This difference was calculated by first averaging the accuracies of larger (1*1.25 and 1*1.5) and smaller ratios (1/1.25 and 1/1.5), then calculating the difference between the latter and the former. A positive value indicated a higher accuracy for smaller ratios. In other words, participants overestimated since they preferred the larger ratio when presented with the correct answer. A negative value indicated higher accuracy for larger ratios: participants preferentially selected the smaller set.

Of all the possible correlations between the over-estimation index and the Weber value, only two were significant and both for the digits-dots operations (addition: R^2 = .25, p < .05;

subtraction: $R^2 = .43$, $p < .005$). Although surprising that only two correlations reached significance, it is worth highlighting that these two were the most likely to be related to the Weber value. Indeed, in the digits-dots condition, participants could exactly calculate the result and would only have to discriminate between the two numerosities presented as outcomes. For the other tasks, other factors could have hidden the relation. Indeed, for the dots-dots combination estimation biases could emerge at both levels of the task, whereas in the dots-digits combination the discrimination ability for the answer has probably been hidden by the exact representation of the outcomes. As will be discussed in the conclusion, the dots-digits condition was probably biased and is not well representative of pure approximate calculation. Because the range of answers was always the same for all combinations, one participant overtly declared at the end of the second session that he had understood the range of correct answers. Thus, participants could indeed infer the probable answers to the dots-digits trials by exactly calculating the result in the digit-dots trials. In support to this to this speculation it can be noted that this format combination yielded the highest percentages of accuracy.

I.2. Conclusions

The aims of this study were to estimate the individual value of internal Weber fraction through a numerosity discrimination task and to relate this value to other tasks that putatively all rely on a common (log-Gaussian) representation as shown by Weber's law signature.

As for the reliability of the individual Weber value, two versions of the discrimination tasks were administered to the same participants one week apart. Unfortunately, participants' values for the test and retest sessions failed to correlate but group analysis also failed to show a statistical difference between the two sessions. Although it is speculative to draw conclusions from null effects, we may consider that the absence of correlation was due to the small number of participants and trials. This shortcoming should easily be resolved by increasing the pool of participants and therefore increasing statistical power.

As for the primary aim of this study, we have tested the same participants on two different numerical tasks, number comparison and approximate addition and subtraction, that should both rely on a common internal representation. The former showed classical effects observed by other authors for double digit comparison. That is, participants' performance was influenced by both distance and size of the pair as well as by the compatibility of the digits (Reynvoet & Brysbaert, 1999; Butterworth et al., 2001; Nuerk et al., 2001; Verguts & De Moor, 2005). Surprisingly this task also failed to correlate with the ability of discriminating between sets (i.e., the Weber value), which has been considered as a basic competence upon which further symbolic knowledge would be constructed (e.g., Feigenson, Dehaene & Spelke, 2004). Indeed, as observed by the distance and size effect, comparing digits requires the activation of a numerical representation that obeys to Weber's law. The failure to correlate could be explained by the fact that, the indices used might not be directly comparable because accuracy the Weber value reflects discrimination accuracy while performance in the number comparison task was measured by RTs. Alternatively, it could be argued that the symbolic comparison task relies only partially on the internal log-Gaussian representation. In support of the latter hypothesis, different models of numerical processing are now leading to the conclusion that adult numerical abilities are underpinned by two separate systems: one would be language-based and activated for all exact numerical tasks and the second would be language independent, necessary for numerical approximate processing and phylogenetically determined (Verguts, Fias, & Stevens, 2005; Verguts & Fias, 2004; Spelke, 2000; Zorzi, Stoianov, Becker, Umiltà & Butterworth, submitted). Further support to this hypothesis comes from a neuropsychological study (Lemer, Dehaene, Spelke, & Cohen, 2003) on two patients with numerical deficits. The first patient had preserved abilities in approximation tasks but failed in symbolic tasks. The second patient showed the reverse pattern of deficit, with spared exact calculation and deficits in approximate tasks. Therefore, in our data, when comparing Arabic numbers instead of numerosities, an exact numerical system would be active and therefore hide the numerical representation described by the Weber value.

The approximate addition and subtraction task was constructed in order to observe a format influence on performance. Symbolic and non-symbolic numerosities were presented as either the problem or as the outcome. To avoid exact calculation, the symbolic to symbolic format combination was excluded. The interesting observation was that the *operational momentum* defined by McCrink and collaborators (2007) and replicated by Knops and collaborators (submitted) was found in the non-symbolic to non-symbolic format combination. When the task was to add two sets of dots and choose the correct answer among two proposed sets, participants tended to select the larger of the two outcomes (overestimation), whereas if the task was to subtract, the preferred answer was the smaller set (underestimation). The patterns observed for the other two format combinations were different. When the problem format was symbolic and the proposed outcomes were non-symbolic, participants tended to overestimate the correct answer, regardless of the operation, selecting the larger set. For the reverse combination of formats the opposite was true: when participants had to choose among two symbolic outcomes after viewing a non-symbolic problem, they systematically chose the smaller value. Further analyses revealed that in the symbolic to non-symbolic combination the overestimation was nevertheless influenced by the operation. The tendency to overestimate was statistically smaller for subtractions than for additions, highlighting a subtle *operational momentum* effect. For the non-symbolic to symbolic combination this effect did not appear. It is arguable that this absence was due to the stimuli employed. Indeed, several participants reported that they had guessed the range of possible correct answers by exactly calculating the solutions in the symbolic to non symbolic condition and inferring that the same answers held true for the non symbolic to symbolic condition. This strategy used by participants has probably influenced overall performance, diminishing a possible *operational momentum* for this format combination.

Even if the *operational momentum* seems to be confirmed in this experiment, there is a second explanation for the observed patterns of results, which in our view is more economical although not mutually exclusive with the former. Our alternative explanation is based on numerous observations that dot patterns are underestimated and the degree of underestimation is proportional to set size (Ginsburg, 1978; Izard & Dehaene, 2007; Allik & Tuulmets, 1991; Durgin, 1995). Larger sets are more underestimated than smaller ones. In the present experiment, when the combination was non-symbolic to symbolic, participants mentally approximated the result and since it was in a non-symbolic format they underestimated the outcome. Therefore, when they had to select the answer in the symbolic format they preferred the smallest value. For the opposite combination, symbolic to non-symbolic formats, participants could obviously calculate the exact result but when they had to chose among sets they tended to underestimate the proposed outcomes and therefore preferred the largest set (perceived as less numerous). Finally, also for the purely non-symbolic combination the data can be accounted by an underestimation bias. When solving additions, the operands as well as the outcomes are underestimated and since the proposed outcomes are numerically larger than the operands (always the case in additions), the latter were more underestimated, leading participants' choice towards the larger sets (overestimation). In subtraction, the numerically larger set, the more underestimated, was in the operation and led participants to a much smaller approximated answer than the correct one. Therefore, when choosing among the possible outcomes, participants' choice preferentially fell on the numerically smaller set (Figure IV.8.).

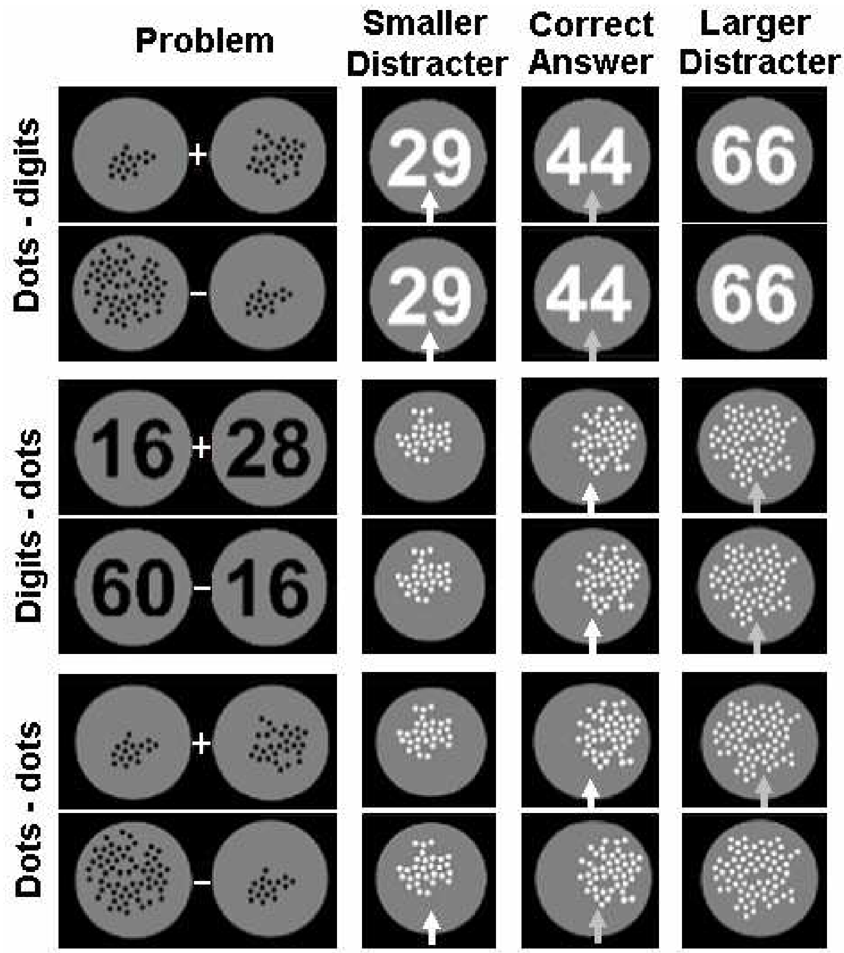

Figure III.8.: Example of problem and outcomes for each format combination and operation. The problem is on the right hand of the picture and the correct answer with the two types of possible distracters (smaller or larger) are on the left hand. Arrows indicate what outcomes were preferentially chosen. For example, in the dots-digits combination, participants chose the correct answer when it was presented with the larger distracter (dark arrow), whereas they chose the distracter when it was smaller than the answer (light arrow).

Finally, the Weber value obtained for each participant correlated with the performance in the approximate addition and subtraction when the format combination was symbolic to non-symbolic. This result is encouraging since this combination of format was probably the closest to the numerosity discrimination task. Indeed, after calculating the answer to the symbolic problem, participants had a simple discrimination task between the two proposed outcomes. The correlations with the other tasks could be explained by the double source of underestimation in the purely non-symbolic task and by the bias observed for the condition in which the outcomes where symbolic. To overcome this bias, a larger number of trials and stimuli should be created where the correct answer could also be the wrong answer to some problems therefore disrupting the use of non-numerical strategies. This solution would nevertheless leave open a question regarding the performance on symbolic to symbolic approximate calculation. Therefore, a fully crossed experiment could be designed by presenting several distracters differing by a larger range of ratios and by avoiding to present the correct answer as in Knops and colleagues (submitted).

Chapter 5: Numerical Representation in Synaesthesia.

> I must, however, beg them not to consider their own minds as identical with those of every other sane and healthy person. Psychologists ought to inquire into the mental habits of other men with as little prejudice as if they were inquiring into those of animals of different species to their own, and should be prepared to find much in many cases that is quite unlike their own personal experience.

(Galton (1880), Nature, 22, p. 494)

In 1880, Galton was amongst the first to discover that some people may experience sensory stimulations in various different ways. For instance, he pointed out that “the various ways in which numerals are visualised is but a small subject, nevertheless it is one that is curious and complete in itself (Galton, 1880b, p. 252)”. In his seminal work, Galton describes personal reports of people who vividly experience numbers in their mind’s eye. They visualize them in space, colour or both, and sometimes are able to manipulate these images to perform arithmetical operations. This particular type of associations have been termed number-colour and number-form synaesthesia. In Ancient Greek, *σύν* (syn) means “with” and αἴσθησις (aisthēsis) means “sensation” thus indicating the union of sensations. Synaesthesia is defined today as a neurologically-based phenomenon where the stimulation of one sensory modality leads to an automatic and involuntary secondary experience in the same or in a different modality. Various types of associations exist (e.g., grapheme-colour, sound-colour, word-taste, ect.) and probably not all have yet been described. However, here the focus will only be on the association where the perception of numbers (heard or read) induces a sensation or the visualisation of a colour.

Because the earliest (1880-1930) investigation of these phenomena was difficult and relied primarily on introspection, the rise of behaviourism hampered research in this field until almost completely disappearing from scientific journals. It is only around the 80’s that adequate experimental techniques allowed to reconsider synaesthesia as a phenomenon to investigate.

Several experiments have tested synaesthetes’ claims and given proof of their introspective reports (Nikolic, Lichti, & Singer, 2007; Wollen & Ruggiero, 1983; Paulsen & Laeng, 2006; Mills, Boteler, & Oliver, 1999; for a review see Hubbard & Ramachandran, 2005). Moreover, with the increasing accessibility to neuroimaging techniques and studies in genetics the neurological bases of the phenomena are starting to be drawn (Ramachandran &

Hubbard, 2001a; Nunn et al., 2002; Hubbard & Arman, 2005; Barnett et al., 2007; Smilek et al., 2001; Hancock, 2006).

In addition of being "curious", as described by Galton, synaesthesia is another path to understand human cognition. Just as it is informative to study neurologically impaired patients, synaesthesia through its "abnormality" sheds light on how a "normally functioning" system works. In the field of numerical cognition, synaesthetic experiences have given support to some models of numerical representation by overtly describing their perceived associations.

As seen in Chapter 1, several models of numerical representations postulate that numbers are represented on a mental number line (Dehaene, 1992; Verguts, Fias, & Stevens, 2005; Zorzi, Stoianov, Becker, Umilta, & Butterworth, 2008). Moreover, the distance, size and SNARC[1] effects suggest that the line might be compressed towards larger numbers and be oriented from left to right (Dehaene, Bossini, & Giraux, 1993; Ashcraft, 1992; Moyer & Landauer, 1967). Synaesthetic reports of number-form associations have shown that in about 63 to 66 % of the cases, (Seron, Pesenti, & Noël, 1992; Sagiv, Simner, Collins, Butterworth, & Ward, 2005), the orientation of the line where numbers are visualized follow a left to right orientation. Curiously, in a study of individual performance, only 65% of the participants showed a SNARC effect (Wood, Nuerk, & Willmes, 2006). In addition to reporting numbers on a line, several synaesthetes describe their numbers as becoming more fuzzy the larger the value (Figure V.1., Seron et al., 1992; Galton, 1880b). The compressed numerical representation was also demonstrated by interference effects in a Stroop type paradigm (Cohen Kadosh, Tzelgov, & Henik, 2008). Number-colour synaesthetes were asked to name digits that could be congruently or incongruently coloured with their personal associations (if "2" was associated to "red" for one participant, it would have been incongruent if presented with the colour associated to "3" which was perceived as "blue") or to name the colour ignoring the digit. Reaction times showed that the congruency effect was modulated by the numerical magnitude: the larger the number the stronger the interference. According to the authors, this would challenge the idea of a linear number line.

Potentially, besides these examples in numerical cognition, synaesthesia may inform on many other domains as automaticity, crossmodal interactions, modularity, brain development, attention, etc. This is why an apparently unrelated study is included in this thesis. The

---

[1] Spatial Association of Response Codes effect: responding is faster for small numbers with the left key and to large numbers with the right key, in western societies.

heterogeneity of adopted perspectives should bring complementary knowledge to the understanding of numerical representation.

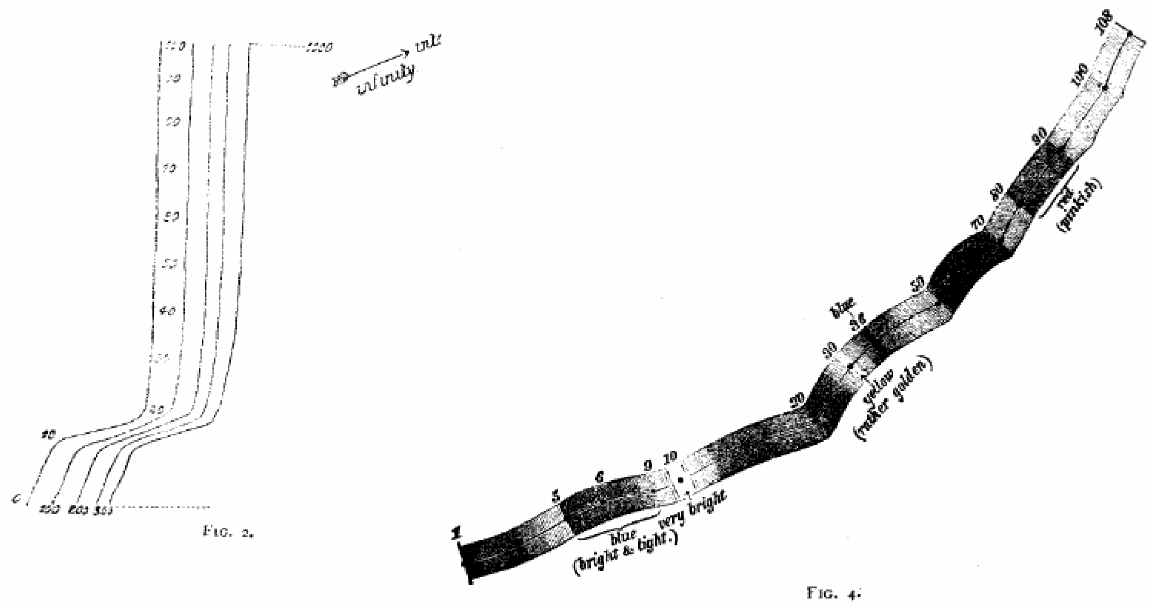

Figure V.1.: Number-form associations form two synaesthetes interviewed by Galton (Galton, 1880b). The representation on the left was also followed by remarks highlighting that larger numbers were closer to each other both for decades and hundreds. On the right, the spatial representation is accompanied by different colours for each decade.

Ⅰ: Experiment 4: Implicit versus Explicit Interference Effects in a Number-Colour Synaesthete.[2]

Abstract

One of the fundamental questions in the study of consciousness is the connection between subjective report and objective measures. Here, we explore this question by testing an individual with grapheme-color synesthesia, NM, who reports an associated conscious experience of seeing colors when viewing digits, but not when viewing dot patterns. Synesthesia research has traditionally used variants of the Stroop paradigm as an objective correlate of these subjective synesthetic reports. We used two paradigms: a standard synesthetic Stroop paradigm, in which digits were colored congruently or incongruently with the colors NM reports, and a numerosity Stroop paradigm, in which random dot patterns were colored congruently or incongruently with the colors that NM reports for digits. We observed longer response times in the incongruent condition for both the standard and the numerosity Stroop paradigms, despite the fact that NM denies experiencing colors for dot patterns. This constitutes a clear dissociation between subjective and objective measures of synesthetic experience. Based on these results, we argue that establishing the presence of synesthesia in an individual should depend primarily on the presence of subjective reports, validated by objective measures, and more generally, that consciously and unconsciously mediated interference may arise from qualitatively different mechanisms.

[2]This work is currently submitted as Berteletti, Hubbard and Zorzi.

I.1. Introduction.

Over a century ago, Galton (Galton, 1880b; Galton, 1880a) described a particular neurological condition in which a sensory experience would automatically evoke an additional percept (concurrent) in the same or a different sensory modality (i.e. colored graphemes, colored music, etc.). Different etiologies have been advanced to account for this phenomenon, ranging from a deficit of neuronal pruning during development to overactive binding between colors and letter forms (Baron-Cohen, Burt, Smith-Laittan, Harrison, & Bolton, 1996; Hubbard & Ramachandran, 2005). However, regardless of which explanation is ultimately supported, all models of synesthesia agree that cerebral connections are the physiological basis for these unusual perceptual experiences.

To demonstrate the genuineness of synesthesia, numerous behavioral paradigms have been developed; among these the most widely used is a modified version of the Stroop paradigm (Wollen & Ruggiero, 1983; Mills et al., 1999; Odgaard, Flowers, & Bradman, 1999; Dixon, Smilek, Cudahy, & Merikle, 2000; Mattingley, Rich, Yelland, & Bradshaw, 2001). The classical version of the Stroop paradigm is a conflict task in which color names are written in various ink colors (i.e. the word RED in either red or green ink; Stroop, 1935; for a review see MacLeod, 1991). The participant is required to name the ink color while ignoring the color name. Reaction times (RTs) to name the ink color are longer when the ink color and word name are incongruent, compared to when they are congruent, which suggests that automatic reading process interfere with the color naming process and thus increase RTs. Using the same logic, the paradigm has been adapted to test grapheme-color synesthetes who report perceiving colors (photisms) when viewing or thinking about letters and numbers (Wollen & Ruggiero, 1983). When the graphemes are incongruently colored with respect to the photisms reported by an individual synesthete, RTs are higher for naming the ink color compared to the condition where the same stimuli are congruently colored (e.g. Mills et al., 1999; Dixon et al., 2000; Mattingley, Payne, & Rich, 2006; Nikolic et al., 2007). This result is one of the strongest pieces of evidence for the authenticity of synesthesia, with the interference being interpreted as the result of an automatic and involuntary process triggering the perception of color.

However, it is clear that synesthesia is not a unitary phenomenon, and it is possible that synesthetic color experiences are elicited at different stages of processing. Ramachandran and

Hubbard (Ramachandran & Hubbard, 2001b; Hubbard & Ramachandran, 2005) have introduced a distinction based on the representational level at which the concurrent synesthetic experience is evoked (perceptual vs. conceptual). The authors suggest that if different stimuli sharing the same meaning (e.g. digits, number words, Roman numerals) also induce the same synesthetic experiences, then the association probably occurs at a semantic level (“higher” synesthesia), whereas if lower-level properties of the stimulus elicit concurrents, (e.g., a digit and number word elicit different colors) then the association probably arises at a perceptual level (“lower” synesthesia).

In order to explore this question, some authors have used versions of the Stroop paradigm to test the representational level at which the concurrent is induced. In the first study of this kind, Dixon *et al*. (Dixon et al., 2000) presented a synesthetic participant, C, with a “mathematical Stroop” paradigm in which arithmetic problems, such as 2 + 5 were presented, followed by a color patch. C’s reaction times were slower when the color patch was incongruent with the arithmetic result, demonstrating that the digit does not have to be physically presented to elicit interference. However, these experiments do not allow us to identify the representational level at which the concurrent was elicited, as the participant could have depended on either a semantic representation or on an image of the Arabic digit. In a follow-up study, Jansari *et al*. (Jansari, Spiller, & Redfern, 2006) tested whether interference was present both in the visual and auditory modality in three synesthetic participants. One of their participants showed interference only in the visual modality, while the other two showed interference only in the auditory modality, contrary to the hypothesis that the synesthetic experiences for these individuals are elicited purely from higher-level conceptual representations, which should be triggered by both auditory and visual stimuli, as has previously been shown for numerical representations (Eger, Sterzer, Russ, Giraud, & Kleinschmidt, 2003).

More recently, Ward and Sagiv (2007) presented a single-case study of a synesthete, TD, who reported conscious experiences of colors when viewing Arabic digits, when counting fingers, and when viewing dice patterns, which according to the current definition constitutes a case of higher synesthesia. Using a Stroop-like paradigm, the authors showed that TD's synesthetic colors interfered with incongruently colored stimuli for all types of numerical stimuli. Moreover, incongruent colors also interfered when TD was asked to estimate the number of dots in a dice pattern (the reverse of the standard task). Interestingly, colors not only impaired color naming in the classical Stroop but also in the numerosity judgment suggesting that colors convey numerical information for TD. Based on the results, the authors conclude not only that TD's synesthesia is elicited at a semantic level ("higher" synesthesia) but also that it is a result of bi-directional links between colors and numbers. However, as Ward and Sagiv note, random dot patterns and dice patterns may not be strictly comparable given that dice patterns may be overlearned, and may be treated as familiar visual patterns rather than pure numerosity stimuli.

In this study, we present a synesthete (NM) who reports colors for digits 1 to 9. Unlike TD, NM reports that he does not experience synesthetic colors when shown dice patterns or random dot configurations. Despite the fact that he reports no subjective colors, we tested NM on two Stroop-like tasks. The first one, a digit Stroop task, aimed at replicating previous interference findings. The second one was a numerosity Stroop task where a canonical (dice patterns) and two non-canonical configurations of dots (NC1 and NC2) were presented colored congruently or incongruently with the color of the corresponding digit. Since the digits elicit conscious experiences of color for NM, we predicted the same interference effects as already reported in previous studies. For the dice and non canonical patterns, however, whether interference is observed will depend on the locus at which the synesthesia is elicited. If NM's synesthesia is elicited at lower level, then we would not predict any interference with dot patterns. However, if NM's synesthesia is elicited at a higher level, then we would expect to observe interference in at least the dice pattern condition which is known to trigger the numerical information faster than non-canonical configurations (Wolters, van Kempen, & Wijlhuizen, 1987). The presence of Stroop interference from random dot patterns would constitute the first demonstration that synesthetic triggers in number-color synesthesia can occur at a semantic level, since they are not subject to the above concern about overlearned pattern recognition.

Moreover, since the dot patterns are not overtly associated to color, we would predict that if interference is found, it will be smaller than that shown in the digit Stroop task. The presence of such interference despite the absence of conscious synesthetic reports would also suggest that the connection between semantic stimuli and synesthetic experiences is triggered automatically. Finally, the possibility of finding interference even in the absence of consciously experienced colors raises the question first posed by Merikle and colleagues (Merikle, Smilek, & Eastwood, 2001; Cheesman & Merikle, 1986), as to whether explicit and implicit Stroop-interference arises from qualitatively different mechanisms. In the discussion we thus explore an alternative account for the implicit effects of synesthesia.

I.1.1. Method.

Participant

NM answered an announcement placed in the lecture halls on the University of Padova campus. When we first interviewed him, NM was 29 years old and was finishing a PhD program. He reported being a grapheme-color and a number-color synesthete with number-forms for digits and week of the days. He visualizes the numbers 1 to 9 from left to right and the smaller the digit the glossier the color. Moreover, the smaller digits are bold and become progressively more normal font with numerical size.

NM's synesthesia was first assessed with a broad self-report questionnaire and the genuineness of his number-color associations for digits 0 to 9 were assessed with a test-retest procedure. Color selections for each number were recorded in the form of RGB triplets, values ranging from 0 to 255 on each dimension. For the color test-retest, a group of five controls with no synesthesia also performed the task and were instructed to remember their associations since they would be retested one week later. NM, on the other hand, was retested two months apart without notice. The analysis of the RGB values, using the city block distance procedure[3], showed that NM was less variable than the controls and scored close to 2 *SD* deviations below the mean of the non-synesthetes (NM's mean distance for the 10 digits was .095 in RGB space, the controls' mean distance was .533 in RGB space, with a standard deviation (*SD*) of .247; $t_{(4)} = -1.62$, $p = .09$ one-tailed[4]; (Crawford & Garthwaite, 2002)).

We then performed a second test to objectively verify NM's superior consistency by randomly presenting colors pairs to external judges and asking them to rate the similarity (1 = completely different to 5 = completely similar) of the colors chosen at Time 1 and Time 2. The judges rated the colors NM selected across the two sessions as significantly more similar than those of control participants (NM mean = 4.57, controls mean = 3.05 and $SD = .59$; $t_{(4)} = -2.343$, $p < .05$ one-tailed; (Crawford & Garthwaite, 2002)).

---

[3] The city block distance analysis permits calculation of distances between coordinates in a 3-dimensional space.

[4] Because the t-test tends to be liberal when comparing a single sample against a population, we used Crawford and Garthwaithe's (2002) program to test whether NM's individual score is significantly different from the control sample: http://www.abdn.ac.uk/~psy086/dept/SingleCaseMethodsComputerPrograms.HTM

Procedure

NM participated in two different tasks, a Digit Stroop and a Numerosity Stroop task in three experimental sessions. Each session started by asking NM to pick his personal colors for the numbers 1 to 6. This allowed us to ensure that the stimulus colors would be as close as possible to NM's photisms in the specific experimental lighting conditions. In order to eliminate the possibility of carry-over effects from the Digit Stroop task, the Numerosity Stroop Task was run in the first two sessions and the Digit Stroop Task in the third, although we describe the tasks in the opposite order for clarity.

Digit Stroop Task

In order to replicate previous studies demonstrating synesthetic Stroop effects when Arabic digits are presented in colors inconsistent with those reported by the synesthetes (Wollen & Ruggiero, 1983; Mills et al., 1999; Odgaard et al., 1999; Dixon et al., 2000; Hancock, 2006; Paulsen & Laeng, 2006), we presented NM with Arabic digits colored either congruently or incongruently with his reported photisms. NM's task was to name the ink color as quickly and accurately as possible, while ignoring the identity of the digit. We also tested the two baseline conditions, which were to name the ink of large colored disks and name black digits. Given that the simple digit naming task can be performed through the non-semantic pathway, we would not expect RTs to be influenced by numerical size, and thus no numerical effects were expected in the digit naming task (Dehaene, 1992; Butterworth, Zorzi, Girelli, & Janckheere, 2001).

Numerosity Stroop Task

In order to determine whether the synesthetic Stroop effect is limited to Arabic numerals, or whether it generalizes to semantic representations of number, we presented NM with dot patterns colored either congruently or incongruently with the photisms that he reported for the corresponding Arabic digit. As mentioned above, NM did not report perceiving colors for any kind of dot pattern. These dot patterns were either canonical (dice) or non-canonical patterns. Numerosities ranged from one to six and in order to minimize any potential learning effects for the non-canonical configurations, we constructed two different sets, non-canonical 1 (NC1) and non-canonical 2 (NC2)[5].

[5] The non-canonical patterns covered the same area on the screen as the dice patterns.

To control for any potential color naming effects, a baseline condition was run in which NM had to name the ink color of large central disks. The colors were the same ones used for the Stroop task. A second baseline, an enumeration task, was run to control for any possible familiarity effects of the dice patterns, which are assumed to be over-learned compared to the non-canonical configurations, and for any potential learning effects for the non-canonical configurations due to repeated presentation of the stimuli during the experiment. In this control condition, the same dot configurations as those for the Stroop task were used, but were presented in black on a gray background. Previous studies of dot enumeration have shown that reaction time increases with the numerical size of the set (numerical size effect (Wolters et al., 1987). Moreover, enumeration of the largest set of dots is usually faster than for the next smaller set (end-effect) because the largest set has only one competitor whereas all others have two competitors (van Oeffelen & Vos, 1982; Wolters et al., 1987). For the over-learned dice configurations, on the other hand, the RT curve is flat since participants merely need to recognize the pattern, rather than enumerate the number of dots (Wolters et al., 1987). Therefore, for the enumeration base line, we expected to observe numerical size effects and the end-effect (i.e. for numerosity 6) for the non-canonical configurations whereas we expected to observe a flat RT curve for the dice configurations.

General procedure

In both Stroop tasks each stimulus was presented 20 times, resulting in 240 total trials (2 congruity conditions x 6 numerosities) for the Digit Stroop and 720 total trials (3 pattern types x 2 congruity conditions x 6 numerosities) for the Numerosity Stroop, divided into two sessions of 360 trials. For the baseline conditions, there were 120 trials for color naming (6 colors), 120 trials for digit naming (6 digits) and 360 for dot enumeration (3 pattern types x 6 numerosities). Crucially, for the incongruent condition, in both the Numerosity and the Digit Stroop, the colors used were the colors associated with each of the other numbers (i.e. the numerosity or digit 3 in the incongruent condition was colored with the colors associated with 1, 2, 4, 5, and 6).

The trial sequence was the same for both tasks: a fixation cross was presented for 1000 ms after which the stimulus appeared until the voice key device detected the response. After the response was detected, the image was reduced in size, and a code appeared in the left lower corner allowing the experimenter to code, by means of the keyboard, the accuracy of the answer and any possible voice key errors (e.g., coughs or hesitations). Regular breaks were scheduled during the task and the participant could also choose to rest between trials before the answer was coded. The experiment was programmed using E-Prime 1.1 experimental software (Schneider, Eschman, & Zuccolotto, 2002b; Schneider, Eschman, & Zuccolotto, 2002a) and run on a PC desktop computer (AMD Athlon 1800+) running Windows 2000. The experimental room was dimly lit and quiet. Stimuli were presented on a 17-inch CRT screen (1024 x 768 resolution at a 75 Hz refresh rate) centrally in the visual field in less than 5° of visual angle. NM was seated 1 meter from the screen.

I.1.2. Results.

Overall Numerosity and Digit Stroop Tasks

Overall accuracy, collapsed across tasks, was 98.3% for congruent trials and 95.4% for incongruent trials. After errors and voice key detection errors were excluded (3.1% of trials) from the data set, naming latencies exceeding two standard deviations from the mean per condition were eliminated (4.9% of the remaining trials). For the two Stroop Tasks, an overall ANOVA was run on the mean latencies for the cleaned data. The factors introduced were stimulus type (digits, dice pattern, NC1 and NC2), congruency (congruent or incongruent ink color) and numerical value (1 to 6). All post-hoc comparisons were Bonferroni corrected.

Figure V.2. shows the mean reaction times for each of the four stimulus types as a function of whether the ink color was congruent or incongruent with NM’s reported photism. The main effect of congruency across digits and dots was significant ($F_{(1, 834)} = 190.12$, $p < .001$). This result replicates and extends previous research on digit and dot patterns demonstrating a synesthetic Stroop effect when ink colors are incongruent with reported photism colors. Mean RTs were 441 ms (*SD* = 121 ms) in the congruent condition and 549 ms (*SD* = 228 ms) in the incongruent condition. In addition, the main effect of stimulus type was significant ($F_{(3, 834)} = 358.61$, $p < .001$) consistent with previous research in numerical cognition. Mean RTs (collapsed across congruency) were 718 ms (*SD* = 197 ms) for digits, 430 ms (*SD* = 94 ms) for the dice pattern, 404 ms (*SD* = 105 ms) for the NC1 pattern and 424 ms (*SD* = 131 ms) for the NC2 pattern. Post-hoc comparisons revealed significant differences between digits and all dot patterns (all *p*s < .001) and between the dice patterns and the NC1 patterns ($p < .05$).

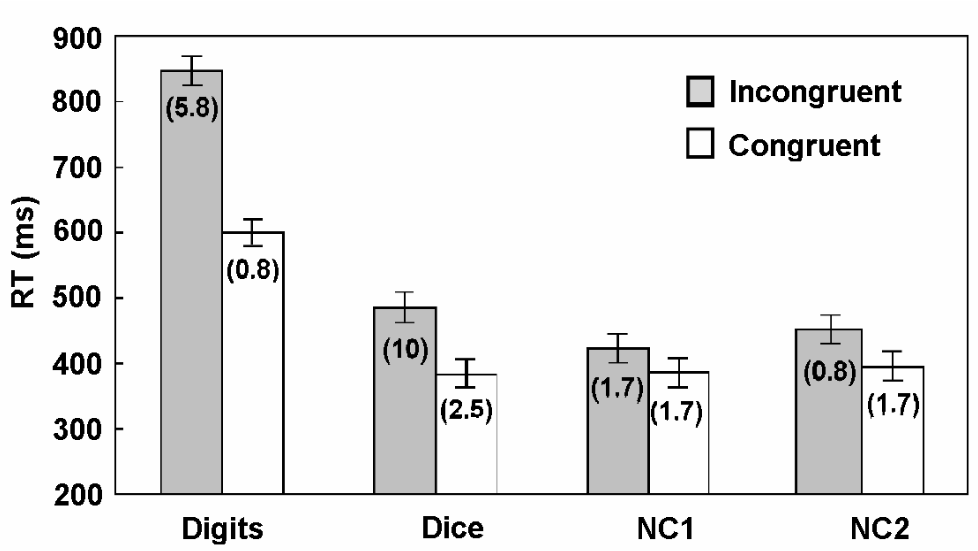

Figure V.2.: Mean RTs as a function of congruency across type of stimuli. The shaded bars represent the incongruent condition, while the white bars represent the congruent condition. Error bars represent 2 standard deviations from the mean.

The stimulus type x congruency interaction was also significant ($F_{(3,\ 834)} = 35.09$, $p < .001$), with digits and dice patterns showing greater interference (247 ms for digits and 102 for dice patterns) than the two non-canonical patterns (38 ms and 56 ms for the NC1 and NC2 patterns, respectively), suggesting that the congruency effect is modulated by stimulus type. Crucially, however, the incongruent condition was significantly slower for all four stimulus types (all *p*s < .01), demonstrating the presence of a synesthetic Stroop effect for each of the stimulus configurations, despite the fact that NM denied experiencing colors for dot patterns. Finally, the stimulus type x congruity x numerosity interaction also reached significance ($F_{(15,\ 834)} = 2.09$, $p < .01$). Separate analyses for each stimulus type showed that congruency x numerosity interaction was significant only for dice and NC2 patterns ($p = .028$ and $p = .015$, respectively).

Although the analysis of variance is a very robust statistical method, running it on a single case violates basic assumptions of data independency (Basso, Salmaso, & Pesarin, 2006; Basso, Chiarandini, & Salmaso, 2007). We thus compared a simplified ANOVA (with congruency and stimulus type as factors) with a permutation test analysis (10000 permutations) to provide independent verification of the observed congruency and stimulus type effects. Since only a limited number of factors may be introduced in a permutation analysis we focused our permutation analysis on those stimuli that yielded the smallest effects in our ANOVA; thus the demonstration of a significant effect in the permutation analysis allows us to conclude that larger effects would also be significant if we were to test them with the permutation analysis. Both the permutation analysis and the ANOVA yielded substantially similar results, with both main effects and the interaction being significant at $p < 0.005$, confirming the robustness of observed effects.

Numerosity Influence

Since the congruency x numerosity interaction was significant for the dice and the NC2 patterns in the overall analysis, a second ANOVA was run on the Numerosity Stroop task only. To this end, numerosities were subdivided in two ranges: small (numerosities 1, 2 and 3) and large (4, 5 and 6). Thus the factors introduced in the analysis were: stimulus type (Dice pattern, NC1 and NC2), congruency (congruent or incongruent) and numerical range (small and large numerosities). Mean RTs for the congruent and incongruent conditions (across numerosities and stimulus types) were 388 ms (*SD* = 85 ms) and 451 ms (*SD* = 126 ms) respectively, whereas for the small and large numerosities (across congruence levels and stimulus types) RTs were 414 ms (*SD* = 114 ms) and 424 ms (*SD* = 108 ms), respectively. The factors stimulus type and congruency were both significant ($F_{(2, 649)} = 4.74$, $p < .01$ and $F_{(1, 649)} = 62.50$, $p < .001$) demonstrating that responses were faster for the dice patterns than for the two non-canonical patterns and replicating the significant synaesthetic Stroop effect reported above. As in the overall ANOVA, the stimulus type x congruency interaction was significant ($F_{(2, 649)} = 5.25$, $p < .01$) with the synaesthetic Stroop effect being stronger in the dice condition than in the other two conditions. Finally the congruency x range interaction was also significant ($F_{(5, 649)} = 9.76$, $p < .005$; see Figure V.3.) indicating that the interference was stronger for smaller numerosities than for larger numerosities (90 ms and 36 ms, respectively). In separate analyses for both ranges, congruency remained significant (both *p*s < .001). These results provide two arguments in favor of a semantic interpretation of NM's experiences. First, the slower observed RTs in all four incongruent conditions demonstrate that the interference occurs independent of notation, a hallmark of semantic processing. Second, the finding that smaller number elicited stronger effects indicates that the concurrents are elicited at a semantic level, since small numerosities are processed faster and with greater accuracy (a process known as subitizing (Mandler, 1982) than larger numbers thus eliciting the synesthetic color faster, and correspondingly greater interference with the physically presented color.

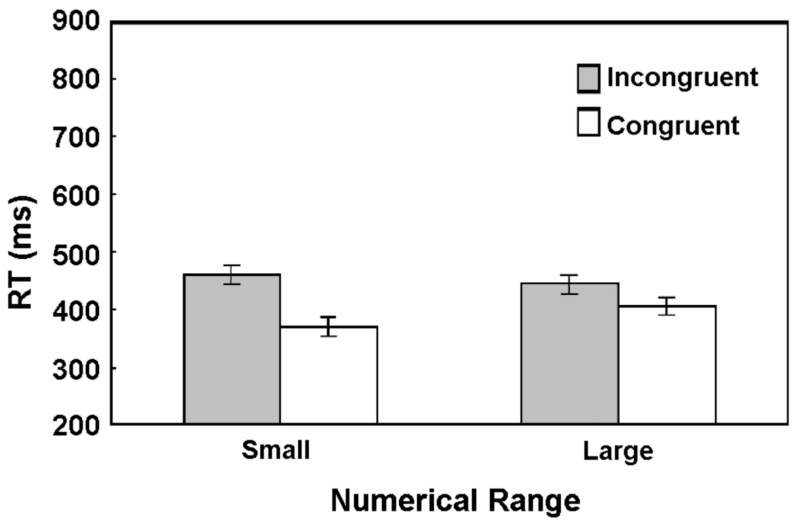

Figure V.3.: Mean RTs as a function of congruency across the numerical range. The shaded bars represent the incongruent condition, while the white bars represent the congruent condition. Error bars represent 2 standard deviations from the mean.

Baseline tasks

Given that dice patterns are over learned, and are therefore likely to be processed differently than the non-canonical patterns, after the Stroop task we asked NM to verbally report the number of dots for neutrally colored (black on a gray background) versions of the dice, NC1 and NC2 patterns used in the main experiment. This baseline task also allowed us to test the learning effect of the non-canonical configurations used during the Numerosity Stroop task. An ANOVA for the enumeration baseline was run introducing stimulus type (dice, NC1, NC2) and numerosity (1 to 6) as factors. Both main effects were significant and as expected, RTs are longer for larger numerosities (numerosity: $F_{(5, 325)} = 108.70$, $p < .001$; stimulus type: $F_{(2, 325)} = 83.42$, $p < .001$; Figure V.4.). Moreover, post-hoc comparisons for stimulus type indicated that all types of configurations were significantly different from each other (all *ps* < .05) with the dice patterns being enumerated faster than the non-canonical patterns and NC2 patterns also being enumerated faster than NC1 patterns. Furthermore, the main effect of numerosity showed a classical increasing curve for increasing numerosities. The stimulus type x numerosity interaction was also significant ($F_{(10, 325)} = 10.018$, $p < .001$). The increasing curve was only true for the NC1 and NC2 patterns since RTs for the dice patterns, as expected, were constant for all numerosities (Wolters et al., 1987). The difference between the enumeration curves for the NC1 and NC2 patterns and that for the dice patterns indicates that the non-canonical configurations were indeed unfamiliar to NM and could have not been previously associated with colors in long term memory.

However, it could be argued that the association with colors was due to repeated exposure during the experiment and that patterns were better recognized at the end compared with the beginning of the experiment. In order to assess whether NM had learned the patterns and could have recalled them directly from memory in the second session, we performed a second ANOVA on enumeration RTs including session (first or second) as a factor. In this new analysis, in addition to the previous main effects, the main effect of session was significant ($F_{(1,\ 307)}$ = 10.429, $p < .01$) indicating that during the second session NM was overall faster (session 1: mean = 346 ms, *SD* = 98 ms; session 2, mean = 331 ms, *SD* = 98 ms). However none of the interactions with session as a factor approached significance ($p > .5$). Separate analyses confirmed the presence of a significant stimulus type x numerosity interaction in both sessions (both *ps* < .001).

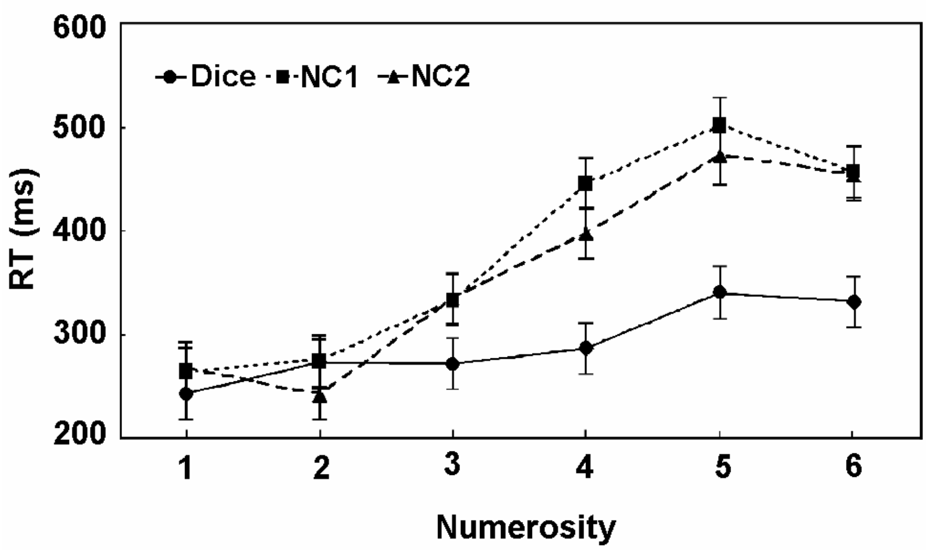

Figure V.4.: Mean RTs for the six numerosities for each of the three dot configuration (dice, NC1 and NC2). The dots represent the dice patterns, the squares represent the NC1 patterns, and the triangle the NC2 patterns. Error bars represent 2 standard deviations from the mean.

Taken together, the findings that enumeration of both non-canonical patterns was significantly slower than for dice patterns and that they did not become over-learned during the experimental sessions strengthens the argument that the results obtained in the Stroop task are not simply due to an association between particular patterns and colors. Rather these results argue that synesthetic interference generalizes to other numerically based stimuli (at least for NM), including novel dot patterns, further arguing for a semantic locus for the observed effects.

In addition, we tested NM on a Color Naming baseline to test for the presence of differences in the time required to name each color. Color naming RTs (after exclusion of errors and outlier removal using a 2 *SD* cutoff) were analyzed with a one-way ANOVA with color as a factor (6 different color names). The main effect of color was significant ($F_{(5, 223)} = 6.934$, $p < .01$) and post-hoc comparisons indicated that the differences were a result of color names brown and grey being slower to name than the others (brown was significantly slower than blue and green, $p < .05$ and .01 respectively, whereas grey was significantly slower than all colors except for brown, all $ps < .01$). It seems unlikely that the interference effects obtained in both the Digit and the Numerosity Stroop tasks were mediated by the slower reaction times to brown and grey, given that we presented all of the numbers and dots in all of the colors. However, in order to rule out this possibility, we performed a new ANOVA introducing mean naming time for each color (according to each experimental session) as a covariate. The factors included were congruency (congruent and incongruent), stimulus type (digits, dice, NC1 and NC2) and number (1 to 6). The covariate did not reach significance ($p = .623$) whereas the congruency and stimulus type effects were still significant ($F_{(1, 833)} = 190.075$, $p < .001$ and $F_{(3, 833)} = 266.211$, $p < .001$). Moreover, the interactions also remained significant: Congruency x Stimulus Type ($F_{(3, 833)} = 35.054$, $p < .001$) and Congruency x Stimulus Type x Number $F_{(15, 833)} = 2.087$, $p < .01$). These results rule out the possibility that the interference effects were the result of differences in color naming times.

Permutation analysis and result validation

Although the analysis of variance is a very robust statistical method, running it on a single case violates basic assumptions of data independency (21, 22). For this reason, we performed a permutation test analysis (10000 permutations) and an ANOVA with congruency (congruent and incongruent) and stimulus type (dice, NC1 and NC2) as factors[6] to compare and validate the results of the ANOVA. We focused our permutation analysis on those stimuli that yielded the smallest effects in our ANOVA; thus the demonstration of a significant effect in the permutation analysis allows us to conclude that larger effects would certainly be significant if we were to test them with the permutation analysis. Both the permutation analysis and the ANOVA yielded substantially similar results, with both main effects and the interaction being significant at a p-value < 0.005, confirming the robustness of observed effects.

---

[6] Only a limited number of factors and levels can be introduced into a particular permutation analysis.

Color Naming Baseline

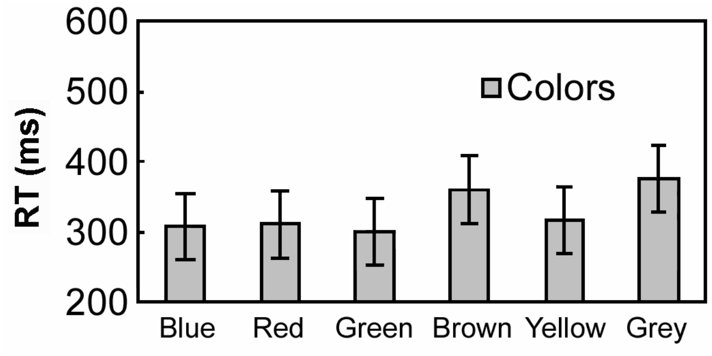

Figure V.5.: Mean for the six colors that NM associates to each number from one to six. Error bars represent 2 standard deviations from the mean.

NM was tested on a Color Naming baseline to test for the presence of differences in the time required to name each color (Figure V.5.). Color naming RTs (after exclusion of errors and outlier removal using a 2 *SD* cutoff) were analyzed with a one-way ANOVA with color as a factor (6 different color names). The main effect of color was significant ($F_{(5, 223)} = 6.934$, $p<.01$) and post-hoc comparisons indicated that the differences were a result of color names brown and grey being slower to name than the others (brown was significantly slower than blue and green, $p < .05$ and .01 respectively, whereas grey was significantly slower than all colors except for brown, all *ps* < .01). It seems unlikely that the interference effects obtained in both the Digit and the Numerosity Stroop tasks were mediated by the slower reaction times to brown and grey, given that we presented all of the numbers and dots in all of the colors. However, in order to rule out this possibility, we performed a new ANOVA introducing mean naming time for each color (according to each experimental session) as a covariate. The factors included were congruency (congruent and incongruent), stimulus type (digits, dice, NC1 and NC2) and number (1 to 6). The covariate did not reach significance ($p = .623$) whereas congruency and stimulus type effects were still significant ($F_{(1, 833)} = 190.075$, $p < .001$ and $F_{(3, 833)} = 266.211$, $p < .001$). Moreover, the interactions also remained significant: Congruency x Stimulus Type ($F_{(3, 833)} = 35.054$, $p< .001$) and Congruency x Stimulus Type x Number $F_{(15, 833)} = 2.087$, $p < .01$). These results rule out the possibility that the interference effects were the result of differences in color naming times.

I.2. Conclusions.

In this paper we presented the case of NM, a grapheme-color synesthete for whom digits, but not dot patterns, elicited the subjective experience of colors. Our results demonstrate Stroop-like interference for incongruently colored stimuli both when NM performed a digit Stroop task and when he performed a numerosity Stroop task with dice and non-canonical patterns, despite the fact that he actively denies any conscious experiences of color for dot patterns. Moreover, both interference and facilitation were stronger for the smaller numerosities than for larger ones. These results suggest that NM may be a “higher” synesthete for whom the associations are explicit for digits but implicit for other numerical stimuli.

These results differ from those presented by Ward and Sagiv (2007) even though both studies suggest the same synesthetic locus of induction. Their synesthetic participant, TD, explicitly reported colors for digits, dice patterns and fingers, and demonstrated interference for all three types of stimuli. However, he reported that he did not experience colors for random dot patterns, and correspondingly did not show an interference effect with these patterns, resulting in a tight correspondence between subjective experience and objective measures. In contrast, we find for the first time a synesthetic Stroop effect with non-cannonical dot patterns. Thus, our participant NM demonstrates a dissociation between these measures in the case of dot patterns. One possible explanation for the difference between the current findings and Ward and Sagiv is a difference in statistical power, as we presented twice as many stimuli per cell. However, given that the non-significant numerical trend in their data was the opposite of what we observed with NM, we do not believe that power alone can explain these different findings. A second possibility is that the way their random configurations were created could have generated some overlap with the dice patterns introducing an undesired bias in RTs (N. Sagiv, personal communication 2007).

Cohen Kadosh *et al.* (Cohen Kadosh et al., 2005; Cohen Kadosh & Henik, 2006a; Cohen Kadosh & Henik, 2006b; Cohen, Cohen, & Henik, 2007) have argued that the interference due to higher synesthesia could be bi-directional, even though conscious reports of synesthetic experiences are almost universally uni-directional. In a single-case study, Cohen Kadosh *et al*. (Cohen Kadosh et al., 2005) tested the claim that numbers elicit colors, but not the other way around. They presented two digits that were colored either congruently or incongruently relative to the synesthete’s photisms, and asked participants to indicate which of the two digits represented the larger magnitude with a key press on the corresponding side. In the incongruent condition, the ink colors corresponded either to digits closer together or further apart than those presented for the numerical judgment (e.g. a 4 and a 6 in the colors of a 1 and a 9). In the incongruent condition, where the colors were those of numbers further apart, RTs were faster than when the same digits were congruently colored. They concluded that the incongruent colors were thus facilitating the judgment by eliciting a greater numerical distance, suggesting an implicit bi-directional activation.

Cohen Kadosh and collaborators have subsequently presented additional data supporting an implicit bi-directional link between colors and analog magnitude in conditions where colors not only created interference with numerical processing but also influenced physical magnitudes judgments in absence of numerical information (Cohen Kadosh & Henik, 2006a; Cohen Kadosh & Henik, 2006b). In an attempt to rule out a possible learning-based account of their findings, Cohen Kadosh *et al.* (Cohen Kadosh et al., 2005) trained non-synesthetic participants in five one-hour sessions to associate numbers with colors. However, it is clear that five hours of training cannot mimic a lifetime of synesthetic experiences. Although these results have been taken as evidence for an implicit bi-directional activation between colors and digits, we argue that such conclusions are premature. Based on our own results in a uni-directional paradigm, we have shown that, despite the absence of overt color report for dice and non-canonical patterns, NM was slower for those stimuli when they were colored incongruently with his corresponding digit photisms.

Given the presence of both implicit uni- and bi-directional interference effects in synesthesia, some account of how such interference arises must be given. One possibility is that implicit bi-directionality in synesthesia might be due to neural connections between color and numerical representations, but which are strong enough to lead to interference, but weak enough that such connectivity does not elicit a conscious experience (Hubbard & Ramachandran, 2005; Cohen Kadosh & Henik, 2007). However, another possibility, which has not been sufficiently considered, is that these interference effects are cognitive consequences of the primary synesthetic connections, which do lead to conscious experiences (see Figure V.6.). That is, the semantic effects we have observed here are not a direct consequence of synesthesia *per se* (implicit synesthesia), but rather are secondary consequences of a lifetime of associations between digits, colors and numerical magnitudes (pseudosynesthesia).

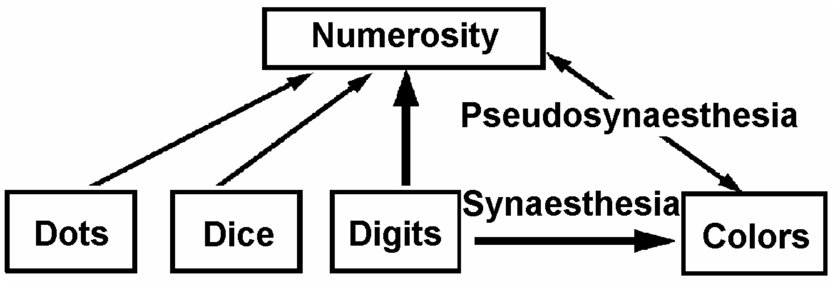

Figure V.6.: Schematic representation of synesthetic and pseudosynesthetic connections between areas. The bold arrows indicate the direct pathways between Arabic numerals and numerosity processing and between Arabic numerals and colors, present in grapheme-color synesthesia. The thinner arrows from dice and dot patterns to numerosity indicate other possible pathways to access numerosity information, while the double-headed arrow indicates secondary pseudosynesthetic associations between numerical information and colors, as a result of a lifetime of experience in which Arabic numerals simultaneously elicit color and magnitude information.

Neuroimaging and neuroanatomical methods suggest that the primary linkage in number-color synesthesia is due to cross-activation between graphemic representations and color representations in the fusiform gyrus (Hubbard & Ramachandran, 2005; Rouw & Scholte, 2007) Given the existence of such primary linkages, each time a number-color synesthete looks at an Arabic digit, he or she also automatically experiences a color and simultaneously activates the numerical magnitude associated with that digit. Because of the constant association between magnitudes and colors, the two may become associated within a broader cognitive system, despite the absence of conscious links between them (for a similar line of reasoning applied linkages between different types of synesthesia, see Simner & Hubbard, 2006). Indeed, both fMRI adaptation in human participants (Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004) and single-unit recordings in monkeys (Diester & Nieder, 2007) have demonstrated that numerosity stimuli (dot patterns) and Arabic digits map onto the same neurons in parietal and prefrontal cortices. In this manner, spreading activation within the semantic network may account for the presence of color interference for dot patterns, and similarly, for the findings of “implicit bi-directionality” reported in previous studies. In order to distinguish between conscious synesthetic reports and non-conscious associations that may develop with repeated associative learning, we suggest that the latter be referred to as pseudosynesthesia, indicating the fact that they mimic synesthesia, without giving rise to conscious experiences, one of the defining features of synesthesia. It is therefore important to differentiate between primary synesthetic connections between areas, which give rise to secondary experiences and are probably genetically induced, from those connections that are the consequence of the consistent experience of color each time a digit is seen, creating secondary semantic links.

In the only study to directly compare Stroop effects due to conscious synesthetic associations with equally strong pseudosynesthetic associations, Elias and colleagues (Elias, Saucier, Hardie, & Sarty, 2003) compared a single synesthete to two types of controls: a non-synesthete expert in cross-stitching with eight years of experience for whom colored threads where associated to digits, and a group of four non-synesthetic control participants with no color-number associations. In three behavioral tasks which tested for the automaticity of color-digit associations – a standard digit Stroop task, a mathematical Stroop task, and a priming task – the congruency between color and digit was manipulated according to the individual associations of each participant. When considering only simple RT measures, the synesthete and the expert in cross-stitching showed equivalent patterns of interference in incongruent conditions in all three tasks, despite the fact that the cross-stitcher reported no conscious experience of colors in response to digits. As defined here, the cross-stitcher is a clear example of pseudosynesthesia. Moreover, this study suggests that individuals who learn number-color associations over a sufficiently long period of time may perform similarly to synesthetes and be subject to interference when performing synesthetic Stroop tasks.

Similarly, MacLeod and Dunbar (1988), trained a group of 22 participants to name geometrical shapes using color names. The authors report that five hours of training were sufficient to create an interference effect while performing a Stroop type task with congruently/incongruently colored shapes. After 20 hours of practice, not only was the interference very strong, but one participant even claimed that the white shapes began to take on the colors of their associated color names. The fact that one of MacLeod and Dunbar participants reported colors to shapes after 20 hours of training suggests that more extensive training may mimic synesthetic associations in particular individuals.

A similar associative learning explanation could account for results of Dixon *et al.'s* (Dixon et al., 2000) mathematical Stroop task described in the introduction (see also Jansari et al., 2006). Certain models of numerical cognition suggest that numbers activate a rich network of associations that, in the context of a given task, includes both relevant and irrelevant information (Campbell, 1994; Campbell, Parker, & Doetzel, 2004). As experience increases, retrieval of simple arithmetical problems becomes specialized to optimize performance, creating preferential links between related digits (such as between 5, 2, and 7). Every time a synesthete retrieves a given arithmetic problem, viewing or thinking of digits will also elicit the relevant colors creating a link not only between the operands and the results, but also for the appropriate sequence of colors. Therefore the interference observed by Dixon *et al*. may not be due to a conscious experience of color elicited by the arithmetic solution (and indeed, Dixon *et al*. do not state whether C reported any conscious photisms) but rather may be due to associative priming. Rather than numbers activating a conceptual representation, we argue that the operands may have activated both the numerical information and the color which would generate the interference in parallel. This conclusion is supported by similar results in an identical paradigm tested on Elias *et al.'s* cross-stitcher who may have similar cognitive associations, but has never reported any synesthetic experience (Elias et al., 2003).

These considerations highlight that Stroop tasks, when used as objective markers for synesthesia in the absence of corresponding subjective reports, must be treated with caution (for a related argument see Smilek & Dixon, 2002). We are not questioning the use of Stroop-like paradigms as a method of validating subjective reports (indeed they have been highly useful), but question their direct application to the structure of synesthetic representations in the absence of such subjective reports, given that the entire cognitive system will be modified by the repeated synesthetic experiences. In the absence of longitudinal and/or developmental studies of synesthesia, it is impossible with a Stroop-like task to disentangle primary, direct consequences of synesthesia, from secondary adaptations to a lifetime of altered sensory experience. We argue that results obtained with Stroop paradigms only demonstrate the presence of an association, which could be either synesthetic or pseudosynesthetic. Distinguishing between these two possibilities can only be done through the use of subjective reports.

In sum, we stress that synesthesia is traditionally defined as the union of the senses where a perception in one modality triggers a second conscious experience in a non-stimulated modality. We suggest that in the quest for understanding synesthesia, reports of additional sensations should be explicit, since it is only then that we can distinguish between synesthetic phenomena and over-learned associations. More generally, following Merikle and colleagues (Cheesman & Merikle, 1986; Merikle et al., 2001), we argue that studies of unusual experiences should depend on not only objective measures, but also on subjective report, especially given that explicit and implicit processing may yield qualitatively different effects (Cheesman & Merikle, 1986), or, as we suggest here, may even arise from qualitatively different mechanisms.

Chapter 6: General Conclusions.

The general topic of this thesis, that is the representation of numerical information in humans, was investigated from three different perspectives. The first was to study how numbers are represented by children. Two questions were addressed: the first was to observe the developmental pattern of numerical representation in preschoolers; the second was to differentiate a specific numerical representation from the representation of non-numerical ordered sequences. The results of the first study demonstrated that the understanding of how numbers map onto space develops long before formal education begins. At first, when a young child is asked to position numbers on a line she will rely on an intuitive representation where smaller numbers are overrepresented and larger numbers are clustered together. With familiarization to numbers and practice, the child learns progressively to position them regularly in a linear manner. Moreover, the ability to position numbers linearly is dependent of the context. If the context given is too large or unfamiliar, the child will go back to a intuitive representation. Support to the claim that context plays an important role can be found in the second study, where children failed on a numerically small interval (1-7) that did not conform to the base-10 reference system. Indeed, for this interval not only the numbers had to be known but children had to deploy a numerical representation independent from their familiar reference scheme.

The focus of the second study was on the similarity between numerical and non-numerical sequences. On the one hand, the developmental pattern observed with non-numerical sequences showed to be similar to the development of the numerical representation when tested with the positioning task. Indeed, the youngest children seemed to have trouble in positioning items on the lines and progressively relied on an informal representation and then shifted to a linear positioning. The intermediate stage of logarithmic positioning was present, although in a small percentage. This was explained by the testing period that occurred at the end of the school year allowing extensive practice to 1st graders compared to preschoolers, and by the smaller interval range for the non-numerical sequences. Additionally, the ability in the numerical domain was predictive of the ability in the non-numerical domain. On the other hand, the results were different when children were tested on the mental bisection task. Different performance patterns were observed as a function of the tested sequence. Only the bisection of numerical intervals showed the classical signature of compressed numerical representation. The younger children tended to underestimate more the true midpoint compared to the older children and they were also influenced by the length of the interval: that is, the longer the interval the more they were biased.

The different performance observed for the different sequences in the mental bisection task and the predictive character of the numerical ability on performance in the non-numerical positioning tasks suggests that children generalize the linearity principle. They first learn to represent items linearly in the numerical domain and only after they extend this concept to other sequences. This result is of particular interest because the processing of non-numerical sequences has been shown to produce effects that are remarkably similar to those observed in the numerical domain. For example, Gevers and collaborators (2003; 2004) have shown a SNARC-like effect with non-numerical material. Moreover, others (Fias et al., 2007; van Opstal et al., 2007) have shown that ordinal sequences activate the same brain regions (i.e., the intraparietal sulcus) that are usually activated when manipulating numerical information or performing arithmetic tasks. These results could be explained by the generalization assumption: children start to understand the concept of linearity in the numerical domain and with time they generalize it to other domains such that a stable associations between the non-numerical sequence and the numerical concept is created; as a consequence to this continuous association the brain areas specific to number processing start to be active for non-numerical sequences even in the absence of numerical information.

The second perspective adopted to understand humans' numerical representation was to study skilled adults with task requiring to manipulate both symbolic and non-symbolic numerical information. The first aim was to estimate the individual value of internal Weber fraction through a numerosity discrimination task and to relate this value to other tasks that putatively rely on a common (log-Gaussian) representation as shown by Weber's law signature. The second was to further understand the influence of the format presentation in an approximate calculation task.

Surprisingly, although performance in the number comparison task shows Weber's law signature (size and distance effect), reaction times failed to correlate with the Weber value. One could argue that this failure was the result of incompatible indices. Indeed, performance in the number comparison task was indexed by RTs, whereas the Weber value was estimated from the accuracy in the numerosity discrimination task. Alternatively, it could be argued that the symbolic comparison task relies only partially on the internal log-Gaussian representation. In support of the latter hypothesis, different models of numerical processing are now leading to the conclusion that adult numerical abilities are underpinned by two separate systems: one would be language-based and activated for all exact numerical tasks and the second would be language independent, necessary for numerical approximate processing and phylogenetically determined (Verguts, Fias, & Stevens, 2005; Verguts & Fias, 2004; Spelke, 2000; Zorzi, Stoianov, Becker, Umiltà & Butterworth, submitted). For instance, the numerosity code model assumes a linear representation of numerosity, in a similar way to a "thermometer" representation (Zorzi et al., 2008). Each numerosity set is represented by a corresponding number of nodes which contains the smaller sub-sets. This model successfully explains the distance and size effect in number comparison. Moreover, the authors suggest that this model should coexist with an approximate representation since cardinality is not the only type of mental representation of numbers. Further support to the hypothesis of a dual system comes from a neuropsychological study (Lemer, Dehaene, Spelke, & Cohen, 2003) on two patients with numerical deficits. The first patient had preserved abilities in approximation tasks but failed in symbolic tasks. The second patient showed the reverse pattern of deficit, with spared exact calculation and deficits in approximate tasks. Therefore, in our study, when comparing Arabic numbers instead of numerosities, an exact numerical system would be active and therefore hide the numerical representation described by the Weber value.

Moreover, in the study with adults, the data allowed to formulate a new hypothesis on what McCrink and collaborators (2007) termed the *operational momentum*. They observed that when performing approximate additions, participants tended to overestimate the correct result whereas they preferred a smaller outcome when performing subtractions. To them, this phenomenon was comparable to the *representational momentum* which is the tendency to remember a position of a moving object further in the direction of movement than it really was. McCrink and collaborators (2007) suggest that computing additions and subtractions would be equivalent to moving on the internal numerical representation as if it was a physical line. Therefore, the movement induced by the operation (forward for addition and backward for subtraction) would yield the overestimation and underestimation. However, by crossing symbolic and non-symbolic formats in our experiment, the results suggested an alternative explanation to the calculation biases although not mutually exclusive with McCrinck and collaborator's hypothesis. Indeed, numerous studies have shown that dot patterns are generally underestimated and the degree of underestimation is proportional to set size (Ginsburg, 1978; Izard & Dehaene, 2007; Allik & Tuulmets, 1991; Durgin, 1995). Larger sets are more underestimated than smaller ones. In a purely non-symbolic approximation task, the data can be accounted for more economically by an underestimation bias. When solving additions, the operands as well as the outcomes are underestimated and since the proposed outcomes are numerically larger than the operands (always the case in additions), the latter were more underestimated, leading participants' choice towards the larger sets (overestimation). In subtraction, the numerically larger set, the more underestimated, was in the operation and led participants to a much smaller approximated answer than the correct one. Therefore, when choosing among the possible outcomes, participants' choice preferentially fell on the numerically smaller set.

Finally, in the third perspective adopted in this work we studied a NM, a number-colour synaesthete, for whom digits, but not dot patterns, elicited the subjective experience of colors. In a Stroop-like task with numbers and dot patterns interference effects were observed despite the fact that he actively denied any conscious experience of color for dot patterns. These results suggest that NM may be a "higher" synesthete for whom the associations are explicit for digits but implicit for other numerical stimuli.

Other studies carried out by Cohen Kadosh *et al.* (Cohen Kadosh et al., 2005; Cohen Kadosh & Henik, 2006a; Cohen Kadosh & Henik, 2006b; Cohen, Cohen, & Henik, 2007) have argued that the interference due to higher synesthesia could be bi-directional, even though conscious reports of synesthetic experiences are almost universally uni-directional. In a single-case study, Cohen Kadosh *et al*. (Cohen Kadosh et al., 2005) tested the claim that numbers elicit colors, but not the other way around. They presented two digits that were colored either congruently or incongruently relative to the synesthete's photisms, and asked participants to indicate which of the two digits represented the larger magnitude with a key press on the corresponding side. In the incongruent condition, the ink colors corresponded either to digits closer together or further apart than those presented for the numerical judgment (e.g. a 4 and a 6 in the colors of a 1 and a 9). In the incongruent condition, where the colors were those of numbers further apart, RTs were faster than when the same digits were congruently colored. They concluded that the incongruent colors were thus facilitating the judgment by eliciting a greater numerical distance, suggesting an implicit bi-directional activation.

Given the presence of both implicit uni- and bi-directional interference effects in synesthesia, some account of how such interference arises must be given. One possibility is that implicit bi-directionality in synesthesia might be due to neural connections between color and numerical representations, but which are strong enough to lead to interference, but weak enough that such connectivity does not elicit a conscious experience (Hubbard & Ramachandran, 2005; Cohen Kadosh & Henik, 2007). However, we argue that these interference effects are cognitive consequences of the primary synesthetic connections, which do lead to conscious experiences. That is, the semantic effects we have observed in our study are not a direct consequence of synesthesia *per se* (implicit synesthesia), but rather are secondary consequences of a lifetime of associations between digits, colors and numerical magnitudes (pseudosynesthesia).

Finally, we stress that synesthesia is traditionally defined as the union of the senses where a perception in one modality triggers a second conscious experience in a non-stimulated modality. We suggest that in the quest for understanding synesthesia, reports of additional sensations should be explicit, since it is only then that we can distinguish between synesthetic phenomena and over-learned associations. More generally, following Merikle and colleagues (Cheesman & Merikle, 1986; Merikle et al., 2001), we argue that studies of unusual experiences should depend on not only objective measures, but also on subjective report, especially given that explicit and implicit processing may yield qualitatively different effects (Cheesman & Merikle, 1986), or, as we suggest here, may even arise from qualitatively different mechanisms.

FIGURE INDEX

CHAPTER I.

CHAPTER II.

CHAPTER III.

Table Index

Chapter II.

Chapter III.

Chapter IV.

Reference List

Allik, J. & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception and Psychophysics, 49,* 303-314.

Allik, J. & Tuulmets, T. (1993). Perceived numerosity of spatiotemporal events. *Perception and Psychophysics, 53,* 450-459.

Allik, J., Tuulmets, T., & Vos, P. G. (1991). Size invariance in visual number discrimination. *Psychological Research, 53,* 290-295.

Antell, S. E. & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development, 54,* 695-701.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition, 44*, 75-106.

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

Banks, W. P., Milton, F., & Fortunee, K.-S. (1976). Semantic congruity effects in comparative judgment of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 435-447.

Barnett, K. J., Finucane, C., Asher, J. E., Bargary, G., Corvin, A. P., Newell, F. N. et al. (2007). Familial patterns and the origins of individual differences in synaesthesia. *Cognition*.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J., & Bolton, P. (1996). Synaesthesia : prevalence and familiality. *Perception, 25,* 1073-1079.

Barth, H., Kanwisher, N., & Spelke, E. (2003). The construction of large number representations in adults. *Cognition*, *86,* 201-221.

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition, 98.*

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Science U.S.A, 102,* 14116-14121.

Basso, D., Chiarandini, M., & Salmaso, L. (2007). Synchronized permutation tests in IxJ designs. *Journal of Statistical Planning and Inference*, *137,* 2564-2578.

Basso, D., Salmaso, L., & Pesarin, F. (2006). On synchronized permutation tests in two-way anova. In J.López-Fidalgo, J. M. Rodríguez-Díaz, & B. Torsney (Eds.), *mODa 8 - Advances in Model-orientation design and analysis* (pp. 25-32). Heidelberg: Physica-Verlag HD.

Berch, D. B., Foley, E. J., Hill, R. J., & Ryan, P. M. (1999). Extracting parity and magnitude from Arabic numerals: developmental changes in number processing and mental representation. *Journal of Experimental Child Psychology, 74,* 286-308.

Bijeljac-Babic, R., Bertoncini, J., & Mehler, J. (1993). How do 4-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, *29,* 711-721.

Booth, J. L. & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology, 41,* 189-201.

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition, 83,* 223-240.

Buckley, P. B. & Gillman, C. B. (1974). Comparisons of digits and dot patterns. *Journal of Experimental Psychology, 103,* 1131-1136.

Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.

Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46,* 3-18.

Butterworth, B., Zorzi, M., Girelli, L., & Janckheere, A. R. (2001). Storage and retrieval of addition facts: the role of number comparison. *The Quarterly Journal of Experimental Psychology A, 54,* 1005-1029.

Campbell, J. I., Parker, H. R., & Doetzel, N. L. (2004). Interactive effects of numerical surface form and operand parity in cognitive arithmetic. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 30,* 51-64.

Campbell, J. I. D. (1994). Architectures for numerical cognition. *Cognition, 53,* 1-44.

Canfield, R. L. & Smith, E. G. (1996). Number-based expectations and sequential enumeration by 5-month-old infants. *Developmental Psychology, 32,* 269-279.

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,* e125.

Carey, S. (2001). Cognitive foundations of arithmetic: evolution and ontogenesis. *Mind & Language, 16,* 37-55.

Cheesman, J. & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology, 40,* 343-367.

Clearfield, M. W. (2004). Infants' enumeration of dynamic displays. *Cognitive Development, 19,* 309-324.

Clearfield, M. W. & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science, 10,* 408-411.

Clearfield, M. W. & Mix, K. S. (2001). Amount versus number: infants' use of area and contour length to discriminate small sets. *Journal of Cognition and Development, 2,* 243-260.

Cohen Kadosh, R. & Henik, A. (2006a). Color congruity effect: where do colors and numbers interact in synaesthesia? *Cortex, 42,* 259-263.

Cohen Kadosh, R. & Henik, A. (2006b). When a line is a number: Color yields magnitude information in a digit-color synesthete. *Neuroscience, 137,* 3-5.

Cohen Kadosh, R. & Henik, A. (2007). Can synaesthesia research inform cognitive science? *Trends in Cognitive Sciences, 11,* 177-184.

Cohen Kadosh, R., Sagiv, N., Linden, D. E. J., Robertson, L. C., Elinger, G., & Henik, A. (2005). When blue is larger than red: Colors influence numerical cognition in synesthesia. *Journal of Cognitive Neuroscience, 17,* 1766-1773.

Cohen Kadosh, R., Tzelgov, J., & Henik, A. (2008). A synesthetic walk on the mental number line: The size effect. *Cognition*.

Cohen, K. R., Cohen, K. K., & Henik, A. (2007). The Neuronal Correlate of Bidirectional Synesthesia: A Combined Event-related Potential and Functional Magnetic Resonance Imaging Study. *Journal of Cognitive Neuroscience, 19,* 2050-2059.

Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distignuish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin and Review, 8,* 698-707.

Crawford, J. R. & Garthwaite, P. H. (2002). Investigation of the single case in neuropsychology: confidence limits on the abnormality of test scores and test score differences. *Neuropsychologia, 40,* 1196-1208.

Dehaene, S. (1989). The psychophysics of numerical comparison: A reexamination of apparently incompatible data. *Perception and Psychophysics, 45,* 557-566.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44,* 1-42.

Dehaene, S. (2007). Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation. In P.Haggard & Y. Rossetti (Eds.), *Attention & Performance XXII. Sensori-motor foundation of higher cognition.* Cambridge: Harvard University Press.

Dehaene, S. & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 314-326.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122,* 371-396.

Dehaene, S. & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5,* 390-407.

Dehaene, S. & Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Mathematical Cognition, 1,* 83-120.

Dehaene, S., Dupoux, E., & Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 626-641.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20,* 487-506.

Diester, I. & Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. *PLoS Biology, 5,* 1-12.

Dixon, M. J., Smilek, D., Cudahy, C., & Merikle, P. M. (2000). Five plus two equals yellow. *Nature, 406,* 365.

Duncan, E. M. & Mc Farland, C. E. Jr. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory and Cognition, 8,* 612-622.

Durgin, F. H. (1995). Text density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 149-169.

Eger, E., Sterzer, P., Russ, M. O., Giraud, A. L., & Kleinschmidt, A. (2003). A supramodal number representation in human intraparietal cortex. *Neuron, 37,* 719-725.

Elias, L. J., Saucier, D. M., Hardie, C., & Sarty, G. E. (2003). Dissociating semantic and perceptual components of synaesthesia : behavioural and functional neuroanatomical investigations. *Cognitive Brain Research, 16,* 232-237.

Feigenson, L. (2005). A double-dissociation in infants' representations of object arrays. *Cognition, 95,* B37-B48.

Feigenson, L., Carey, S., & Spelke, E. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology, 44,* 33-66.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core system of number. *Trends in Cognitive Sciences, 8,* 307-314.

Freyd, J. J. & Finke, R. A. (1984). Representational momentum. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 126-132.

Freyd, J. J., Kelly, M. H., & DeKay, M. L. (1990). Representational momentum in memory for pitch. *Journal of Experimental Psychology: Learning, Memory and .Cognition, 16,* 1107-1117.

Fias, W., Lammertyn, J., Caessens, B., & Orban, G. A. (2007). Processing of abstract ordinal knowledge in the horizontal segment of the intraparietal sulcus. *Journal of Neuroscience, 27,* 8952-8956.

Gallistel, C. R. & Gelman, I. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences, 4,* 59-65.

Gallistel, C. R. & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43-74.

Galton, F. (1880a). Visualised Numerals. *Nature,* 494-495.

Galton, F. (1880b). Visualised Numerals. *Nature,* 252-256.

Gelman, R. (2006). Young natural-number arithmeticians. *Current Directions in Psychological Science, 15,* 193-197.

Gelman, R. & Butterworth, B. (2005). Number and language: how are they related? *Trends in Cognitive Neurosciences, 9,* 6-10.

Gelman, R. & Meck, E. (1983). Preschoolers' counting: principles before skill. *Cognition, 13,* 343-359.

Gevers, W., Reynvoet, B., & Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition, 87,* B87-B95.

Gevers, W., Reynvoet, B., & Fias, W. (2004). The mental representation of ordinal sequences is spatially organized: evidence from days of the week. *Cortex, 40,* 171-172.

Ginsburg, N. (1978). Perceived numerosity, item arrangement, and expectancy. *American Journal of Psychology, 91*, 267-273.

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology, 76*, 104-122.

Halligan, P. W. & Marshall, J. C. (1988). How long is a piece of string? A study of line bisection in a case of visual neglect. *Cortex, 24,* 321-328.

Halligan, P. W. & Marshall, J. C. (1989). Line bisection in visuo-spatial neglect: disproof of a conjecture. *Cortex, 25*, 517-521.

Hancock, P. (2006). Monozygotic twins' colour-number association: a case study. *Cortex*, *42,* 147-150.

Hannula, M. M. & Lehtinen, E. (2005). Spontaneous focusing on numerosity and mathematical skills of young children. *Learning and Instruction, 15,* 237-256.

Ho, C. S. H. & Fuson, K. (1998). children's knowledge of teen quantities as tens and ones: comparisons of Chinese, British and American Kindergartners. *Journal of Educational Psychology, 90*, 536-544.

Hubbard, E. M. & Arman, A. C. (2005). Individual differences among grapheme-color synesthetes: brain-behavior correlations. Unpublished Work

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Neuroscience, 6,* 435-448.

Hubbard, E. M. & Arman, A. C. (2005). Individual differences among grapheme-color synesthetes: brain-behavior correlations. Unpublished Work

Hubbard, T. L. (2005). Representational momentum and related displacements in spatial memory: A review of the findings. *Psychonomic Bulletin & Review, 12,* 822-851.

Huntley-Fenner, G. & Cannon, R. (2000). Preschoolers' magnitude comparisons are mediated by a preverbal analog mechanism. *Psychological Science, 11*, 147-152.

Indow, T. & Ida, M. (1977). Scaling of dot numerosity. *Perception and Psychophysics, 22*, 265-276.

Izard, V. & Dehaene, S. (2007). Calibrating the mental number line. *Cognition*.

Jansari, A. S., Spiller, M. J., & Redfern, S. (2006). Number synaesthesia: when hearing "four plus five" looks like gold. *Cortex, 42,* 253-258.

Johnson, D. M. (1939). Confidence and speed in the two-category judgement. *Archives of psychology, 241,* 1-52.

Koechlin, E., Dehaene, S., & Mehler, J. (1997). Numerical transformation in five-month-old human infants. *Mathematical Cognition, 3,* 89-104.

Krueger, L. E. (1982). Single judgments of numerosity. *Perception and Psychophysics, 31,* 175-182.

Krueger, L. E. (1984). Perceived numerosity: a comparison of magnitude production, magnitude estimation, and discrimination judgments. *Perception and Psychophysics, 35,* 536-542.

Le Corre, M. & Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition, 105,* 395-438.

Le Corre, M. & Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition, 105,* 395-438.

Le Corre, M. & Carey, S. (2008). Why the verbal counting principles are constructed out of representations of small sets of individuals: A reply to Gallistel. *Cognition, (in press)*.

Lemer, C., Dehaene, S., Spelke, E., & Cohen, L. (2003). Approximate quantities and exact number words: dissociable systems. *Neuropsychologia, 41,* 1942-1958.

Lipton, J. & Spelke, E. (2003). Origins of number sense: large-number discrimination in human infants. *Psychological Science, 14,* 396-401.

Lipton, J. & Spelke, E. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development, 76,* 978-988.

Lipton, J. & Spelke, E. (2006). Preschool children master the logic of number words meanings. *Cognition, 98,* B57-B66.

Lucangeli, D., Iannitti, A., & Vettore, M. (2007). *Lo sviluppo dell'intelligenza numerica.* (1st ed.) Rome: Carocci.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychological Bulletin, 109,* 163-203.

MacLeod, C. M. & Dunbar, K. (1988). Training and Stroop-like interference: evidence for a continuum of automaticity. *Journal of Experimental Psychology. Learning, .Memory & Cognition, 14,* 126-135.

Mandler, G. (1982). Subitizing: An analysis of its component processes. *Journal of Experimental Psychology: General, 111,* 1-22.

Margolis, E. & Laurence, S. (2007). How to learn the natural numbers: Inductive inference and the acquisition of number concepts. *Cognition*.

Marshall, J. C. & Halligan, P. W. (1989). When right goes left: an investigation of line bisection in a case of visual neglect. *Cortex, 25,* 503-515.

Mattingley, J. B., Payne, J. M., & Rich, A. N. (2006). Attentional load attenuates synaesthetic priming effects in grapheme-colour synaesthesia. *Cortex, 42,* 129-320.

Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature, 410,* 580-582.

McCrink, K., Dehaene, S., & Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Perception and Psychophysics, 69,* 1324-1333.

McCrink, K. & Wynn, K. (2004). Large-number addition and subtraction by 9-month-old infants. *Psychological Science, 15,* 776-781.

McCrink, K. & Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychol.Sci., 18,* 740-745.

Meck, W. H. & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 320-334.

Merikle, P. M., Smilek, D., & Eastwood, J. D. (2001). Perception without awareness: Perspectives from cognitive psychology. *Cognition, 79,* 115-134.

Mills, C. B., Boteler, E. H., & Oliver, G. K. (1999). Digit synaesthesia : A case study using a stroop-type test. *Cognitive Neuropsychology, 16,* 181-191.

Minturn, A. L. & Reese, T. W. (1951). The effect of differential reinforcement on the discrimination of visual number. *Journal of Psychology: Interdisciplinary and Applied, 31,* 201-231.

Mix, K. S. (1999). Similarity and numerical equivalence: appearance counts. *Cognitive Development*, *14,* 269-297.

Mix, K. S. (2002). The construction of number concepts. *Cognitive Development*, *17,* 1345-1363.

Mix, K. S., Huttenlocher, J., & Levin, S. C. (1996). Do preschool children recognize auditory-visual numerical correspondences? *Child Development, 67,* 1592-1608.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy : is number one of them? *Psychological Bulletin, 128*, 278-294.

Mix, K. S., Levine, S. C., & Huttenlocher, J. (1997). Numerical abstraction in infants: another look. *Developmental Psychology, 33,* 423-428.

Moore, D., Benenson, J., Reznick, J. S., Peretson, P., & Kagan, J. (1987). Effect of auditory numerical information on infants' looking behavior: Contradictory evidence. *Developmental Psychology, 23,* 665-670.

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

Mussolin, C. & Noel, M. P. (2007). The nonintentional processing of Arabic numbers in children. *Journal of Clinical and Experimental Neuropsychology, 29,* 225-234.

Nieder, A. (2004). The number domain - Can we count on parietal cortex? *Neuron, 44,* 407-409.

Nikolic, D., Lichti, P., & Singer, W. (2007). Color opponency in synaesthetic experiences. *Psychological Sciences, 18,* 481-486.

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C. R., Parslow, D. M., Morgan, M. J. et al. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words. *Nature Neuroscience, 5,* 371-375.

Nuerk, H.-C., Weger, U., & Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition, 82,* B25-B33.

Odgaard, E. C., Flowers, J. H., & Bradman, H. L. (1999). An investigation of the cognitive and perceptual dynamics of a colour-digit synaesthete. *Perception, 28,* 651-664.

Opfer, J. E. & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*.

Paulsen, H. & Laeng, B. (2006). Pupillometry of grapheme-colour synaesthesia. *Cortex, 42,* 290-294.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44,* 547-555.

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

Priftis, K., Zorzi, M., Meneghello, F., Marenzi, R., & Umiltà, C. A. (2006). Explicit vs. implicit processing of representational space in neglect: dissociation in accessing the mental number line. *Journal of Cognitive Neuroscience, 18,* 680-688.

Ramachandran, V. S. & Hubbard, E. M. (2001a). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society of London, 268,* 983.

Ramachandran, V. S. & Hubbard, E. M. (2001b). Synaesthesia - A window into perception, thought and language. *Journal of Consciousness Studies, **8**,* 3-34.

Rips, L. J., Asmuth, J., & Bloomfield, A. (2006). Giving the boot to the bootstrap: how not to learn the natural numbers. *Cognition, 101,* B51-B60.

Rips, L. J., Asmuth, J., & Bloomfield, A. (2007). Do children learn the integers by induction? *Cognition*.

Reynvoet, B. & Brysbaert, M. (1999). Single-digit and two-digit Arabic numerals address the same semantic number line. *Cognition, 72,* 191-201.

Rousselle, L., Palmers, E., & Noël, M. P. (2004). Magnitude comparison in preschoolers: What counts? Influence of perceptual variables. *Journal of Experimental Child Psychology, 87,* 57-84.

Rouw, R. & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience, 10,* 792-797.

Sagiv, N., Simner, J., Collins, J., Butterworth, B., & Ward, J. (2005). What is the relationship between synaesthesia and visuo-spatial number forms? *Cognition*, 1-15.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002a). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002b). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

Sekuler, R. & Mierkiewicz, D. (1977). Children's judgment of numerical inequality. *Child Development, 48,* 630-633.

Seron, X., Pesenti, M., & Noël, M. P. (1992). Images of numbers, or "When 98 is upper left and 6 sky blue". *Cognition, 44,* 159-196.

Siegel, L. S. (1968). The role of experience in judgments of complexity of random shapes. *Canadian Journal of Psychology, 22,* 474-483.

Siegler, R. S. & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development, 75,* 428-444.

Siegler, R. S. & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237-243.

Simner, J. & Hubbard, E. M. (2006). Variants of synesthesia interact in cognitive tasks: evidence for implicit associations and late connectivity in cross-talk theories. *Neuroscience, 143,* 805-814.

Smilek, D. & Dixon, M. J. (2002). Toward a synergistic understanding of synaesthesia. Psyche [On-line]. Available: http://psyche.cs.monash.edu.au/v8/psyche-8-01-smilek.html

Smilek, D., Moffatt, B. A., Pasternak, J., White, B. N., Dixon, M. J., & Merikle, P. M. (2001). Synaesthesia: A case study of discordant monozygotic twins. *Neurocase, 8,* 338-342.

Spelke, E. (2000). Core knowledge. *American Psychologist,* 1233-1243.

Starkey, P. & Cooper, R. G. (1980). Perception of numbers by human infants. *Science, 210,* 1033-1035.

Starkey, P., Spelke, E. S., & Gelman, R. (1983). Detection of intermodal numerical correspondences by human infants. *Science, 222,* 179-181.

Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition, 36,* 97-127.

Strauss, M. S. & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development, 52,* 1146-1152.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Trick, L. M. & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80-102.

van Loosbroek, E. & Smitsman, A. W. (1990). Visual perception in numerosity in infancy. *Developmental Psychology, 26,* 916-922.

van Oeffelen, M. P. & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception and Psychophysics, 32,* 163-170.

Van Opstal, F., Verguts, T., Orban, G. A., & Fias, W. (2007). A hippocampal-parietal network for learning an ordered sequence. *NeuroImage*.

Verguts, T. & De Moor, W. (2005). Two-digit comparison. Decomposed, Holistic, or Hybrid? *Experimental Psychology, 52,* 195-200.

Verguts, T. & Fias, W. (2004). Representation of number in animals and humans: a neural model. *Journal of Cognitive Neurosciences, 16,* 1493-1504.

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review, 12,* 66-80.

Ward, J. & Sagiv, N. (2007). Synaesthesia for finger counting and dice patterns: a case of higher synaesthesia? *Neurocase, 13*, 86-93.

Whalen, J., Gallistel, C. R., & Gelman, R. (1999). Nonverbal counting in humans: The psychophysics of number representation. *Psychological Science, 10,* 130-137.

Wollen, K. A. & Ruggiero, F. T. (1983). Colored-letter synaesthesia. *Journal of Mental Imagery, 7*, 83-86.

Wolters, G., van Kempen, H., & Wijlhuizen, G.-J. (1987). Quantification of small numbers of dots: subitizing or pattern recognition? *American Journal of Psychology, 100,* 225-237.

Wood, G., Nuerk, H. C., & Willmes, K. (2006). Crossed hands and the SNARC effect: a failure to replicate Dehaene, Bossini and Giraux (1993). *Cortex, 42*, 1069-1079.

Wynn, K. (1990). Children's understanding of counting. *Cognition, 36,* 155-193.

Wynn, K. (1992). Children acquisition of the number words and the counting system. *Cognitive Psychology, 24,* 220-251.

Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science, 7,* 164-169.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representation. *Cognition, 89,* B15-B25.

Xu, F. & Arriga, R. I. (2007). Number discrimination in 10-months-old infants. *British Journal of Developmental Psychology, 25,* 103-108.

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

Zorzi, M. & Butterworth, B. (1999). A computational model of number comparison. In M. Hahn & S. C. Stoness (Eds.), (pp. 772-777). Mahwah, NJ: LEA.

Zorzi, M., Priftis, K., Meneghello, F., Marenzi, R., & Umiltà, C. A. (2006). The spatial representation of numerical and non-numerical sequences: evidence from neglect. *Neuropsychologia, 44,* 1061-1067.

Zorzi, M., Priftis, K., & Umiltà, C. A. (2002). Neglect disrupts the mental number line. *Nature, 417,* 138-139.

Zorzi, M., Stoianov, I., Becker, C., Umilta, C., & Butterworth, B. (in press). The numerosity code for the representation of numerical magnitude.

Zorzi, M., Stoianov, I., & Umiltà, C. A. (2005). Computational modelling of numerical cognition. In J.I.Campbell (Ed.), *Handbook of mathematical cognition* (pp. 67-84). New York: Psychology Press.

Zur, O. & Gelman, R. (2004). Young children can add and subtract by predicting and checking. *Early Childhood Research Quarterly, 19,* 121-137.