UNIVERSITE CATHOLIQUE DE LOUVAIN

Faculté de Psychologie et des Sciences de l'éducation

Centre de Neurosciences Système et Cognition

The representations of numbers:
From the non-symbolic to the symbolic numerical systems

Virginie CROLLEN

Jury:

Prof. Pierre FEYEREISEN, UCL, President

Prof. Mauro PESENTI, UCL, Promoter

Prof. Wim FIAS, Gent Universiteit

Prof. Xavier SERON, UCL

Prof. Marie-Pascale NOËL, UCL

Prof. Jacques GRÉGOIRE, UCL

Thesis submitted for the degree of "Docteur en Sciences Psychologiques"

Louvain-la-Neuve, Belgique, Mai 2011

UNIVERSITE CATHOLIQUE DE LOUVAIN

Faculté de Psychologie et des Sciences de l'éducation

Centre de Neurosciences Système et Cognition

The representations of numbers:
From the non-symbolic to the symbolic numerical systems

Virginie CROLLEN

Jury:

Prof. Pierre FEYEREISEN, UCL, President

Prof. Mauro PESENTI, UCL, Promoter

Prof. Wim FIAS, Gent Universiteit

Prof. Xavier SERON, UCL

Prof. Marie-Pascale NOËL, UCL

Prof. Jacques GREGOIRE, UCL

Thesis submitted for the degree of "Docteur en Sciences Psychologiques"

Louvain-la-Neuve, Belgique, Mai 2011

Remerciements….

Je tiens à remercier tout particulièrement mon Promoteur, Xavier Seron, qui m'a intégrée dans le groupe calcul il y a déjà 5 ans et qui, même pensionné, s'est toujours intéressé à mon travail de recherche. A coup de « *c'est bien petite* », « *comment ça va ma grande* » et « *salut vilaine* », il n'a cessé de me soutenir et de me pousser dans les limites de mon raisonnement. Il a été pour moi un encadrant présent et respectueux de mes idées et de mes projets. Je lui en suis extrêmement reconnaissante ! J'ai aimé travailler avec lui mais j'ai également aimé déguster Duvel et spaghettis bolo en sa compagnie. Xavier… merci de t'être intéressé non seulement à mes activités recherche mais aussi à mes activités extra-recherche. Merci de m'avoir téléphoné de temps en temps pour me demander non pas comment allait ma thèse mais comment, moi, j'allais…

I wish to express my deep gratitude to Professors Wim Fias and Jacques Grégoire who accepted to be members of my jury. Merci à Pierre Feyereisen d'avoir accepté de présider mon jury de thèse. Merci à Mauro Pesenti d'avoir repris le rôle officiel de Xavier dans le courant de l'année 2010. Merci à Marie-Pascale Noël de m'accorder sa confiance et de soutenir ma candidature au poste de chargé de recherches… Notre collaboration est une perspective enthousiasmante !

Merci à toutes les personnes voyantes et non-voyantes qui ont accepté de participer à mes études…

Merci aux « anciens » et « nouveaux » chercheurs de la faculté de psychologie avec qui j'ai partagé d'agréables discussions professionnelles et personnelles : Alice, Auriane, Caro, Charlotte, Fred, Gaëlle, Gilles, Laurence, Mandy, Maria, Michael, Nicolas Ma., Nicolas Mi., Pierre, Renaud, Salva, Sandrine, Thomas, Valéry. Merci à Valérie et à Christine de m'avoir soutenue et encouragée sans relâche au cours de ces derniers mois. Toutes leurs petites attentions m'ont sans aucun doute aidée à clôturer ce travail de thèse dans les délais impartis. Un merci tout particulier à Julie Castronovo avec qui j'ai réalisé mes premières expériences, d'abord en tant que mémorante et ensuite en tant que thésarde. Sa contribution à la rédaction des articles a été très enrichissante. Je voudrais également remercier Xavier De Fays, Rachel Mahe et Stéphane Grade pour leur implication dans mes travaux de recherche.

Merci à Olivier Collignon de m'avoir accueillie au CERNEC, à l'Université de Montréal. Son intérêt pour la recherche m'a poussée à étendre et à clôturer mes expériences sur les non-voyants; son enthousiasme m'a donné la motivation nécessaire pour entamer la rédaction de cette thèse. Un grand merci également à Giulia Dormal pour son soutien, son dynamisme et sa générosité. Merci de m'avoir fait découvrir tous les endroits les plus excitants de Montréal et merci d'avoir rendu mon séjour aussi stimulant intellectuellement que culturellement.

Merci à toutes les personnes « de l'ombre » : Brigitte Champagne, Yolande Nivaille, Patricia Badoux et Vinciane Desnoeck, nos secrétaires; Roland Jaumain et Pierre Mahau, nos informaticiens; et enfin, merci à Nathalie Lefèvre pour tous ses conseils statistiques.

Spéciale dédicace à « *Ritalino* » avec qui j'ai partagé mes joies et mes peines chaque temps de midi depuis maintenant 4 ans… Merci de m'avoir écoutée avec beaucoup d'attention et de patience, merci de m'avoir fait découvrir différents univers artistiques, cinématographiques et musicaux et merci d'avoir été là chaque fois que j'en avais besoin.

Merci à mes deux grands amis, Valérie & Karim, pour leur présence et les fous rires partagés. Les soirées « chez Jules » et les soirées « Vodka Redbull » sont des souvenirs inoubliables… Merci aussi à la « bande des neuropsy » : Amélie, Caro et Marie. Les moments passés en leur compagnie ont toujours été très agréables.

Merci à toute l'équipe du restaurant Del Castello, et plus particulièrement merci à Gaby et à Giovanni, qui, depuis 8 ans maintenant, sont comme ma deuxième famille. Merci à tous les membres de l'EPO, en particulier : Dominique, Marc, Michel, Paul, Pascal & Renaud. Ils m'ont appris à vider mon esprit aussi bien que mon masque de plongée…

Enfin, last but not least, merci à ma famille : mes parents, ma grand-mère, mon frère, ma belle-sœur et ma filleule. S'il y a bien des gens qui m'ont toujours soutenue, ce sont eux ! Avoir supporté mes sautes d'humeur, mes envies de tout arrêter et puis de tout recommencer n'a pas dû être chose aisée… Thomas, merci pour ton soutien informatique, à moyenne et à grande distance… Mam's, merci pour ta présence bienveillante et merci de t'être intéressée à la neuropsy en général et aux représentations numériques en particluier. Maman, Papa, merci d'avoir cru en moi et merci de m'avoir offert ce que je n'ai pas (encore) trouvé dans le milieu professionnel: un environnement stable et sécurisant. Les dimanches soirs en famille m'ont souvent aidée à prendre des décisions importantes … C'est à vous et à Xavier que je dois la réalisation de cette thèse…

Table of contents

PART II

GENERAL OVERVIEW

The present thesis is divided in two different parts. The first one is concerned with the symbolic and non-symbolic representations of numbers. The second one is concerned with the realtionships between number processing and visuo-motor functions. Regarding the first question, we examined whether the numerosity estimation performance could reflect the qualitative features of symbolic and non-symbolic representations of numbers. Regarding the second question, we examined whether vision could influence the interactions between numbers and space and between numbers and fingers. Each part comprises some theoretical and empirical chapters. At the beginning of each part, an overview will describe the research background and provide a description of the structure of the chapters.

PART I

OVERVIEW

Human and non-human animals come to life equipped with an evolutionary numerical system which allows numerosity to be represented in an approximate and compressed fashion (Dehaene, 1992). This approximate number sense obeys Weber's law and is therefore ratio-dependent: the smaller the ratio separating two numerosities the easier are the numerosities to discriminate. While obedience to Weber's law is very pervasive in numerical cognition, the availability of number symbols varies across cultures (Pica, Lemer, Izard, & Dehaene, 2004) and arises at 2 or 3 years old in human development (Fuson, 1988). However, once acquired, the numerical symbols entail profound changes in the way numerical information is mentally represented and processed: while numerical symbols are assumed to be processed by a place-coded and linear representation, non-symbolic quantities are assumed to be processed by a summation-coded and logarithmic representation (Dehaene, 2009). The first part of the present thesis aims to examine whether these qualitative differences could be reflected in numerosity estimation performance and is organized as follows:

In Chapter 1, the main features of the symbolic and non-symbolic representations of numbers are exposed.

In Chapter 2, the numerosity estimation process is described and defined as a quantification process that requires mapping a numerical symbol onto a non-symbolic numerical representation.

In Chapter 3, we present a study that aimed to determine whether the mapping and the differential scaling of the symbolic and non-symbolic representations of numbers could be the origin of the under- and overestimation found in numerosity estimation tasks (the bi-directional mapping hypothesis; Castronovo & Seron, 2007a).

In Chapter 4, we examine whether the over-estimation error could be due to a left-to-right displacement on the mental number line (i.e., operational momentulm) rather than to a bi-directional mapping process.

In Chapter 5, the possibility to calibrate the numerosity estimation process is examined. The aim of this Chapter is to investigate whether the calibration effect could reflect the summation- and place-coding features of the non-symbolic and symbolic representations of numbers.

The results of our experiments are discussed and integrated in the last chapter of this first part (Chapter 6). Our empirical studies take the form of published or submitted scientific papers.

CHAPTER 1

The semantic representations of numbers

When children acquire numerical skills, they have to learn a variety of specific numerical tools. The most obvious are the numerical codes such as number words (one, two, three, etc.) or Arabic numerals (1, 2, 3, etc.). Other skills will be relatively more abstract: arithmetical facts (i.e., 4 x 2 = 8), arithmetical procedures (i.e., borrowing), or arithmetical laws (i.e., a + b = b + a). The acquisition of these numerical tools is complex and probably not facilitated by the fact that a numerical expression does not have a single meaning. Indeed, numbers can be used as a kind of label or proper name (i.e., Bus 51). They can also be part of a familiar fixed sequence (i.e., 51 comes immediately after 50 and before 52). They can be used to refer to continuous analogue quantities (i.e., 51,2 grams) (Butterworth, 2005; Fuson, 1988) and, most importantly, they can be used to denote the number of things in a set – the *cardinality* of the set. This last function means that numerosity is abstract and not a physical object or the property of an object (such as a colour or a shape). Rather, it seems that numerosity is a property of a set that can have any type of member. In this chapter, we will demonstrate that children are able to understand the special meaning of cardinality because they possess a specific capacity for dealing with quantities. In the first part of the chapter, number processing in preverbal infants and in animals will be examined and compared. Then, in the second part of this chapter, and in order to account for the similarities found between animal and human number processing, the accumulator model and the mental number line model will be explained in detail. In the third and last part of the chapter, the differences between symbolic and non-symbolic representations of numbers will be considered.

1. Numerosity as the basis of numerical abilities

1.1 Number processing in infants

There is compelling evidence that young infants possess a biologically determined “*sense of number*” (Dehaene, 1992). A large set of behavioural studies using the classic method of habituation has indeed revealed sensitivity to small numerosities (e.g., Starkey & Cooper, 1980). In the preliminary study of Starkey and Cooper (1980), for example, slides with a fixed number of 2 dots were repeatedly presented to 4- to 6-month-old infants until their looking time decreased, indicating habituation. At that point, a slide with a deviant number of 3 dots was presented and yielded significantly longer looking times, indicating dishabituation and therefore discrimination between the numerosities two and three. This effect was replicated with newborns (Antell & Keating, 1983) and with various stimuli such as sets of realistic objects (Strauss & Curtis, 1981), targets in motion (Van Loosbroek & Smitsman, 1990; Wynn, Bloom & Chiang, 2002), two- and three-syllable words (Bijeljac-Babic, Bertoncini, & Mehler, 1991) or puppet making two or three sequential jumps (Wynn, 1996; Wood & Spelke, 2005). The recovery effect was even replicated in cross-modal numerosity matching tasks (Féron, Gentaz & Streri, 2006; Izard, Sann, Spelke, & Streri, 2009; Jordan & Brannon, 2006; Kobayashi, Hiraki & Hasegawa, 2005; Starkey, Spelke & Gelman, 1983, 1990) but was not observed with other numerosities such as 4 and 6 (Starkey & Cooper, 1980). Taken together, these data therefore suggest that infants may possess a concept of small numbers which is independent of stimulus modality (auditory or visual, homogeneous or heterogeneous) and stimulus presentation (simultaneous or sequential) but which is dependent on the absolute number of items presented (up to 3 or 4).

Besides their ability to discriminate particularly small numerosities, infants are also able to discriminate large numerosity sets. However, this discrimination ability differs dramatically from that observed with small numerosities: performance no longers depends on the absolute number of items presented but on the numerical ratio that separates the two numerosities to be discriminated. Hence, while 6-month-old infants are able to discriminate numerositites with a 1:2 ratio (4 vs. 8, 8 vs. 16 and 16 vs 32 stimuli: Brannon, Abbott & Lutz, 2004; Lipton & Spelke, 2003; 2004; Xu, 2003 ; Xu & Spelke, 2000; Xu, Spelke & Goddard, 2005; Wood & Spelke, 2005), they fail to discriminate numerosities with a 2:3 ratio (8 vs. 12 and 16 vs. 24: Xu & Spelke, 2000). By 9 or 10 months, however, the precision of the representation improves since infants become able to discriminate between 8 and 12 elements (Lipton & Spelke, 2003; Xu & Arriaga, 2007). Interestingly, this pattern of success and failure was observed across different kind of stimuli: visual arrays (Xu, 2003; Xu & Spelke, 2000), sequences of sounds (Lipton & Spelke, 2003, 2004) and sequences of a puppet's jumps (Wood & Spelke, 2005) and thus supports the idea that infants' numerical discrimination actually depends on abstract representations of numerosity.

Following these two waves of investigations on numerosity discrimination in infancy, it has been suggested that basic numerical abilities could be sustained by two different systems (e.g. Butterworth, 1999a; Carey, 2001; Dehaene, 1992; Feigenson, Dehaene, & Spelke, 2004; Xu, 2003; Xu & Spelke, 2000): **(1)** an object-tracking system (e.g., Scholl, 2001; Trick & Pylyshyn, 1994; Uller, Huntley-Fenner, Carey, & Klatt, 1999) which has a clear limit on set size (3 or 4) and thus allows the precise representation of small number of objects; **(2)** a large numerosity estimation system which has no set size limit and therefore allows the approximate representation of large numerosity sets. These non-verbal quantification systems allow infants to form expectations about the outcomes of simple arithmetic problems such as 1 + 1 = 2 and 2 – 1 = 1 (Wynn, 1992a) or 5 + 5 = 10 and 10 − 5 = 5 (McCrink & Wynn, 2004) and are assumed to constitute the foundation of all further, more elaborate symbolic numerical skills that evolved in numerate humans (Carey, 2001; Dehaene, 2001).

1.2 Number processing in animals

The two biological precursor systems of elementary arithmetic not only exist in human infants but also in many animal species. Rhesus monkeys, for example, systematically choose the larger quantity of 1 vs. 2, 2 vs. 3, or 3 vs. 4 apple slices but are at chance with 3 vs. 8 and 4 vs. 8 (Hauser, Carey, & Hauser, 2000). Untrained monkeys tested in the wild were also shown to be able to detect violation of addition and subtraction events (Hauser, McNeilage, & Ware, 1996). Interestingly, an upper size limit of 4 items was found in this violation of expectation paradigm: monkeys succeeded in discriminating the correct outcomes of 1 + 1, and 2 + 1, but failed with 2 + 1 + 1 (Hauser & Carey, 2003). Moreover, cross-modal extraction of numerosity has been observed in rats (Church & Meck, 1984). Indeed, a rat trained to press a lever on the left in response to two flashes or two sounds, and a lever on the right in response to four flashes or four sounds, spontaneously pressed the right lever when presented with a combination of two sounds and two lights. This observation therefore suggests that the rat’s behaviour was based on the abstract total number of four events, not on modality-specific representations. In all these experiments, non-human animals were shown to be able to exactly represent a maximum of 4 elements. This abrupt limit therefore implicates the object tracking system as the source of their performance.

The object tracking system is nevertheless not the only quantification system that animals possess. In Hauser, Tsao, Garcia, and Spelke’s experiment (2003), for example, cotton-top tamarin monkeys were shown to be able to discriminate sequences of 4 vs. 8, 4 vs. 6, and 8 vs. 12 speech syllables while they systematically failed to discriminate 4 vs. 5 and 8 vs. 10 syllables. This particular pattern of performance shows the ratio signature and therefore indicates the presence of the large numerosity estimation system in animals.

The data presented in this section are far from exhaustive but intended to show that humans and non-human animals present two distinct elementary systems for representing numbers. One of these is precise and limited by its absolute set size (up to 3 or 4), while the other is extensible to very large quantities and thus allows the discrimination and approximate representation of visual and auditory numerosities without verbal counting (Butterworth, 1999a; Dehaene, 1992; Feigenson et al., 2004; Nieder & Miller, 2003; 2004; Wynn, 1995). As we will see in the following sections of the present chapter, only numerate humans acquire additional symbolic numerical systems such as Arabic digits and number words. The developmental relationship linking the two non-symbolic systems to symbolic arithmetic is however still under debate. According to the dominant view, the approximate number system is the evolutionary and ontogenetic foundation of developed numerical concepts (Dehaene, 2001; Piazza, 2010). According to a less widespread theory, the object tracking system is an index of the (exact) symbolic representation of numbers and forms the core numerical ability upon which all others are built (Butterworth, 1999a, 2005; Carey, 2001). In the following sections of the present chapter, we will describe in more detail the parallels between animal and human number processing and the main features of the symbolic and non-symbolic representations of numbers.

1.3 Parallels between animal and human number processing

Support for the hypothesis of a shared evolutionary heritage for elementary arithmetic comes from the finding of two major parallels between human and animal approximate number processing: the numerical distance effect and the number size effect, both arising from the ratio effect.

On the one hand, the ***numerical distance effect*** (see Fig. 1-1) refers to the finding that the ability to discriminate two numerosities improves as the numerical distance between them increases. Distance effects have been reported with various animal species whenever the animal must identify the larger of two numerical quantities or indicate whether two numerical quantities are the same or not (Gallistel & Gelman, 1992). In the same way, the distance effect has been obtained with adult humans, in tasks requiring them to compare not only the numerosity of two sets of dots (Buckley & Gillman, 1974; van Oeffelen & Vos, 1982) but also when processing Arabic digits, number words (Buckley & Gillman, 1974; Dehaene, 1996; Dehaene, Dupoux & Mehler, 1990; Moyer & Landauer, 1967) and two-digit Arabic numerals (Dehaene et al., 1990). In all cases, comparison times and error rates are a continuous, convex upward function of distance.

The ***number size effect*** (see Fig. 1-2) refers to the finding that, for equal numerical distance, discrimination of two numerosities worsens as their numerical size increases. This effect holds when various animal species are presented with different numbers of visual objects or sounds in various discrimination or comparison tasks (Gallistel & Gelman, 1992). In human adults, a growing imprecision for increasingly large numerosities was found during the identification or discrimination of two sets of dots (Buckley & Gillman, 1974; van Oeffelen & Vos, 1982). Interestingly, a similar number size effect was found when humans compare or calculate with numbers presented as Arabic digits or as number words (Buckley & Gilman, 1974; Moyer & Landauer, 1967; Aschcraft & Battaglia, 1978). Thus, cognitive psychological evidence indicates that, in various number processing tasks, humans and non-human animals quickly access a representation of numerical quantities which is organized by numerical proximity and gets increasingly fuzzier for increasingly larger numbers.

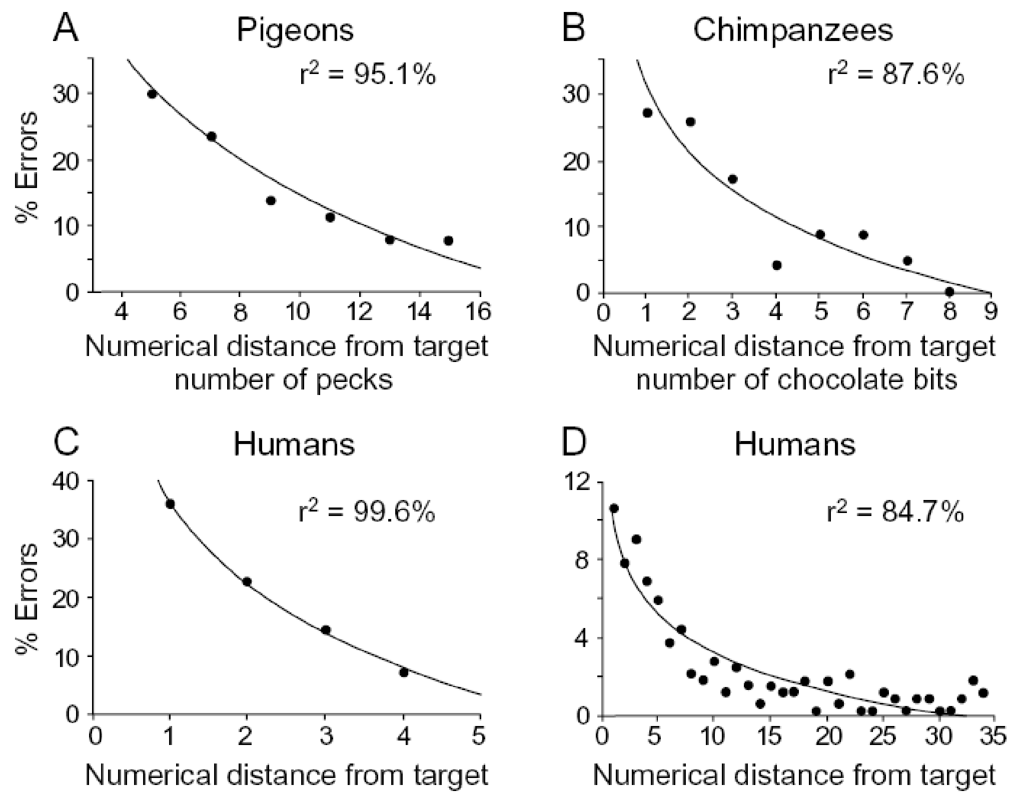

Figure 1-1. Distance effect. (A) Distance effect observed with pigeons in a task requiring comparison of the number of pecks with a fixed standard of 50 (Rilling & McDiarinid, 1965). **(B)** Distance effect observed with chimpanzees in a task requiring selection of the larger of two small numbers of chocolate bits (Washburn & Rumbaugh, 1991). **(C)** Distance effect observed with humans in a task requiring decision as to whether a pattern contained exactly 12 dots (van Oeffelen & Vos, 1982). **(D)** Distance effect observed with humans in a task requiring decision as to whether a two-digit Arabic numeral is larger than 65 (Dehaene et al., 1990). *Reproduced from Dehaene, Dehaene-Lambertz & Cohen (1998).*

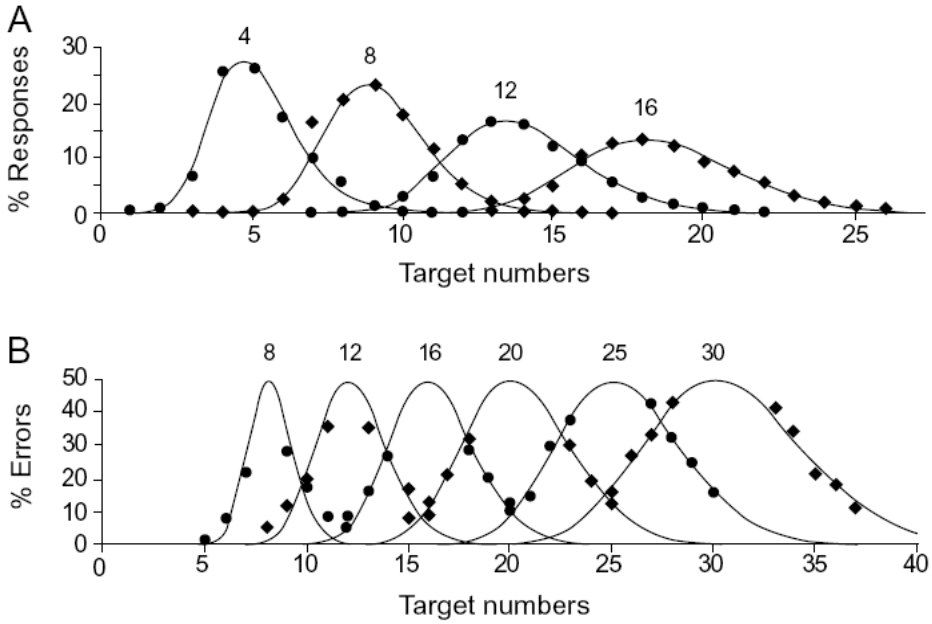

Fig. 1-2. ***Size effect.*** **(A)** Size effect observed with rats in a task requiring comparison of the numbers of lever presses to a fixed standard of 4, 8, 12 or 16 (Mechner, 1958). **(B)** Size effect observed with humans when they have to compare the numerosity of a dot pattern to a fixed standard of 8, 12, 16, 20, 25 or 30 (van Oeffelen & Vos, 1982). In both cases, the size effect is revealed by an increasing dispersion of the distribution of responses. In general, when humans process symbolic numerals, the size effect is reflected only by lengthened response times and a small increase in error rates for increasingly larger numbers (Aschcraft & Battaglia, 1978; Buckley & Gillman, 1974; Moyer & Landauer, 1967). *Adapted from Dehaene et al. (1998).*

Both the distance and the size effects reflect the fact that numerical processing obeys Weber's law in the same way as the perception of physical stimuli such as loudness or brightness does in psychophysics (Dehaene, 2001). Weber's law states that the change in stimulus intensity needed for an organism to detect a change is a constant proportion of the original stimulus intensity rather than a constant amount. The same logic holds true for non-verbal number discrimination: the ability to discriminate between two numerosities is influenced by both the linear distance and the absolute magnitude of the values compared. Therefore, the larger the stimuli and the smaller the difference between them, the worse their processing. An alternative formulation is that numerosity discrimination (and hence, the distance and size effects) depends on the ratio of the numbers involved, not their absolute values. The closer this ratio is to 1, the more difficult the number discrimination. So, it will be easier to compare 1 and 2 than 8 and 9 because the ratio separating the first two numbers is 0.50 vs. 0.88 for the last two numbers.

2. Theoretical models

In this section, we will consider the two major theoretical models that have been proposed to conceptualize how numerical processing obeys Weber's law: the accumulator model and the mental number line model. We will also look in some detail at the neuronal model and the cerebral networks which are assumed to sustain the implementation of the mental number line.

2.1 The accumulator model

The accumulator model was initially proposed by Meck and Church (1983) in order to account for numerical competences observed in rats and was then extended by Gallistel and Gelman (1992, 2000) in order to integrate human and animal numerical performances. According to this model, numerosities are represented by magnitudes on an analogue linear representation with scalar variability. The mental magnitudes are generated by a non-verbal counting process so that each enumerated unit is represented in a mental accumulator by an additional fixed increment of magnitude. The cardinal value of the counted set or sequence is represented by the final magnitude, which is then either compared to a magnitude stored in memory or mapped to symbols for quantities (see Fig. 1-3a). As the magnitudes in memory are noisy, the trial-to-trial variability in remembered magnitudes is proportional to these magnitudes (i.e., scalar variability). Discrimination failures therefore occur when the signal distributions for nearby numerositites overlap. This model is a summation-coding system as the mental representation of a numerosity is analogue to the magnitude it represents (i.e., a given number activates a portion of the numerical representation and this portion constitutes a subset of the portion activated for a larger number) (see Fig. 1-3b).

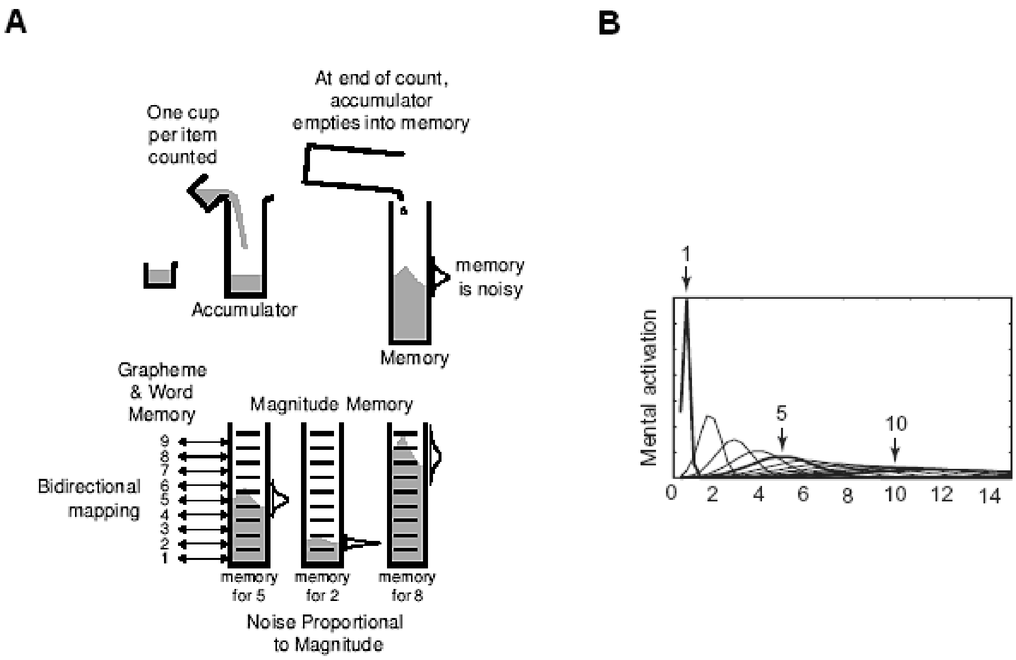

Fig. 1-3. (A) The accumulator model. A numerosity is represented through the accumulation of fixed increments of magnitude. Accumulated magnitudes may then be compared to a magnitude stored in memory or mapped to symbols for quantities. However, magnitude read from memory is noisy. So, the greater the magnitude, the more likely an error. *Reproduced from Cordes, Gelman, Gallistel, & Whalen (2001).* **(B) Linear model with scalar variability.** *Reproduced from Feigenson et al. (2004).*

The accumulator model has been supported by Whalen, Gallistel and Gelman’s (1999) study, in which human adults undertook two numerosity estimation tasks. In the first, participants had to approximately produce the magnitude of a target number by pressing a response key at a rate that made vocal or subvocal counting impossible. In the second, participants had to estimate the number of flashes in a rapid and randomly timed sequence. In both tasks, the mean responses increased in proportion to the target number, as did the trial-to trial variability, thus suggesting that adults, like animals, use a non-verbal counting mechanism that generates magnitude representations with scalar variability (see Cordes et al., 2001 for similar results).

2.2 The mental number line

2.2.1 Psychophysical features

The second model proposed to account for the recurrent observation of the distance and size effects (Aiken & Williams, 1968; Buckley & Gillman, 1974; Dehaene, 1996; Dehaene et al., 1990) is the mental number line model (e.g., Restle, 1970; Dehaene, 1992).

According to this model, the representation of numerosities would present two major features, each explaining one of the behavioural effects (i.e., the distance and size effects). Firstly, it seems that numbers are represented on the mental number line by equal distributions of activation (see Fig. 1-4). Therefore, the larger the distance between the quantities being compared, the more distant their distributions of activation on the mental number line and the easier it is to discriminate between them (i.e., distance effect; Dehaene, 1992, 1997, 2001, 2003). Secondly, the mental number line appears to be logarithmically scaled (see Fig. 1-4), so that small numerical magnitudes are further apart on the number line than large numerical magnitudes. This leads to less accurate representation and poorer discrimination of large than of small numbers and thus explains the behavioural size effect. A third feature of the mental number line, namely its orientation from left-to-right, will be explored in Chapter 3.

The imprecision of the mental number line decreases with age (Halberda & Feigenson, 2008; Piazza, Facoetti, Trussardi, Berteletti, Conte, Lucangeli et al., 2010; Xu & Spelke, 2000), is subject to individual differences (Halberda, Mazzocco, & Feigenson, 2008) and is expressed as a Weber fraction (*w*). Whereas infants can detect changes in numerosity for ratios of 1:2 at 6 months (Lipton & Spelke, 2003; Xu & Spelke, 2000) and 2:3 at 9 months (Xu and Arriaga, 2007), the Weber fraction for adults is approximately 0.11, yielding successful numerical discrimination of arrays differing by a ratio as little as 9:10 (Halberda & Feigenson, 2008).

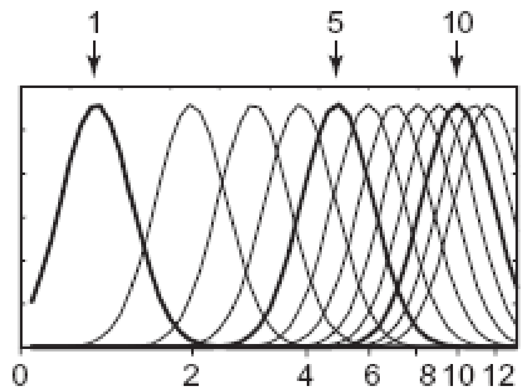

Fig. 1-4. Representation of the log-Gaussian model of the mental number line. *Reproduced from Feigenson et al. (2004).*

At a behavioural level, scalar variability (i.e., the accumulator model) and compressed coding (i.e., the mental number line model) lead essentially to the same predictions (Dehaene, 2003): the amount of signal overlap between representations of any two numbers is a function of the size and distance between them. However, as we will see in the following section, the discovery of single numerosity detector neurons (Nieder, Freedman, & Miller, 2002; Nieder & Miller, 2003) provided further support for the mental number line model.

2.2.2 Neuronal model

The discovery of single neurons tuned to numerosity in the macaque monkey (Nieder et al., 2002; Nieder & Miller, 2003, 2004), and of neurons plausibly implementing random-walk accumulation models of decision (Gold & Shadlen, 2002) have provided insight into the neuronal mechanisms from which the distance and size effects arise. More particularly, Dehaene and Changeux (1993) proposed that numerosity is extracted from a retinotopic map in three successive stages (see Fig. 1-5): **(1)** *object location map*: object locations are coded in a retinotopic way, regardless of object identity and size - a normalization which could be performed by the occipito-parietal "where" pathway (Izard, Dehaene-Lambertz, & Dehaene, 2008); **(2)** *accumulation neurons*: the total activity is represented by accumulation neurons which simply sum the activation on the object location map. These neurons have limited receptive fields and thus respond solely to the local numerosity within a certain retinotopic area, not to the total numerosity across the whole visual field (Roitman, Brannon, & Platt, 2007); **(3)** *numerosity detector neurons*: the summed activation of the accumulation neurons stimulated a numerosity-specific detector neuron which is "number selective" (i.e., tuned to a specific numerosity) instead of being "number sensitive" as the object location map (i.e., there is more neural activity when more objects are presented).

According to the neural model, numbers are represented cortically by a population of neurons, each coarsely tuned to a preferred quantity. The coarse tuning implies that the closer two numbers are, the more similar the coding scheme of their neuronal populations. It thus explains why numerical distance is a major determiner of number comparison performance. Furthermore, the model assumes that the preferred numbers for different neurons are distributed on a logarithmic scale, thus permitting a wide range of quantities to be encoded with a small population of cells. This compressive property implies an increasingly coarser encoding of larger numbers. Thus, it can explain the behavioural observation that, as the numbers increase, it takes a proportionally larger difference between them for discrimination performance to remain at a constant level (Weber's law).

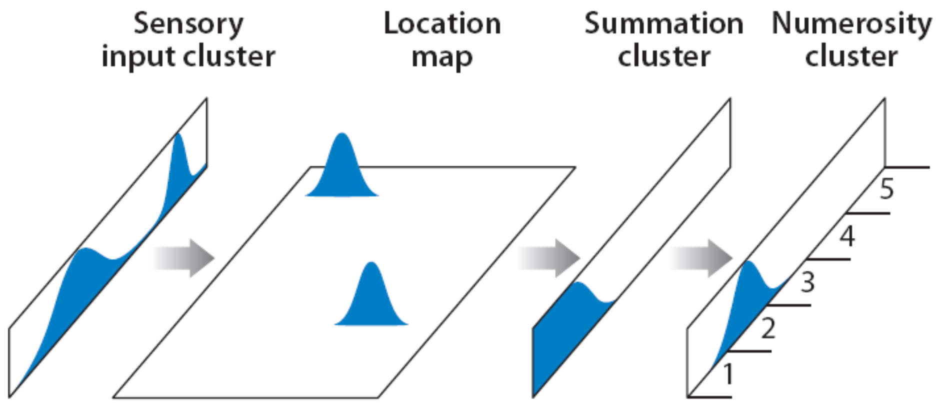

Figure 1-5. Neuronal model. Sensory input cluster is mapped onto the location map, where each stimulus is represented as a similar level of activity regardless of the size or the shape of the objects. The location map then projects to the summation or accumulation cluster which simply sums the activation of the location map. The summed activation then stimulated a specific numerosity detector neuron. *Reproduced from Nieder & Dehaene (2009).*

2.2.3 Cerebral networks

Neuro-imaging studies rapidly identified the horizontal segment of the intra parietal sulcus (hIPS) as the neuroanatomical correlates of numerical processing. This region is actually always activated whenever a numerical task requires a quantitative representation of numbers. For example, it is more active when subjects calculate than when they merely have to read numerical symbols (Chochon, Cohen, van de Moortele, & Dehaene, 1999; Pesenti, Thioux, Seron, & De Volder, 2000a). Parametric studies also revealed that the activation of the hIPS is modulated by semantic parameters such as the absolute magnitude of the numbers and their value relative to a reference point. Thus, intraparietal activity is larger and lasts longer during operations with large numbers than with small ones (Kiefer & Dehaene, 1997; Stanescu-Cosson, Pinel, van de Moortele, Le Bihan, Cohen, & Delaene, 2000). It is also modulated by the numerical distance separating the numbers in a comparison task (Dehaene, 1996; Pinel, Dehaene, Rivière, & Le Bihan, 2001), and by the degree to which novel value deviates from habituation value (Piazza, Izard, Pinel, Le Bihan & Dehaene, 2004).

Finally, the activation of the hIPS is independent of the particular modality of input used to convey the numbers. Arabic numerals, spelled-out number words, and even non-symbolic stimuli like sets of dots or tones can activate this region (Le Clec'H, Dehaene, Cohen, Mehler, Dupoux, Poline et al., 2000; Piazza, Mechelli, Price, & Butterworth, 2006; Pinel et al., 2001). Taken together, these data suggest that the hIPS is essential for the semantic representation of numbers. This representation may therefore provide a foundation for our "numerical intuition" (Dehaene, 1992, 1997; Dehaene & Marques, 2002).

3. Symbolic and non-symbolic representations of numbers

3.1 The triple code model

The metaphor of the mental number line takes place in what Dehaene (1992, 1997; see also Dehaene & Cohen, 1995, 1997) called the triple code model. This model assumes that numbers are represented in three different codes that are related to specific tasks and have distinct neuroanatomical correlates (Dehaene, Piazza, Pinel & Cohen, 2003). First, there is the analogue magnitude code which corresponds to the mental number line and which is, according to the model, the only code that includes semantic knowledge about numbers. This code is therefore used in magnitude comparison and approximation tasks and is supposed to be implemented in the hIPS. Second, there is a verbal word frame which is activated whenever sequences of number words are manipulated. It would therefore be used for retrieving well learned arithmetic facts such as multiplication tables and is predicted to engage the left angular gyrus. Third, there is the visual Arabic number form which represents numbers as strings of Arabic digits and which is used for multi-digit calculation and parity judgments (Dehaene & Cohen, 1991). The neuro-anatomical correlate of this last code is located in the bilateral inferior ventral occipito-temporal regions (Dehaene, 1992) (see Fig. 1-6 for an illustration of the brain regions involved in numerical processing).

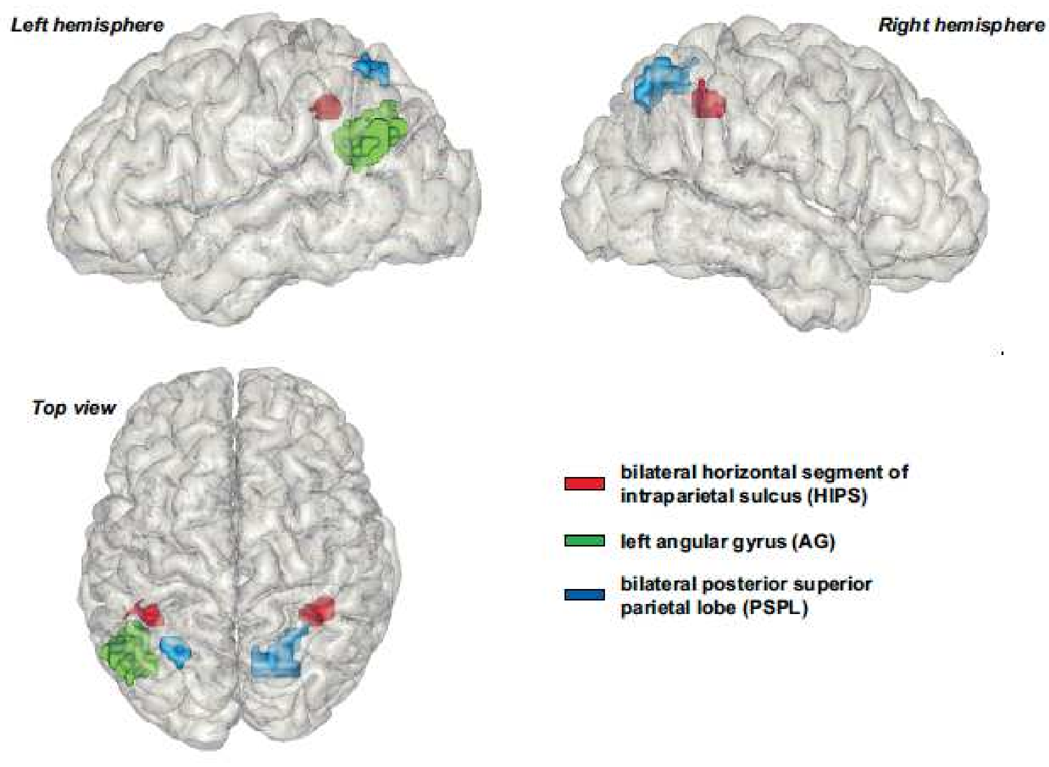

Fig. 1-6. Brain regions involved in numerical processing. According to Dehaene et al. (2003), the analogue magnitude code is located in the hIPS whereas the verbal word form engages the left angular gyrus (AG). The third region, the bilateral posterior superior parietal lobe (PSPL), is not specific to the number domain but plays a central role in visuo-spatial processing. *Reproduced from Dehaene et al. (2003).*

According to the triple code model, various transcoding paths enable numerical information to be translated from one code to the other. Overall, two major routes were proposed by Dehaene and Cohen (1995, 1997): **(1)** the direct asemantic route that translates written numerals to auditory verbal numbers without semantic mediation; and **(2)** the indirect semantic route that transcodes written to auditory numerals through the analogue magnitude code and which is therefore specialized for the manipulation of numbers. As adults were found to be able to compare numerosities across stimulus modality (i.e., verbal vs. visual) and format (i.e., simultaneous vs. sequential), the properties of the modality-specific stimulus were assumed to undergo a transformation into a quantity representation that is independent of the modality of the stimulus (Barth, Kanwisher, & Spelke, 2003).

To conclude, the idea that numbers may be represented in a non-symbolic format (i.e., the analogue code) and in two symbolic formats (i.e., Arabic digits and verbal numerals) already existed in the 90s. As already stated, the non-symbolic representation of numbers precedes the acquisition of numerical symbols and is present in adults of all cultures, as well as in infants and animals. In the course of the numerical development, however, this quantity representation is progressively connected to different numerical symbols. In other words, the arbitrary properties of a specific symbol will progressively be attached to a relevant quantity representation. According to the initial version of the triple code model (Dehaene & Cohen, 1995, 1997), the connection between symbolic and non-symbolic processing does not change the way people represent and process numbers. However, in the following sections, we will provide evidence that symbolic and non-symbolic representations may differ in terms of precision (section 3.2), neural coding (section 3.3) and hemispheric asymmetry (section 3.4).

3.2 Approximate vs. accurate?

Several pieces of evidence suggest that the acquisition of numerical symbols provides access to a more accurate representation of numbers. The first data supporting this idea probably date back several years when Buckley and Gilman (1974) demonstrated that the distance and size effects were smaller in tasks involving symbolic numbers than in tasks involving non-symbolic quantities. It was therefore assumed that the representation of numerical symbols was less approximate (less compressed) than the representation of non-symbolic quantities. Moreover, in a more recent study, no significant correlations were found between the symbolic and the non-symbolic distance effects (calculated on the basis of reaction times), this lack of correlation suggesting that symbolic and non-symbolic numerical quantities may be processed by different underlying representations (Holloway & Ansari, 2009).

While both symbolic (De Smedt, Verschaffel & Ghesquière, 2009; Holloway & Ansari, 2009) and non-symbolic (Gilmore, McCarthy & Spelke, 2010; Halberda et al., 2008) arithmetic abilities were shown to be in a close relationship with mathematical achievement, only the lack of symbolic numerical representations (in animals and in some indigenous human cultures) appeared to preclude the development of higher mathematical skills (Dehaene, Izard, Spelke, & Pica, 2008; Gallistel & Gelman, 2000; Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004). For instance, whereas speakers of Mundurukú, an Amazonian language with a very small lexicon of number words, were able to add, subtract and compare approximate representations of large numbers far beyond their naming range, they were unable to solve exact arithmetic problems comprising numbers larger than 4 or 5. These data therefore suggest that the acquisition of numerical symbols plays a special role in the emergence of accurate numerical competence during child development.

A recent study further supported this idea by showing that the neural representation of symbolic numerical magnitudes was more finely tuned than that of non-symbolic numerical magnitudes (Piazza, Pinel, Le Bihan, & Dehaene, 2007). In this study, brain activity was measured with fMRI while adults participants passively observed non-symbolic (i.e., dot patterns) or symbolic (i.e., Arabic digits) quantities. Once the cerebral activity showed habituation to the repetitive presentation of a specific quantity, a deviant magnitude was presented, either as a set of dots or as an Arabic numeral. The presentation of the deviant stimulus was shown to provoke an asymmetric rebound effect in the left hIPS. Indeed, a distance-dependent recovery effect was observed after a distant Arabic digit was presented amongst dots, but an abnormal distance-independent recovery effect was observed after a set of dots was presented among Arabic digits. Interestingly, these data were interpreted as reflecting the accuracy of the symbolic and non-symbolic tuning curves. When habituation stimuli were presented as sets of dots, the adaptation effect touched a large population of neurons, given the neuron's broad tuning curve. This adaptation effect therefore transferred to close symbolic numerals presented as deviants, as the tuning curve for this numerical format was sharper than the one of the non-symbolic format. Inversely, when habituation stimuli were presented as Arabic digits, the adaptation effect touched a small population of neurons, given the neuron's sharp tuning curve. So, the adaptation to digits did not lead to the adaptation of the population neurons for dots; and a distance-independent recovery effect was thus observed.

A developmental transition from logarithmic to linear numerical representation has also been documented in several studies suggesting that children's representation of numbers changes over time with increasing experience with number symbols (Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Booth & Siegler, 2006;

Siegler & Booth, 2004; Siegler & Opfer, 2003). In these studies, children and adults had to place different numbers on a visual number line with 0 at one end and either 10, 20 (Berteletti et al., 2010), 100 (Siegler & Booth, 2004) or 1000 (Booth & Siegler, 2006; Siegler & Opfer, 2003) at the other end. All these studies reported consistent results: the youngest children positioned numbers logarithmically when the experimental condition was unfamiliar (e.g., 1000) but positioned them linearly when the experimental condition was familiar (e.g., 100). In contrast, older children and adults positioned numbers linearly in both experimental conditions. The dissociation between familiar and unfamiliar conditions therefore reveals that children's performance changes with age and shifts from a logarithmic to a linear representation. Interestingly, the linearity of the numerical representation appeared to be correlated to mathematical achievement (Booth & Siegler, 2008; Siegler & Booth, 2004).

Taken together, the data presented in this section suggest that there is a differential precision of number coding depending on the stimulus format, the symbolic numerical representation being accurate/linear and the non-symbolic representation being approximate/logarithmic.

3.3 Summation-coding vs. place-coding?

Non-symbolic and symbolic representations of numbers differ not only in terms of level of precision but also in terms of neural coding. Verguts and Fias (2004) were the first to propose this distinction by showing, with a simulation study, that a shared representation of numbers was reached through different neural coding steps depending on the stimulus format: **(1)** a summation-coding step for non-symbolic quantities and **(2)** a place-coding step for symbolic numbers. Summation-coding refers to the idea that a non-symbolic quantity activates a portion of the mental number line that includes all the units smaller and equal to the target quantity. Place-coding, on the other hand, refers to the fact that a symbolic number activates its corresponding unit on the mental number line with a maximal strength, but also activates smaller and larger neighbouring units with gradually decreasing strength as a function of distance (see Fig. 1-7). Whereas Verguts and Fias (2004) suggested that a logarithmic scale may characterize the summation-coded representation, Verguts, Fias and Stevens (2005) demonstrated that a linear scale may characterize the place-coded representation.

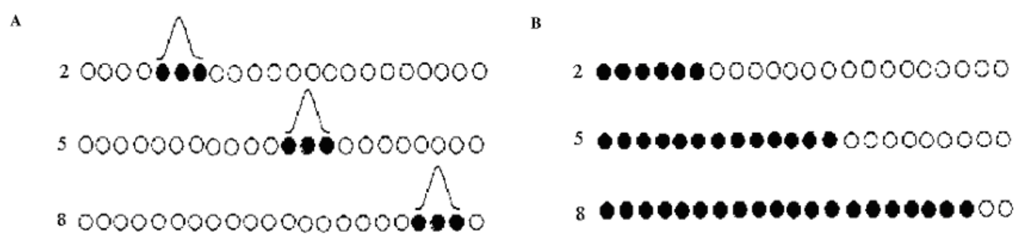

Fig. 1-7. Schematic representation of **(A)** place-coding and **(B)** summation coding. *Reproduced from Roggeman, Verguts and Fias (2007).*

A recent fMRI study by Santens, Roggeman, Fias, and Verguts (2010) supported Verguts and Fias's model (2004). This experiment consisted of the sequential presentation of dot patterns (10 blocks) or the sequential presentation of Arabic digits (10 blocks). At the end of each block, two stimuli were presented left and right of a fixation cross and participants were instructed to indicate the number/numerosity corresponding to the previous display. Brain activity was registered during the stimuli presentation and a functional connectivity analysis showed shared activation for symbolic and non-symbolic stimuli in the intraparietal sulcus but also demonstrated that the pathway to reach this area was modulated by the stimulus format: whereas non-symbolic quantities activated an area in the posterior superior parietal cortex, the processing of symbolic quantities activated a pathway that does not rely on this intermediate stage. The IPS was therefore identified as the neural substrate of the number-selective representation and the posterior superior parietal cortex as the neural substrate for the summation step required for the processing of non-symbolic quantities (see Fig. 1-8).

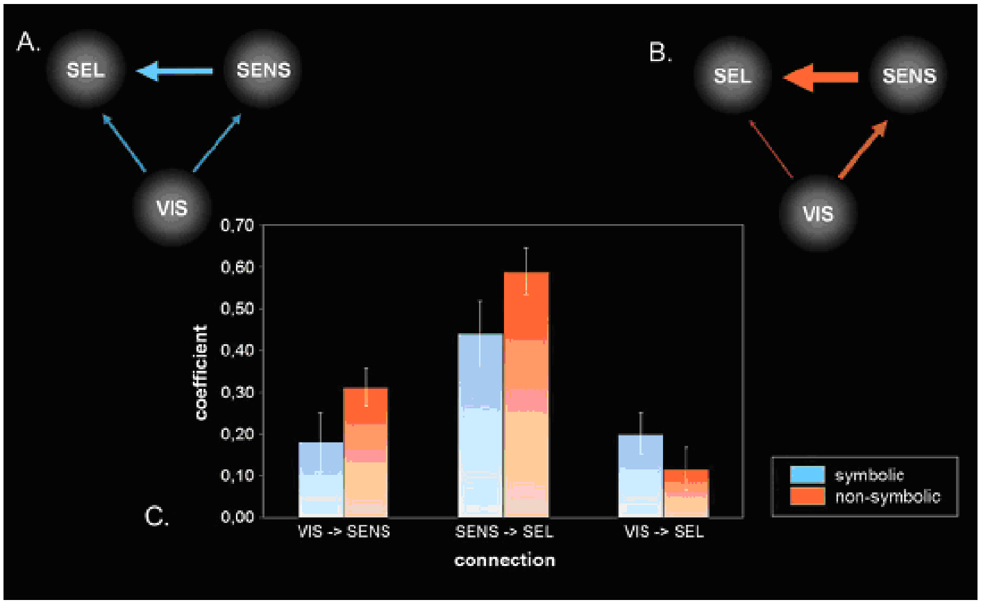

Fig. 1-8. VIS: primary visual region. **SENS**: number-sensitive region in the superior parietal lobe. **SEL**: number-selective region in IPS. Schematic illustration of the connectivity analysis for **(A)** symbolic and **(B)** non-symbolic stimuli. The weight of the lines reflects the strength of the connectivity. **(C)** Path coefficients resulting from the connectivity analysis for different connections and different formats. *Adapted from Santens et al. (2010).*

Moreover, evidence in favour of the distinction between summation and place-coding comes from a behavioural priming task of Roggeman et al. (2007). In this study, participants were asked to name a target quantity (from 1 to 5) after a prime had been presented. In the first condition, prime and target were presented as Arabic digits, whereas in the second they were presented as dot patterns. The authors demonstrated that the priming effect crucially depends on the experimental condition. In the digit format condition, the priming effect was *V*-shaped (see also Reynvoet & Brysbaert, 1999, 2004; Reynvoet, Brysbaert, & Fias, 2002), indicating that an Arabic digit primed smaller and larger numbers as a function of distance (see Fig. 1-9a). In the dot patterns condition, in contrast, the priming pattern was stepwise, suggesting that a dot pattern primed all the quantities smaller or equal to the prime value (see Fig. 1-9b).

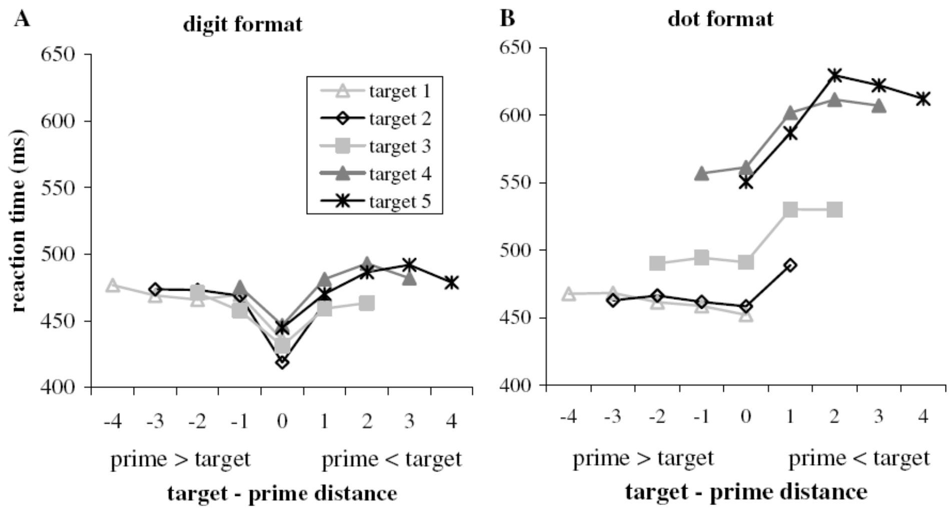

Fig. 1-9. (A) *V*-shaped and **(B)** stepwise priming effects for Arabic digits and dot patterns respectively. *Reproduced from Roggeman et al. (2007).*

These observations were thus in accordance with the model of Verguts and Fias (2004). In the digit condition, the prime induces a distance-dependent activation of the neighbouring units around the prime value (as expected with a place-coding representation). Therefore, the processing of targets close to the prime value was facilitated in comparison to the processing of targets far from the prime value. In the dot format condition, on the other hand, the prime activated and thereby facilitated the processing of all the targets smaller than or equal to the prime value (as expected with a summation-coding representation). As we will see in Chapters 5 and 6, the same pattern of results was observed with canonical and non-canonical finger configurations (Di Luca, Lefèvre & Pesenti, 2010).

Taken together, these data can be used to refine the neuronal model (Dehaene & Changeux, 1993) exposed in section 2.2.2. This model, in its original version, already assumed that numerical magnitudes, regardless of their input format, were represented by a number-selective coding system (i.e., the numerosity detector neurons). The data reported in the present section situated this representation in the IPS and demonstrated that symbolic and non-symbolic quantities reached this area through different pathways. Whereas non-symbolic inputs access the number-selective representation through a summation coding system located in the posterior superior parietal cortex, the processing of symbolic inputs does not require this intermediate step. The summation coding system corresponds to the accumulation neurons described in the neuronal model and therefore summates, in a logarithmic way, the number of elements displayed in a set. The linear scaling and the place-coding features of the symbolic numerical representation may stem from the fact that primary visual regions contact IPS directly or via a visual number form (as suggested in the triple code model: Dehaene & Cohen, 1991; see section 3.1).

According to Verguts and Fias (2004), however, only the left parietal representation acquires a refined precision following the acquisition of numerical symbols, the right parietal cortex keeping a coarse representation for both symbolic and non-symbolic notations. In the following section, we will provide converging evidence for this idea.

3.4 Hemispheric asymmetry?

The number processing capacities of both hemispheres were first investigated in visual half field (VHF) experiments. In these behavioural studies, participants were instructed to look at a central fixation cross while target stimuli were presented to the left or to the right of a fixation cross. Given the neuro-anatomical organization of the brain, stimuli presented in the left visual field (LVF) are firstly processed in the right hemisphere while stimuli presented in the right visual field (RVF) are firstly processed in the left hemisphere. VHF experiments are therefore an efficient tool to identify hemisphere specific processing mechanisms. Using this methodology, Ratinckx, Brysbaert and Reynvoet (2001) instructed their participants to classify a target Arabic number as small (1, 2) or large (5, 6). The target stimuli were presented in the right (RVF) or in the left visual field (LVF), simultaneously with a distractor in the opposite visual field. A comparison distance effect and a priming distance effect (i.e., responses were faster when the target and distractor were numerically close) were observed in both VFs, indicating that target and distractor were similarly processed in both hemispheres. The same conclusion was reached with a cross-notation design (Arabic digits as targets and written verbal numerals as distractor or inversely Arabic digits as distractor and written verbal numerals as target; Ratinckx & Brysbaert, 2002), when prime and target were simultaneously (Ratinckx & Brysbaert, 2002) or sequentially (Reynvoet & Ratinckx, 2004) presented in the same VF or when the prime was centrally presented (Notebaert & Reynvoet, 2009). Contrastingly, with two digit-numbers, a steeper priming curve was observed when targets were presented in the LFV compared to targets presented in the RVF (Notebaert & Reynvoet, 2009). So, while the first VHF studies failed to reveal any distinctive properties of the magnitude representation in the left and right hemispheres, the last study reported here suggested that the representation of numbers may be less precise in the RH than in the LH (Notebaert & Reynvoet, 2009).

Several previous experiments, using more recent experimental methods, lend some support to this last conclusion. In a developmental fMRI study, for example, Rivera and colleagues (Rivera, Reiss, Eckert, & Menon, 2005) found that the neural responses which sustain mental arithmetic change significantly with age. When asked to judge whether the result of an arithmetic operation was correct or incorrect, children showed brain activation in the prefrontal cortex, indicating that they relied on working memory and attentional resources to achieve the task. Older participants, on the contrary, showed more brain activation in the left supramarginal gyrus and in the left lateral occipito-temporal cortex, indicating a stronger reliance on automatic and language mediated processes (see Fig. 1-10) (see also Ansari, Garcia, Lucas, Hamon, & Dhital, 2005).

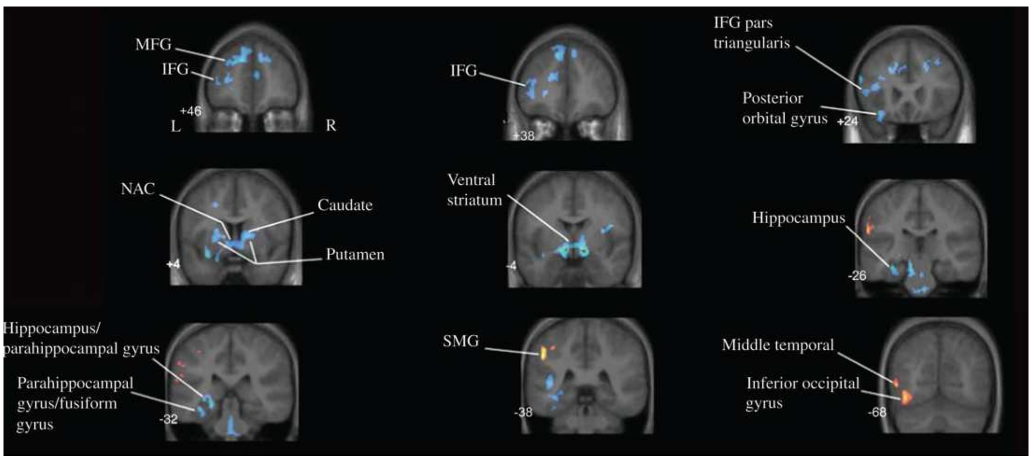

Fig. 1-10. Brain areas that showed significant increase (in red) and decrease (in blue) in activation with age in an arithmetic task. *Reproduced from Rivera et al. (2005).*

Brain activation during mental calculation was also found to be modulated by individual mathematical competences in adults. Indeed, individuals with higher mathematical competences were shown to display more activation in the left angular gyrus than those with lower mathematical competences (Grabner, Ansari, Reishofer, Stern, Hebner, & Neuper, 2007). Finally, the activation of the left and right hemispheres was shown to be dependent on the stimulus format used in a magnitude comparison task; the left angular gyrus being preferentially activated with numerical symbols ant the right superior parietal cortex being preferentially activated with non-symbolic stimuli (Holloway, Price, & Ansari, 2010).

The two hemispheres were also shown to be differentially involved in approximate and exact numerical judgments: approximate calculation leading to stronger activation in the right IPS and exact calculation involving stronger activation in the left IPS (Cohen & Dehaene, 1996; Dehaene & Cohen, 1991; Piazza et al., 2006). Besides, a transcranial magnetic stimulation experiment suggested that a disruption of the left parietal cortex was sufficient to produce deficits in discriminating close digits while a disruption of both parietal cortices was necessary to alter the discrimination of far digits (Andres, Seron, & Olivier, 2005). Finally, Piazza and colleagues demonstrated that the Weber fraction was smaller in the left than in the right hemisphere (Piazza et al., 2004), as was the neurons' tuning curve (Piazza et al., 2007; see section 3.2).

Another study providing evidence of notation-specific representations of numbers in the parietal cortex is the fMRI adaptation experiment of Cohen Kadosh, Cohen Kadosh, Kass, Henik, and Goebel (2007). In this study, the left IPS was found to adapt to quantity regardless of the format of the stimulus (i.e., Arabic digits, number words, mixed format). The right IPS, on the opposite, exhibited an adaptation effect with Arabic digits but not with number words. At first glance, these data may seem contradictory to the observations made by Piazza et al. (2007) according to which the left IPS may contain a more precise numerical representation than the right IPS. However, as suggested by Ansari (2007), both studies could be seen as converging on a central idea: the specialization of the left IPS for the representation of enculturated numerical symbols (see also Pinel & Dehaene, 2010).

4. Conclusions

The human understanding of numbers is rooted in our ability to make judgments about numerosity. Numerosity is an abstract property of a set, since it is independent of the sensory attributes of its members and of the physical parameters of the set, such as shape, luminance, density, duration or frequency. Despite its abstractness, the ability to represent and discriminate numerosities does not depend on learning a symbolic system, as it spontaneously emerges in pre-linguistic infants, and is observable in non-human species. At this stage, the processing of numbers is characterized by two psychophysical effects: distance and size effects. Following the recurrent observation of these two effects in various tasks and with different stimulus formats, it was postulated that numbers could be represented on a logarithmically-scaled mental number line. Although the mental number line model was supported by a formal neuronal network, recent experiments demonstrated that this shared mental representation could be reached through different pathways: **(1)** an accurate, place-coded and left-lateralized step for symbolic numbers and **(2)** an approximate, summation-coded and right-lateralized (or bilateral) step for non-symbolic quantities (Dehaene, 2009).

For the ease of our purposes and following the proposal of Verguts et al. (2005) and Siegler and collaborators (Booth & Siegler, 2006; Siegler & Booth, 2004; Siegler & Opfer, 2003), we will assume, in the following chapters of the present thesis, that the precision level of the two representations actually refers to different kind of scaling: the symbolic numerical representation being linearly scaled (i.e., accurate) and the non-symbolic numerical representation being logarithmically scaled (i.e., approximate).

CHAPTER 2

The numerosity estimation process

In the literature on numerical cognition, three different processes have been defined to account for the quantification of elements. These three processes are: subitizing, counting and estimation. Subitizing refers to rapid, effortless and accurate judgments of small sets of entities (up to three or four items) (e.g., Kaufman, Lord, Reese, & Volkmann, 1949). Number judgments for larger set-sizes involve either counting or estimating, depending on the number of elements presented within the display and the time given to respond. Counting is a verbal time-consuming process of enumeration that allows a precise evaluation of the cardinality of a set of elements (e.g., Gelman, 1978), while estimation is an approximate and rapid judgment of a set's numerosity. In this chapter, we will focus on the numerosity estimation process. We will first describe the tasks usually used to study this process and then we will detail the different hypotheses that have been put forward to account for the under- and overestimation phenomena often observed in these tasks.

1. Estimation tasks

Two kinds of tasks are usually used to study the numerosity estimation process: perception and production task, each of which is tied to a specific input and output code. In a typical perception task, non-symbolic stimuli, such as sets of different numerical size (e.g., collections of dots or sequences of sounds), are presented to participants, who have to estimate their numerosity by the use of symbolic outputs, such as oral verbal or Arabic numerals (e.g., Bevan & Turner, 1964; Kaufman et al., 1949; Krueger, 1972; Mandler & Shebo, 1982). In a typical production task, on the contrary, symbolic numerical stimuli, such as oral verbal or Arabic numerals, are presented to participants who have to approximately reproduce their magnitude by the production of a non-symbolic output, such as a sequence of key presses (e.g., Castronovo & Seron, 2007a; Whalen et al., 1999).

In both tasks, participants' patterns of performance present the signature of Weber's law, suggesting that numbers and numerosities are converted into analogue mental magnitudes on the mental number line (Dehaene, 1992, 1997; Gallistel & Gelman, 1992, 2000; Moyer & Landauer, 1967; Whalen et al., 1999): the means of estimates and their variability increase proportionally with target magnitude, resulting in constant coefficients of variation (i.e., standard deviation/mean) across target size (e.g., Castronovo & Seron, 2007a; Cordes et al., 2001; Whalen et al., 1999). Moreover, and in line with the triple code model, numerosity estimation could be defined as an activity requiring a semantic transcoding process: while perception tasks involve a non-symbolic to symbolic numerical mapping, production tasks involve a symbolic to non-symbolic numerical mapping.

Until now, this ability to map two different representations of numbers has only been examined in a few studies. In a recent experiment by Mundy and Gilmore (2009), for example, the mapping abilities of 6- and 8-year-old children were evaluated in both directions. In half the trials, a target Arabic number was presented followed by two sets of dots (symbolic to non-symbolic numerical mapping). In the other half, a target set of dots was presented followed by two alternative Arabic digits (non-symbolic to symbolic numerical mapping). The target quantities ranged from 20 to 50 and the alternative choices consisted of the correct response and a distractor differing from the target quantity by a ratio of .50 or .67. Children were instructed to map the target quantity to one of the two alternative choices. The results indicated that the ability to map two numerical representations was not symmetrical across direction of mapping: children were unable to map a symbol to a non-symbolic quantity when the ratio was .67, but they were able to do the mapping in the opposite direction at the same ratio. According to the authors, this asymmetry originated from the level of precision of each representation: whereas the symbolic to non-symbolic version of the mapping task involved one activation of the accurate symbolic representation (for the target Arabic digit) and two activations of the approximate non-symbolic representation (for the two alternative sets of dots), the non-symbolic to symbolic version of the task involved one activation of the approximate representation and two activations of the accurate numerical representation. Children were therefore more accurate in the non-symbolic to symbolic numerical mapping because this condition involved less approximation than the symbolic to non-symbolic numerical version of the task.

The ability to map the symbolic and non-symbolic representations of numbers was also studied in adults and was found to be disrupted after a frontal lobe damage. A patient with this kind of lesion was indeed shown to present difficulties in cognitive (i.e., the capacity to give approximate answers to questions of general semantic knowledge for which no precise answer is readily known; Shallice & Evans, 1978) and verbal numerosity estimation tasks (Revkin, Piazza, Izard, Zamarian, Karner, & Delazer, 2008): the patient’s responses fell systematically outside the range of answers given by the healthy participants. Interestingly, as the patient’s performance was comparable to healthy participants on all tasks evaluating the semantic representations of numbers (i.e., dots comparison, dots addition, digit comparison and number size stroop digit comparison), the patient’s deficits were assumed to reflect an impairment of the numerical mapping process rather than an impairment of the semantic representation of numbers itself. In this sense, this case study lends further support to the idea that a numerical mapping process is involved in numerosity estimation tasks. As we will see in the following section, this mapping process may nevertheless be influenced by different effects, therefore producing predictable under- and over-estimation errors.

2. Under- and over-estimation

2.1 <u>The perceptual effects</u>

There is evidence that numerical estimation may be based on factors other than numerosity per se. In dot pattern estimation tasks, for example, participants' performance was shown to be influenced by the Gestalt principles of good form, according to which sets of elements which constitute a good "Gestalt" should appear more numerous (Frith & Frith, 1972; Ginsburg, 1980; Ginsburg & Nicholls, 1988). These principles are: **(1)** *the clustering effect*: dot arrays are judged more numerous when they are spread over a large area than when they are more densely spaced (Krueger, 1972; Sophian & Chu, 2008); **(2)** *the regular-random numerosity illusion*: regular displays are perceived as being larger than randomly arranged displays (Ginsburg, 1978, 1980; Ginsburg & Pringle, 1988); **(3)** *the solitaire illusion*: a large array is judged more numerous than several small arrays (Frith & Frith, 1972; Ginsburg, 1982); **(4)** *the item size effect*: large items are thought to be more numerous than small items (Ginsburg & Nicholls, 1988; Miller & Baker, 1968).

In an attempt to take these numerosity illusions into account in a quantitatively measurable way, Allik and Tuulmets (1991) proposed an *area-based "occupancy" model*, according to which perceived numerosity (essentially in dot pattern estimation tasks) depends on the area occupied by the pattern to be processed. This total occupied area (i.e., the occupancy index) may be quantitatively modelled by surrounding each item with a circular territory of radius *R* centred at the dot and independent of the spatial proximity of dots. As a consequence, if two dots are close to each other, their occupancy territories overlap, resulting in a lower occupancy index (see Fig. 2-1). According to this model, perceived numerosity is therefore not directly linked to the number of dots included in the visual pattern, but to the size of the overlap of the dots' occupied territories.

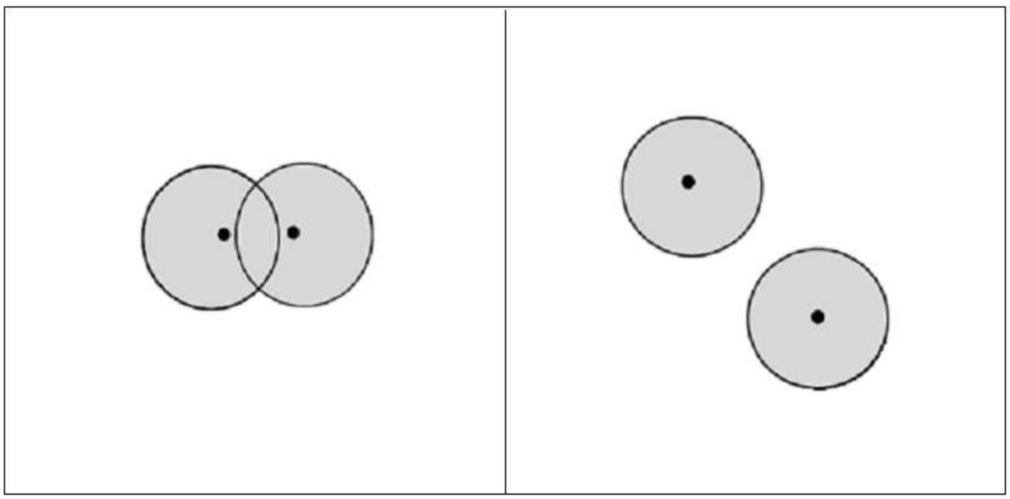

Fig. 2-1. Occupancy model (Allik & Tuulmets, 1991). When 2 dots are closer than 2 *R***,** their occupancy territories overlap (left image), resulting in a lower perceived numerosity than when 2 dots are separated by more than 2 *R* (right image). *Reproduced from Durgin (1995).*

Later, Durgin (1995) proposed that numerosity judgments were made by combining the density and area information of the array. Consistent with this idea, the author demonstrated that numerosity perception was affected by adaptation to a dense array and that this effect increased with increasing numerosity. More specifically, participants were adapted to a dense array of dots in one region of their visual field and then asked to compare the numerosity of two dot patterns, one presented in the adapted visual field, the other presented in the non-adapted field. The results demonstrated that the arrays presented in the adapted field were perceived as less numerous than the same arrays presented in the non-adapted field. This aftereffect is illustrated in Figure 2-2: when participants fixate their gaze on the fixation cross of Figure 2-2a, and then move their gaze to the fixation cross of Figure 2-2b, they have the impression that the right (adapted) display has the same numerosity than the left (non adapted) display. Interestingly, the effect of adaptation on numerosity is range dependent: for large numerosities, adaptation reduces the apparent number of elements presented in the adapted field (see Fig. 2-2). For small numerosities, adaptation does not substantially affect the apparent number of dots.

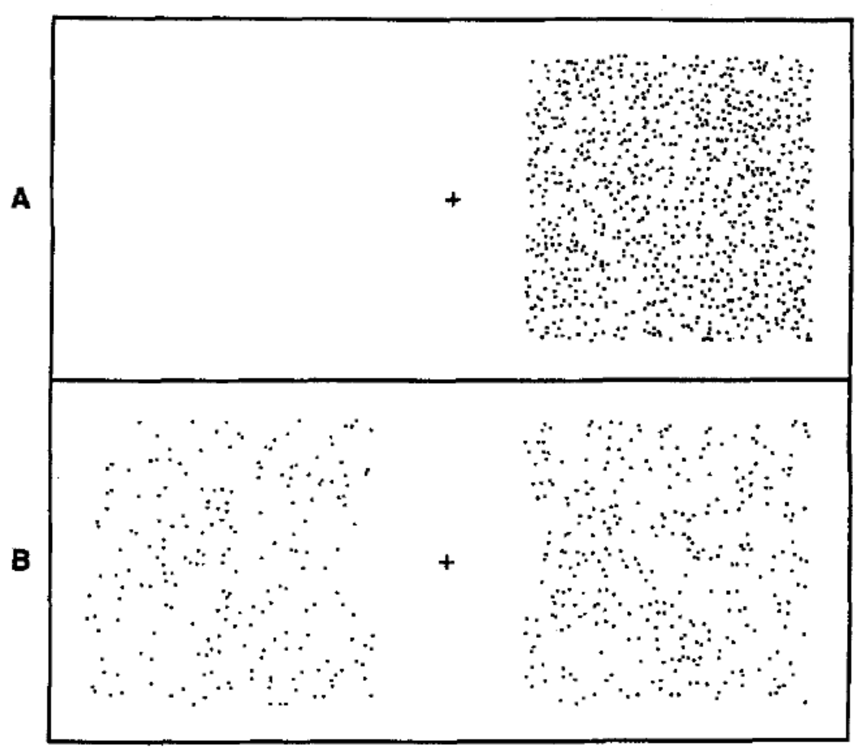

Fig. 2-2. Methodology of the experiment of Durgin (1995). (A) Adaptation stimulus. **(B)** Test numerosities that appear equal after adaptation to (A). The right display actually contains 320 dots while the left display actually contains 224 dots. *Adapted from Durgin (1995).*

Similarly, the perception of sets of dots has been shown to be distorted after participants had been adapted to sets of different numerosities. More particularly, sets containing fewer dots than the adaptation stimulus were perceived as even less numerous whereas sets containing more dots than the adaptation stimulus were perceived as even more numerous (Burr & Ross, 2008). As this effect was restricted to the location of the visual field where the adaptation stimulus had been presented, Dehaene (2009) suggested that it actually reflected the retinotopic property of the accumulation neurons (see Chapter 1, section 2.2.2).

So, although perceptual variables can account for some of the behavioural data on numerical estimation, it cannot be concluded that numerosity judgments are solely based on non-numerical parameters (Sophian & Chu, 2008). Rather, adults probably use a combination of both non-numerical and numerical cues in apprehending numerosities.

2.2 The regression effect

In addition to obeying Weber's law, numerosity perception and numerosity production tasks also present a major difference: numerosity is systematically underestimated in the numerosity perception task whereas it is systematically overestimated in the numerosity production task. This difference has been recently highlighted by Castronovo and Seron (2007a) in a study on numerical estimation but has also been found in studies involving the estimation of psychophysical stimuli, such as loudness, brightness, duration and distance. On one hand, when participants undertake psychophysical scaling tasks that require the attribution of numerical values to stimuli such as brightness (i.e., perception tasks), participants tend to underestimate their target magnitude. On the other hand, when participants have to produce the target magnitude of stimuli, such as particular durations, they tend to overestimate it (i.e., production tasks) (e.g., Stevens, 1953, 1955, 1966; Teghtsoonian, 1973; Teghtsoonian & Teghtsoonian, 1978). The major experimental finding from these psychophysical experiments is that the relation between the objective magnitude of a physical stimulus and its perceived magnitude (i.e., the subjective magnitude) can be conceptualized as a power function such as $\psi = k\ \phi^{n}$, in which ψ is the subjective magnitude's range of the sensation induced by the stimulus, k is a constant depending on the type and the context of the stimulation, ϕ is the range of the objective magnitude of the physical stimulus, and n is a power exponent whose the value depends on the perceptual continuum involved (Stevens, 1957). By investigating several types of perceptual continuous stimuli, Stevens and Greenbaum (1966) showed that perception tasks yield smaller power exponents than production tasks. According to the authors, this phenomenon is due to what they called the *regression effect* in psychophysical judgments. This effect corresponds to a tendency for the subjects to constrict their estimation of the variable under their control when performing the task (i.e., the output stimulus). In perception tasks, the variable under the participants' control is a quantitative continuum (i.e., the subjective magnitude - ψ), the objective magnitude of the stimuli (i.e., ϕ) being controlled by the experimenter. In production tasks, this variable is a physical continuum (i.e., the objective magnitude - ϕ), the subjective magnitude (i.e., ψ) being controlled by the experimenter. Because of the regression effect, when one continuum is matched to the other, two different power functions are obtained according to the adjusted variable on which the participant has control. Each of these functions is a straight line in a log-log plot ($\log \psi = n \log \phi + k$). However, in a perception task, the value of the exponent will be biased below its true value, yielding the subject to underestimate the stimulus magnitude. On the contrary, in a production task, the value of the exponent will be biased above its true value, yielding the subject to overestimate the stimulus intensity (see Fig. 2-3).

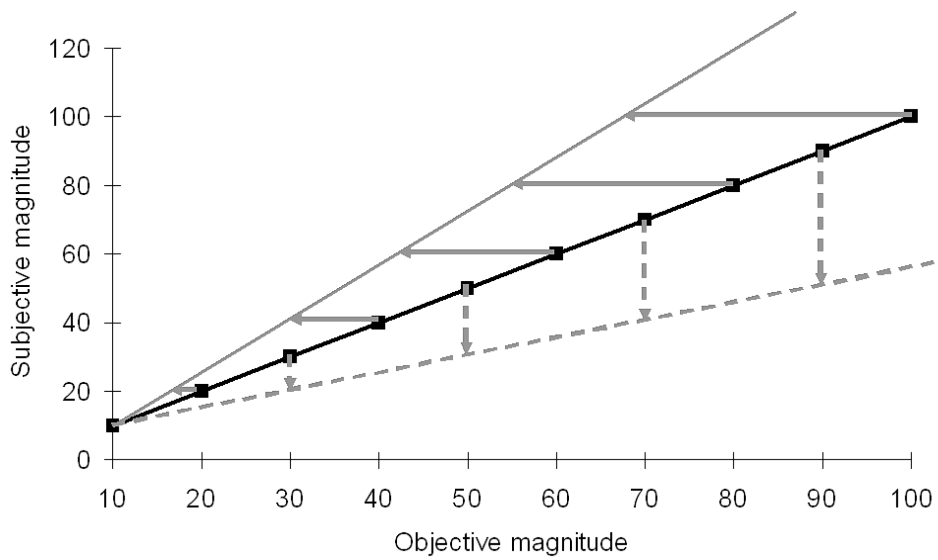

Fig. 2-3. Schematic illustration of the regression effect in psychophysical judgment tasks. According to Stevens and Greenbaum (1966), subjects constrict the range of their estimations on the variable that is under their control. In perception tasks (dotted grey lines), subjects control and thus constrict the variable plotted on the ordinate (i.e., subjective magnitude or ψ), whereas in production tasks (full grey lines), subjects control and thus constrict the variable plotted on the abscissa (i.e., objective magnitude or ϕ). In the former case, the exponent will be biased below its true value (yielding to underestimation errors). In the latter case, the exponent will be biased above its true value (yielding to overestimation errors).

Later, Teghtsoonian (1973) argued that ψ and ϕ cannot constrain a subject's estimation because they are defined only gradually in the course of the experiment. Therefore, according to this view, the regression effect concerns two successive stimulus and response ratios ($r\phi$ and $r\psi$) rather than the absolute range of these parameters (ψ and ϕ). In other words, according to Teghtsoonian (1973), under- and overestimation are due to the fact that subjects avoid giving two successive estimations that are in an extremely large or an extremely small ratio (log $r\psi = n$ log $r\phi + k$). However, no explanation was given regarding the origin of this effect.

In light of the literature on psychophysical and numerical estimation, it appears that the studies conducted in these two fields of research present some similarities. First, they both require participants to use either numerical values to rate their perception of an input stimulus (i.e., perception task) or to produce outputs corresponding to their estimation of the inputs' magnitude (i.e., production task). Second, both psychophysical and numerical estimation studies report that participants' responses vary as a power function of the input magnitude (Indow & Ida, 1977; Krueger, 1972, 1982, 1984, 1989) and that the value of the power function exponent depends on the tasks used: perception tasks lead to a flatter function with a lower exponent, while production tasks lead to a steeper function with a bigger exponent. The regression effect hypothesis could therefore throw some light on our understanding of numerical under- and overestimation. However, this theoretical proposal cannot account for the fact that systematic estimation errors occur even when subjects make their first estimations (Krueger, 1982).

To encompass this issue, a new hypothesis will be proposed in Chapter 3 (the bi-directional mapping hypothesis), according to which under- and overestimation actually reflect the differential scaling of the symbolic and non-symbolic representations of numbers and the direction of the mapping process that takes place between these two representations (Castronovo & Seron, 2007a).

2.3 The momentum effect

Under- and overestimation were also observed in approximate arithmetic tasks. In their preliminary study, McCrink and colleagues (McCrink, Dehaene, & Dehaene-

Lambertz, 2007) showed adults videos of events involving addition or subtraction and required them to judge whether or not different final numerosities were the correct outcome of the operation. Overall, participants appeared to accept as correct a well-defined range of plausible answers, depending on the arithmetic operation that had to be carried out. Although the responses to small problems tended to be centred on the true outcome, as problem size increased, they tended to be biased in the direction of larger numbers for addition and of smaller numbers for subtraction (see Fig. 2-4).

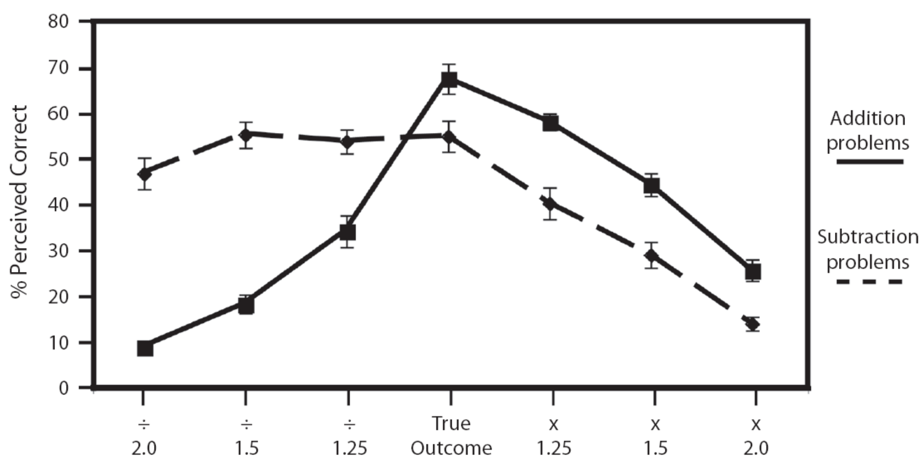

Fig. 2-4. Operational momentum in adults. Participants are more likely to judge as correct an overestimated addition outcome and an underestimated subtraction outcome. *Reproduced from McCrink et al. (2007).*

The so-called operational momentum effect was observed to a lesser extent with Arabic numerals (Knops, Viarouge, & Dehaene, 2009) and in children (McCrink & Wynn, 2009). In this last study, 9-month-old infants were presented videos of rectangles being added or subtracted from one another. In the addition condition, 4 rectangles were serially added to a set of 6 rectangles. In the subtraction condition, 4 rectangles were serially removed from a set of 14 rectangles. Then, three different outcomes were presented to the children: an incorrect outcome of 5 items, a correct outcome of 10 items and an incorrect outcome of 20 items. Interestingly, infants' looking times were significantly longer when the outcome presented violated the momentum of the arithmetic operation (i.e., outcome numerically larger in the subtraction condition and outcome numerically smaller in the addition condition) (see Fig. 2-5).

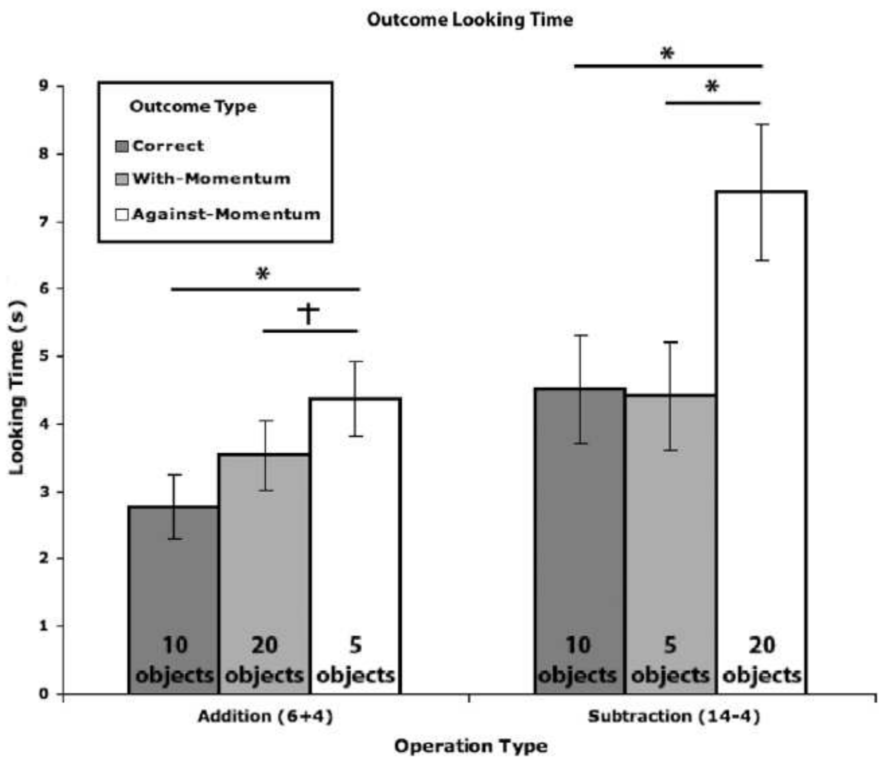

Fig. 2-5. Operational momentum in infants. 9 month-olds looked significantly longer at outcomes that violated the momentum of the arithmetic operation. *Reproduced from McCrink & Wynn (2009).*

The operational momentum was also observed in a study requiring speeded pointing to arithmetic results (Pinhas & Fisher, 2008). In this task, symbolic addition and subtraction problems were presented to participants who were required to compute the result of the operation before locating it on a visually presented number line. Consistently with the previous experiments, pointing was biased leftward after a subtraction problem and rightward after an addition problem.

According to McCrink et al. (2007), the observation of these systematic biases is consistent with the view that numerical operations involve movements on an internal mental number line (Dehaene & Cohen, 1991; Gallistel & Gelman, 1992; Hubbard, Piazza, Pinel, & Dehaene, 2005; Restle, 1970): as small numbers are represented on the left of this continuum and large numbers on the right, addition problems involve a rightward displacement while subtraction problems yield a leftward displacement. In line with this logic, cortical circuits for spatial attention were found to contribute to mental arithmetic in adults (Knops, Thirion, Hubbard, Michel, Dehaene, 2009). The operational momentum hypothesis therefore assumes that under- and overestimation are observed because participants are moving too far on their mental number line while performing approximate arithmetic operations. In other words, under- and overestimation are observed because participants shift their locus of attention along the number line in the direction of the operation (toward larger numbers for addition and toward smaller numbers for subtraction).

As participants have to produce their numerosity production judgments in an iterative way (e.g., production of successive key presses; Castronovo & Seron, 2007a; Cordes et al., 2001; Whalen et al., 1999), it is possible that the numerosity production tasks actually involve a left-to-right displacement along the mental number line (the more the participants produce successive responses, the more they move towards the right of their mental number line). So, the overestimation found in such tasks could be due to a kind of operational momentum rather than to symbolic to non-symbolic numerical mapping. This question will be addressed in Chapter 4.

2.4 The calibration effect

Several pieces of research have demonstrated that it is possible to calibrate the non-symbolic to symbolic numerical mapping. In a dot pattern estimation task, for example, performance was shown to be considerably improved after the participants were told how many dots a particular array contained. This feedback effect was also shown to be strong as it was still present 8 months after the initial testing (Minturn & Reese, 1951). In the same way, Izard and Dehaene (2008) demonstrated that inserting a reference trial was sufficient to improve participants' estimates of the whole range of numerosities presented. In this study, participants were instructed to estimate patterns of dots before (Experiment 1) and after (Experiment 2) a reference trial had been presented. Participants were told that this reference trial contained 30 dots. However, depending on the experimental condition, it actually contained either 25 dots (overestimated inducer), 30 dots (exact inducer) or 39 dots (underestimated inducer). Whereas numerosities were underestimated in Experiment 1, participants' responses fitted the different inducers with a high degree of precision in Experiment 2: the mean stimulus numerosity leading to the response "30" were 27.4, 31.5 and 40.1 in the overestimated, exact and underestimated inducer conditions, respectively. To account for these results, the authors proposed a *numerosity labelling model* according to which the mental number line is divided into several segments, each of them being associated with a distinct verbal label. An optimal strategy to give a numerosity judgment would be to choose the number which is closest to the activation point on the mental number line (i.e., canonical response grid). Therefore, if the perception of a numerosity activates the position "log(40)" on the mental number line, the optimal strategy would be to give the response "40". However, Izard and Dehaene (2008) demonstrated that participants spontaneously use an idiosyncratic affine transformation of this canonical response grid. In other words, they inaccurately determine the amount of activation corresponding to the zero level and the value of one unit on the number line so that the position associated with the response “*n*” is no longer *log(n)* but *a.log(n)+b*. The use of such a response grid can explain why participants underestimated numerosity in Experiment 1. In order to account for the calibration effect observed in Experiment 2, Izard and Dehaene (2008) added a supplementary postulate to their theory: the reference trial does not modify the internal Weber fraction (i.e., the sensitivity of the non-symbolic numerical representation), but induces participants to apply an affine transformation to their response grid. According to the numerosity labelling theory, then, the calibration effect is well captured by a change in the parameters *a* and *b* of the equation *a.log(n)+b*. Because this change implies a global transformation, the idea of an affine transformation very well captures Izard and Dehaene’s experimental results (2008) according to which all numerosities are calibrated at once.

To date, however, we do not know whether the feedback and response format could have an impact on the transformation applied to the stimulus-response association. In the same way as under- and overestimation could reflect the differential scaling of the symbolic and non-symbolic numerical representations, calibration could indeed reflect the differential coding of these representations. In other words, it is possible that calibration occurs at every numerosity, but with a different strength depending on the feedback and response format and on the distance separating the feedback and the target magnitudes. This issue will be the subject of Chapter 5.

3. Conclusions and summary of the experimental questions

Visual estimation is a pervasive process that is modulated by some spatial non-numerical parameters of the stimuli and that has been studied by using two kinds of tasks: perception and production tasks. Although both tasks involve patterns of performance obeying Weber's law, a major difference has been repeatedly observed: participants systematically underestimate the numerosity in perception tasks (Bevan & Turner, 1964; Castronovo & Seron, 2007a; Kaufman et al., 1949; Krueger, 1972; Mandler & Shebo, 1982), whereas they systematically overestimate it in production tasks (Castronovo & Seron, 2007a; Cordes et al., 2001; Whalen et al., 1999). These opposite patterns of performance have also been observed with psychophysical stimuli such as brightness. Until now, however, only one hypothesis has been put forward to account for these phenomena: the regression effect (Stevens & Greenbaum, 1966; Teghtsoonian, 1973) according to which under- and overestimation are observed because individuals tend to use a shortened range of the variable under their control, whether it is the number continuum as in perception tasks, or the physical (e.g., brightness) continuum as in the production tasks. However, such a tendency to "regress" toward the mean value of a serie of controlled stimuli cannot account for the fact that under- and overestimation are observed as soon as participants give their first estimation (Krueger, 1982). A new hypothesis (the bi-directional mapping hypothesis) will therefore be examined in the third Chapter of the present thesis. This hypothesis relies on the assumption that under- and overestimation actually reflect the differential scaling of the symbolic and non-symbolic representations of numbers and the direction of the mapping process that takes place between these two representations.

Under-and overestimation were also observed in approximate arithmetic tasks, but, in this case, these reverse patterns of performance were interpreted as reflecting a leftward or rightward displacement along the mental number line (the momentum hypothesis; McCrink et al., 2007). As a rightward displacement may also occur in the numerosity production task, the fourth chapter of the present thesis will examine whether the momentum hypothesis could be considered as an alternative interpretation to the overestimation error observed in numerosity production tasks.

Finally, it has been found that the general tendency to underestimate numerosity in a numerosity perception task may be calibrated if symbolic feedback is provided to participants (Izard & Dehaene, 2008; Minturn & Reese, 1951). In Chapter 5, we will examine whether the format used to present the feedback trial and the participants' responses may influence the shape of the calibration effect.

CHAPTER 3

Under- and overestimation: a bi-directional mapping process between symbolic and non-symbolic representations of number?[1]

Abstract

Over the last 30 years, numerical estimation has been largely studied. Recently, Castronovo and Seron (2007a) proposed the bi-directional mapping hypothesis in order to account for the finding that dependent on the type of estimation task (perception vs. production of numerosities), reverse patterns of performance are found (i.e., under- and overestimation respectively). Here, we further investigated this hypothesis by submitting adult participants to three types of numerical estimation task: **(1)** *a perception task*, in which participants had to estimate the numerosity of a non-symbolic collection; **(2)** *a production task*, in which participants had to approximately produce the numerosity of a symbolic numerical input; **(3)** *a reproduction task*, in which participants had to reproduce the numerosity of a non-symbolic numerical input. Our results gave further support to the finding that different patterns of performance are found according to the type of estimation task: **(1)** underestimation in the perception task; **(2)** overestimation in the production task; **(3)** accurate estimation in the reproduction task. Moreover, correlation analyses revealed that the more a participant underestimated in the perception task, the more he/she overestimated in the production task. We discussed these empirical data by showing how they can be accounted for by the bi-directional mapping hypothesis (Castronovo & Seron, 2007a).

[1] This section is a modified version of an article with an identical title by Crollen, Castronovo and Seron, which has been published in Experimental Psychology, 2011, 58(1): 39-49. In order to avoid redundancy with the preceding sections, the introduction has been shortened.

1. Introduction

The bi-directional mapping hypothesis was recently proposed by Castronovo and Seron (2007a) to account for the under- and overestimation found in numerosity perception and production tasks. This hypothesis relies on three classical assumptions: **(1)** the existence of several numerical representations; **(2)** the existence of transcoding routes between these representations; **(3)** the existence of difference in precision between the different numerical representations.

Based on these three assumptions, the bi-directional mapping hypothesis assumes that the opposite patterns of performance found in the numerical perception and production tasks (i.e., under- vs. overestimation) are the result of transcoding activities between the analogue magnitude representation and its corresponding symbolic representations (i.e., verbal numerals or Arabic numerals). In a perception task, the numerical process goes from a position on the subjective number line logarithmically compressed to its corresponding symbolic linear representation. As the objective magnitude is always smaller than the subjective one, participants under-estimate the numerosity presented (see Fig. 3-1a). Inversely, in a production task, the numerical process goes from a position on the symbolic linear numerical representation to its corresponding magnitude on the subjective number line logarithmically compressed. As the subjective magnitude is bigger than the objective one, participants over-estimate the target magnitude (see Fig. 3-1b).

While a large body of research has tried to conceptualize the general laws of numerical processing (e.g., the scalar variability principle, the power-shaped distribution of the response function, the compressive scaling of the mental number line), little has been done to study the mapping ability between the symbolic and non-symbolic representations of numbers. The researchers who examined these particular mapping abilities essentially did it in only one direction: the non-symbolic to symbolic mapping (Izard & Dehaene, 2008). Moreover, as the bi-directional mapping hypothesis has been proposed a posteriori by Castronovo & Seron (2007a) to account for some of their results, it certainly needs to be supported by further empirical evidence. Therefore, the current study has three major aims: **(1)** directly address the question of adults' mapping ability in both directions; **(2)** investigate whether the direction of the mapping process has an impact on participants' performance in numerical estimation tasks (i.e., under- and overestimation); **(3)** examine the validity of the bi-directional mapping hypothesis to account for the under- and overestimation errors observed in numerosity estimation tasks.

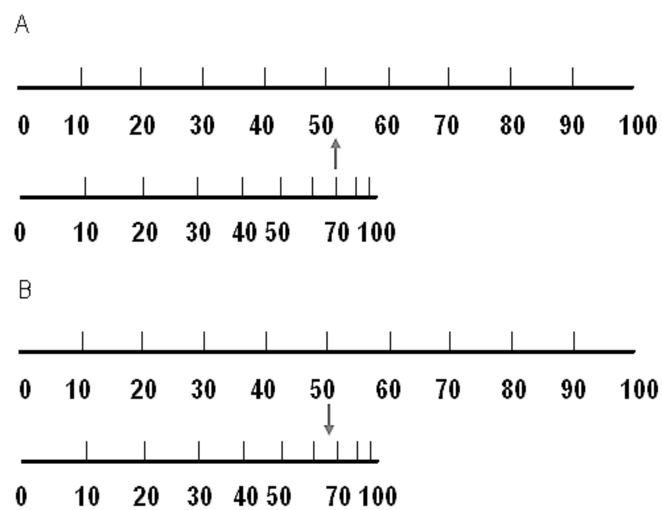

Fig. 3-1. Mapping processes shown by the grey arrows: **(A)** in a perception task, the numerical process starts from a position on the non-symbolic numerical representation which is logarithmic and goes to a position on the linear symbolic numerical representation, which involves underestimation; **(B)** in a production task, the numerical process goes from a position on the symbolic numerical representation to a position on the non-symbolic numerical representation, which involves overestimation.

To this end, the same participants were tested in three different numerical estimation experiments, which involved the same stimuli (i.e., set of dots and Arabic numerals) and response modalities (i.e., use of a potentiometer). Experiment 1 used a non-symbolic to symbolic numerosity perception task that required the estimation of the number of dots presented within a display. Experiment 2 used a symbolic to non-symbolic production task in which participants had to produce dot patterns corresponding to the numerosity of a symbolic Arabic numeral. Finally, Experiment 3 employed a non-symbolic to non-symbolic reproduction task, in which participants had to reproduce, through the production of a set of dots, the numerosity of a non-symbolic input (set of dots). Moreover, as a large body of research demonstrated that estimation may be based on factors other than numerosity per se (e.g., Allik & Tuulmets, 1991; Durgin, 1995; Frith & Frith, 1972; Ginsburg, 1980; Krueger, 1972), the dot patterns were controlled for low-level continuous variables (i.e., luminance, total area, density and dot size).

If, as postulated by the bi-directional mapping hypothesis, the direction of the mapping process between the symbolic and non-symbolic numerical representations has an impact on the estimation patterns of performance, then : **(1)** in the perception task, participants should underestimate the numerosity of the non-symbolic targets; **(2)** in the production task, participants should overestimate the number of dots to produce in order to reach the magnitude of the symbolic target; **(3)** in the reproduction task, more accurate estimations should be found. Indeed, this particularly original estimation task, allowing a greater investigation of the bi-directional mapping hypothesis, could involve: either **(1)** the absence of a mapping process, as only the non-symbolic numerical representation could be activated to solve the task (the non-symbolic input activated a position on the mental number line, which is directly reactivated at the response production stage); or **(2)** a bi-directional mapping with a symbolic mediation (the non-symbolic input is implicitly tagged by a symbolic label which serves, in turn, as a basis for the response production). In both situations, accurate estimates should be found. Indeed, in the first case, no mapping would take place, therefore no systematic bias in the estimates should be found. In the second case, the two mappings would go in opposite directions; as a consequence, they should cancel each other out.

Finally, as the same participants were involved in both types of numerical estimation tasks and following the assumption that there are individual differences in the mental number line acuity (Halberda et al., 2008), participants with more precise "number sense" should show less under- and overestimation, compared to participants with less precise number sense, in both the perception and the production tasks.

2. Experiment 1: numerosity perception Experiment

2.1 Participants

The participants were 15 volunteers (4 males, 11 females; 13 right-handed, 2 left-handed), aged between 19 and 31 (M = 26.3 ± 3.8). They participated in all three experiments, which were conducted in consecutive sessions on different days. Eight participants performed first the perception task and then the reproduction task, whereas the remaining participants performed these two tasks in reverse order. All 15 participants were submitted to the production task in the last experimental session in order to avoid participants gaining information on the experimental number range used.

2.2 Stimuli and procedure

Sixteen different numerosities were used to develop the target sets of dots: [21, 24, 27, 30, 33, 36, 39, 42, 49, 56, 63, 70, 77, 84, 91, 98]. The numerical range started at 21, as previous studies showed that numerical estimates are usually accurate for numerosities smaller than 20 (Indow & Ida, 1977; Krueger, 1982, 1984). Numerosities larger than 21 were chosen by applying a ratio going from .87 to .93, as this ratio corresponds to the discrimination threshold in human adults (Halberda & Feigenson, 2008; Pica et al., 2004; van Oeffelen & Vos, 1982).

Low-level continuous perceptive variables, such as total occupied area, luminance, dot size and density, were controlled in the non-symbolic input stimuli (see Dehaene, Izard & Piazza, 2005), by the introduction of two different sets of stimuli: the Extensive Set and the Intensive set. In the Extensive Set, the sum of the area of all the dots (the luminance) and the total occupied area on the screen were kept constant across numerosity. As a consequence, the density of the array increased with increasing numerosity and the size of the dots decreased with increasing numerosity. In the Intensive Set, the size of the dots and the density of the pattern (i.e., free area around each dot, comprised between 2 and 3 times the dot's radius) were kept constant. As a consequence, the luminance and the total occupied area increased with increasing numerosity. The dots were black on a white background (see Fig. 3-2).

The experiment was conducted on a PC computer. Each trial began with the presentation of a fixation cross (1000 ms). Then, a pattern of dots was flashed on the screen for a duration of 250 ms. After the presentation of the pattern, the sign "=" (500 ms) was displayed on the screen, followed by the Arabic numeral "1", indicating to the participants that they had to start their estimation. To give their answers, participants had to go through the Arabic numerals sequence by turning a potentiometer[2] on a response box. By turning the potentiometer to the right or to the left, participants were able to go through the Arabic numerals sequence forwards or backwards. The response box allowed the production of all the Arabic numerals ranging from 1 to 254. Participants were asked to report their completion by pressing a button on the same response box. Moreover, in order to avoid memory limitations during response production, participants were encouraged to decide on the Arabic numeral to reach before starting turning the potentiometer.

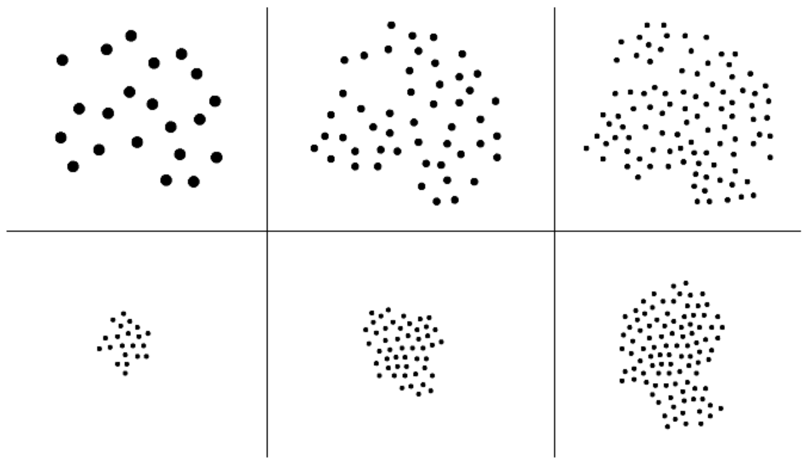

Fig. 3-2. Examples of stimuli belonging to the Extensive (up) and intensive (down) sets.

Each participant completed eight practice trials, followed by eight blocks of 32 experimental trials each. Within a block, each one of the 16 target numerosities of the Extensive set and of the Intensive set was randomly presented once. That is, half of the trials (N = 128) were drawn from the Intensive set and the other half from the Extensive set. Block order was counterbalanced across participants such that eight

[2] The response box measures a relative position (it has no beginning, no end), contrary to a classic potentiometer which only provides a measure of an absolute position. The use of such a “relative” potentiometer does not allow participants to use absolute location references to solve the task.

participants undertook Block 1 to Block 8 and the other seven participants Block 8 to Block 1. Participants received no information about the range of numerosities to be presented and no feedback regarding their performance.

2.3 Results

Data for the trials in which a participant's response was three standard deviations above the participant's overall mean response were excluded from the analysis (1.5% of the data). Moreover, as preliminary analyses showed that the mean estimates and their standard deviations were not normally distributed, the following analyses were conducted on their logarithmic transformation.

Linear mixed models (LMM)[3] with numerosity as a fixed effect and subjects as a random effect were first conducted to determine whether participants' patterns of performance obeyed Weber's law (i.e., constant CVs). The results showed that the log of the mean estimates increased with target magnitude, $F(15,224) = 40.17$, $p < .001$, as did the log of the standard deviation, $F(15,224) = 12.26$, $p < .001$. The coefficients of variation were then computed (CV = standard deviation/mean) and analyzed using a LMM with the same fixed and random effects as previously. The results showed that the CVs were constant across target size (i.e., signature of Weber's law), as they did not significantly vary with numerical size, $p > .9$. Moreover, as the numerosity perception task presents similarities to psychophysical perception experiments, estimates were expected to form a power function of numerosity. If the response function was actually power-shaped ($\psi = k\phi^{n}$), then the log(response) function should be a linear function of the log(numerosity) (log $\psi = n$

[3] The linear mixed model was used because: a) it is an extension of the general linear model but does not require observations to be independent with constant variance; b) it presents the advantage to allow the specification of random (subjects, in our experiments) and fixed effects (numerosity, in our experiments) and thus avoids averaging across all dimensions.

log $\phi + k$). The results showed that this log-log regression analysis fitted the data very well (R^2 = .99, p < .001). Furthermore, the exponent of the power function was smaller than 1 (β = .84) indicating that responses were not linear but shaped as a power function.

Finally, to see whether the participants' responses were accurate, underestimated or overestimated, t-tests with a test value of zero were conducted on the error rate (ER), which was defined as follows: ER = (participant's response – target magnitude)/target magnitude). Accordingly, an ER of zero indicates accurate estimation, a negative ER reflects underestimation of magnitudes, and a positive ER overestimation of magnitudes. As a preliminary paired samples t-test showed that the ER was equivalent in the Extensive (M = −.38 ± .15) and Intensive sets (M = − .37 ± .15), p > .3, data from the two sets were merged. The mean ER (for the whole set of numerosities) was −.38, which was significantly different from zero, $t(15) = -29.30$, p < .001. As shown in Figure 3-3, the underestimation error was observed for each target magnitude (all p_s < .001) and increased with increasing target magnitude, $F(15, 210) = 14.21$, p < .001. Moreover, an analogous t-test of ER for practice trials showed that participants systematically underestimated the first item (Krueger, 1982), $t(14) = -4.19$, p = .001 (M = −.40 ± .37).

2.4 Discussion

Altogether, these first results replicated patterns of performance previously found in studies on numerical estimation: **(1)** both the means and the standard deviations of the numerical judgments increased with the target magnitudes and in direct proportion. This resulted in constant coefficients of variation across target size (Castronovo & Seron, 2007a; Cordes et al., 2001; Whalen et al., 1999), and indicated that the scalar variability principle well described our data; **(2)** participants' responses varied as a power function of the input stimulus (Indow & Ida, 1977; Krueger, 1972, 1982, 1984, 1989); **(3)** participants underestimated magnitudes in a numerical estimation task involving the perception of a numerosity and a non-symbolic to symbolic numerical mapping (e.g., Izard & Dehaene, 2008); **(4)** the underestimation error increased with increasing target magnitudes and occurred as soon as participants made their first numerical judgment (Krueger, 1982).

In order to examine whether the opposite mapping process (i.e., from a symbolic input to a non-symbolic output) would lead the same participants to overestimate the same target magnitudes, as predicted by the bi-directional mapping hypothesis, a second experiment was conducted. This second experiment was a production task requiring participants to produce a pattern of dots corresponding to the magnitude of a target Arabic numeral.

3. Experiment 2: numerosity production Experiment

3.1 Stimuli and procedure

In the numerosity production task, the numerical range (i.e., 21, 24, 27, 30, 33, 36, 39, 42, 49, 56, 63, 70, 77, 84, 91, 98) and materials (i.e., Extensive and Intensive patterns of dots) used were the same as in the numerosity perception task of Experiment 1.

Each trial began with the presentation of a fixation cross (1000 ms), then, the target Arabic numeral appeared on the screen for a duration of 1000 ms. Afterwards, the sign "=" was presented (500 ms), followed by a single dot, indicating to the participants that they had to start the dots' production by turning the potentiometer (to the right to increase the number of dots appearing on the screen, and to the left to decrease the number of dots). The response box allowed participants to produce all the numerosities ranging from 1 to 254. However, participants were encouraged to stop the dots' production as soon as they believed to have reached the target numerosity. Participants were also encouraged to use as few back and forth movements as possible to get to their final response. Half of the pattern of dots produced by participants was initiated from the Extensive set and the other half from the Intensive set in order to control perceptive variables. In order to avoid the use of an overt or covert counting strategy, the dots appeared quickly on the screen (creation of one dot every 1.4° of angular rotation) and participants were explicitly instructed not to count the number of dots they produced. They were also required to perform the task as fast as possible and had to report their completion by pressing a response button on the response box.

The presentation and number of practice and experimental trials was identical to that in Experiment 1. Also, no information about numerosity and no performance feedback were provided.

3.2 Results

Data for the trials in which a participant's response was three standard deviations above the participant's overall mean response were excluded from the analysis (1.3% of the data). Moreover, as preliminary analyses demonstrated that the mean estimates and their standard deviations were not normally distributed, the following analyses were conducted on their logarithmic transformation.

As previously, LMM (with numerosity as a fixed effect and subjects as a random effect) were used to see whether participants' patterns of performance obeyed Weber's law. The results showed that the log of the mean response increased with the target magnitude, $F(15, 224) = 56.96$, $p < .001$, as did the log of the standard deviation, $F(15, 224) = 39.37$, $p < .001$. Moreover, these two measures appeared to increase in direct proportion, as constant CVs were found across target size (i.e., signature of Weber's law, $p > .2$). In psychophysical production experiments, the shape of the response function is generally power shaped. A regression analysis showed that this property extended to the present data: the log(response) function was actually a linear function of the log(numerosity) ($R^2 = .99$, $p < .001$) and the exponent ($\beta = 1.08$) indicated that the larger the target size, the larger the response.

Then, to see whether the participants' responses were accurate, overestimated or underestimated, t-tests with a test value of zero were conducted on the error rate. As in the perception task, data from the Extensive and Intensive stimuli sets were merged as ER appeared to be equivalent in these two sets ($M = .52 \pm .35$ in the Extensive Set; $M = .55 \pm .45$ in the Intensive Set, $p > .6$). The mean ER for the whole set of numerosities was .52, which was significantly different from zero, $t(15) = 26.61$, $p < .001$. As shown in Figure 3-3, participants' overestimation error was found for each target number (all $p_s < .001$) and increased with increasing magnitude, $F(15, 210) = 2.21$, $p < .01$. Moreover, an additional t-test analysis showed for practice trials that the overestimation error occurred as soon as participants make their first numerical judgment ($M = .38 \pm .50$, $t(14) = 2.99$, $p = .01$).

3.3 Discussion

Most of the results of Experiment 1 were replicated in this Production Experiment. Firstly, both the means and the standard deviations of the numerical estimates increased with the target magnitude and in direct proportion, so that the scalar variability principle described the participants' patterns of performance.

Secondly, participants' responses were power-shaped. Thirdly, the estimation error increased with increasing target magnitude.

The only difference found between Experiments 1 and 2 was that participants' responses were, here, overestimated instead of being underestimated as in the perception task. This last finding strengthens the bi-directional mapping hypothesis' postulate according to which under- and overestimation depends on the direction of the mapping process that takes place between two differently scaled numerical representations. In order to further investigate this postulate, a third experiment involving the activation of the non-symbolic numerical representation in input as in output was conducted. In this third task, participants had to produce a dot pattern having the same magnitude as a target pattern of dots.

4. Experiment 3: numerosity reproduction Experiment

4.1 Stimuli and procedure

In the present numerosity reproduction task, the numerical range and materials used were the same as in the two previous experiments. Each trial began with the presentation of a fixation cross (500 ms). Then, a dot pattern was flashed on the screen for a duration of 250 ms. Following the target presentation, the sign "=" (500 ms) was presented. Then a single dot appeared on the screen, indicating to the participants that they had to start their response production, using the same procedure as in Experiment 2 except for the following: for half of the trials, the input stimuli were from the Extensive set and induced the production of an Intensive output. For the other half of the trials, the input stimuli were from the Intensive set and induced the production of an Extensive output. To reproduce approximately, "by feel", new patterns of dots having the same magnitude as the target dot pattern, the potentiometer had to be used like in Experiment 2.

The presentation and number of practice and experimental trials was identical to that in Experiments 1 and 2. Also, no information about numerosity and no performance feedback were provided.

4.2 Results

Data for the trials in which a participant's response was three standard deviations greater than the participant's overall mean response were excluded from the analysis (1% of the data). Again, preliminary analyses showed that the mean estimates and their standard deviations were not normally distributed, therefore, LMM (with numerosity as a fixed effect and subjects as a random effect) were performed on the logarithmic transformation of these measures in order to look for the signature of Weber's law in participants' patterns of performance. In this task, the log of the mean estimates increased with the target magnitude, $F(15,224) = 79.26$, $p < .001$, as did the log of the standard deviation, $F(15,224) = 31.46$, $p < .001$. Moreover, the mean number of dots produced and their standard deviation increased in direct proportion, given constant CVs across target numerosity, $p > .5$ (i.e., signature of Weber's law). Then, to see whether the response function of the task was a power function, a regression analysis was run on the log(numerosity) and log(response). As in the previous experiments, this log-log regression fitted the data perfectly ($R^2 = .99$, $p < .001$). Moreover, the exponent of this regression was smaller than 1 ($\beta = .76$), indicating that responses were shaped as a power function and were not linear.

Finally, to see whether the participants' responses were accurate, underestimated or overestimated, t-tests with a test value of zero were conducted on the error rate. As the ER was equivalent in the Extensive (M = .12 ± .22) and Intensive sets (M = .04 ± .31), p > .4, data from the two stimuli sets were merged. The mean ER (for the whole set of numerosities) was .08, which was significantly different from zero (t(15) = 2.30, p < .05). However, as shown in Figure 3-3, participants' responses were only overestimated for the smallest numerosities from 21 to 36, ($p_s \leq .01$), but accurate for all the other numerosities from 39 to 98 ($p_s > .05$). A repeated measures ANOVA with numerosity as a within subject factor showed that the ER decreased with increasing magnitude, $F(15, 210) = 37.73$, $p < .001$, and an independent sample t-test with a test value of zero conducted on the first practice trial showed that no systematic under- or overestimation error occurs on participants' first estimations (M = −.18 ± .45, p > .1).

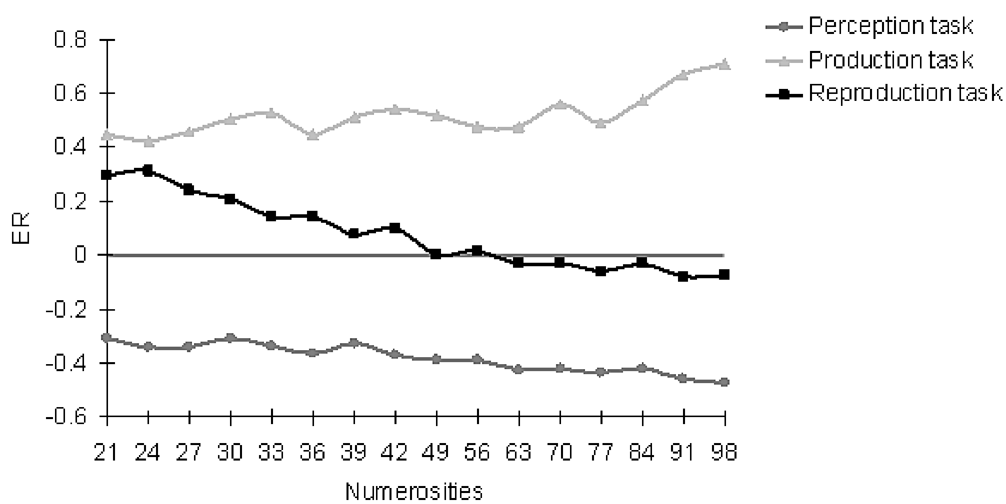

Fig. 3-3. Mean ER for each target magnitude of the numerosity perception, production and reproduction tasks.

4.3 Discussion

Experiment 3 reflects perfectly the access to the semantic non-symbolic numerical representation to perform the task. First, scalar variability principle accounted for the participants' pattern of performance. Second, the response distribution of the task was a power function of the input stimuli.

Moreover, although the mean ER for the whole set of numerosities was significantly greater than 0 (M = .08), responses were nevertheless more accurate in the reproduction task than in the two previous experiments (M = −.38 in the perception task; M = .52 in the production task), both p_s <.001. This original finding gave further support to the bi-directional mapping hypothesis and is therefore further analyzed in the next section of this paper.

5. Comparison between the tasks

As the same group of participants took part in the three numerosity estimation tasks, several analyses were conducted to directly compare the tasks. Firstly, we conducted a repeated measures ANOVA on the ER with the task (Perception vs. Production vs. Reproduction) as a within-subject factor. This analysis confirmed that participants showed different patterns of performance according to the task, $F(2,28) = 45.57$, $p < .001$: they underestimated the target numerosities in the perception task, overestimated them in the production task and had better estimations in the reproduction task (see Figure 3-3). Secondly, a Pearson parametric correlation analysis was conducted to investigate whether the participants' ERs were correlated, especially in the perception and production tasks. The results showed that the more a participant underestimated the target magnitudes in the perception task, the more he/she overestimated similar target magnitudes in the production task, $r(15) = −.53$, $p = .02$ (one-tailed) (see Fig. 3-4a). Thirdly, for each participant, the perception and production response means were submitted to separate linear regressions with numerosity. Each of the 30 equations computed represented the best description for a particular participant of the relation between the response mean and the numerosity variable. A Pearson parametric correlation analysis was then conducted to examine whether the regression coefficients (the slopes of the regression analyses) of the perception and production tasks were correlated. Again, this analysis suggested that the more a participant underestimated target magnitudes in the perception task (the more the slope of the regression analysis was smaller than 1), the more he/she overestimated target magnitudes in the production task (the more the slope of the regression analysis was bigger than 1), $r(15) = -.57$, $p = .01$ (one-tailed) (see Fig. 3-4b).

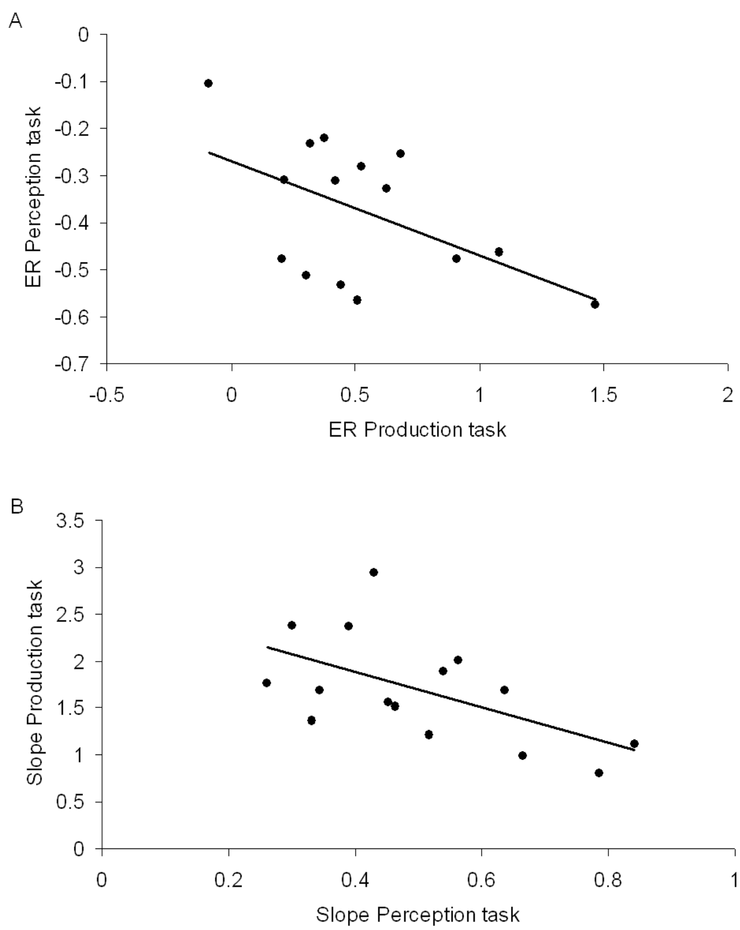

Fig. 3-4. (**A**) Correlation between the error rates of each participant in the numerosity perception and production tasks; (**B**) Correlation between the regression slopes of each participant in the numerosity perception and production tasks.

6. General discussion

Numerical estimation constitutes one of the three main numerical quantification processes with subitizing and counting. However, there is a gap in the literature regarding the study of that particular process, as most of the time it has been investigated in a single direction with perception tasks involving a mapping process from a non-symbolic collection to a symbolic numerical label (e.g., Izard & Dehaene, 2008). Therefore, the aim of our study was to examine whether the direction of the mapping process could have an effect on adults' numerosity estimation performances, as predicted by the bi-directional mapping hypothesis (Castronovo & Seron, 2007a). In order to do so, 15 participants were submitted to three numerosity estimation experiments, each of them involving a different mapping: **(1)** a numerosity perception task requiring participants to estimate the number of dots presented within a display (i.e., non-symbolic to symbolic mapping); **(2)** a numerosity production task requiring participants to produce the number of dots corresponding to a target Arabic numeral (i.e., symbolic to non-symbolic mapping); **(3)** a numerosity reproduction task involving the non-symbolic reproduction of an input dot pattern (i.e., no mapping or a bi-directional mapping).

As expected, different patterns of performance were found according to the task used: underestimation in the perception task; overestimation in the production task; more accurate estimations in the reproduction task. To our knowledge, this last task has never been reported in the literature but has the peculiarity to allow a further reinforcement of the idea that it is the mapping process and its direction that account for the error patterns found in the perception and production tasks. In the same way, the high correlation observed between participants' degree of under- and overestimation is also a strong and original argument in favor of the bi-directional mapping hypothesis. As it is generally assumed that the precision of the mental number line is not the same across individuals (see Halberda et al., 2008) and, within the framework of the bi-directional mapping hypothesis, it was expected that a participant with an “imprecise” mental number line under- and overestimate numerosities in the same proportion in a perception and a production task, and to a greater extent than a participant with a more precise number line. This observation that the greatest “underestimators” in the perception task were also the greatest “overestimators” in the production task may reflect the fact that the experience individuals have with numerical information has an impact on their numerical estimation abilities: the greater their experience, the better their performance (Castronovo & Seron, 2007a; Lipton & Spelke, 2005; Siegler & Opfer, 2003; Verguts et al., 2005).

The results of our three experiments also confirmed that the participants' behavior followed two major and well-established psychophysical laws: **(1)** participants' patterns of performance were well described by the scalar variability principle (i.e., signature of Weber's law) (e.g., Castronovo & Seron, 2007a; Cordes et al., 2001, Whalen et al., 1999); **(2)** the response distribution in the three tasks corresponded to a power-function of the input magnitude (Indow & Ida, 1977; Krueger, 1972, 1982, 1984, 1989).

Therefore, in an attempt to integrate all of our observations in a more general psychophysical background based on the simpler form of the power function $\psi = k\ \phi^{n}$ (Stevens, 1957), we postulated that the value of the exponent *n* might be affected, in numerosity estimation tasks, by: **(1)** the inter-task numerical mapping differences (i.e., exponent lower than 1 in tasks involving a non-symbolic to symbolic mapping and exponent greater than 1 in tasks involving a symbolic to non-symbolic mapping); **(2)** individual differences in the precision of the mental number line (i.e., lower power exponent for participants with imprecise numerical representation). In other words, the psychophysical function of numerosity estimation might result from the integration of several differently weighted types of factor. This idea remains of course speculative and needs to be further investigated in future research.

To conclude, to our knowledge, the present study is the first one to directly investigate in adults different types of numerical estimation tasks, involving different mapping routes and their impact on the patterns of performances found. Our results provide further evidence to the fact that numerical estimation tasks present regular and predictable behavioral data: the signature of Weber's law, the power-shaped distribution of the mean response function and, more importantly, different patterns of performance according to the type of numerical estimation used and the direction of the mapping it involves: underestimation in perception tasks (non-symbolic to symbolic numerical mapping), overestimation in production tasks (symbolic to non-symbolic numerical mapping), more accurate estimations in reproduction tasks (no mapping or a bi-directional mapping process). Our study also showed that the recent bi-directional mapping hypothesis accounts well for these results. Future studies should further investigate this model by exploring other psychophysical dimensions such as time, space, weight, brightness, etc. We already know that Weber's law applies to the discrimination of these dimensions, as it does to the discrimination of numerosities. So, it is likely that the same under- and overestimation phenomenon occurs whenever we need to map continuous quantities and discrete measures.